

Analysis of stable water isotopes in tropospheric moisture during the West African Monsoon

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Physik

des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Christopher Johannes Diekmann

aus Darmstadt (Hessen)

Tag der mündlichen Prüfung:

30. April 2021

Erster Gutachter:

Prof. Dr. Peter Braesicke

Zweiter Gutachter:

Prof. Dr. Peter Knippertz

I did not expect to find that water could behave in such a different manner under so similar conditions. Once again this small, simple molecule amazes and intrigues us with its magic.

*Mpemba effect from a viewpoint of an
experimental physical chemist*

NIKOLA BREGOVIC

Abstract

The West African Monsoon (WAM¹) is one of the most prominent atmospheric features of the West African climate. It represents the main source of rainfall over the semi-arid Sahel zone and thus has a crucial impact on the socio-economics of these regions, which primarily rely on agriculture. However, the complex interaction of large-scale circulation and microphysical rain and cloud processes poses a substantial challenge to reliably quantify the atmospheric branches of the hydrological cycle during the WAM.

This thesis aims to promote the understanding of the variability of tropospheric water vapor and its control mechanisms over the Sahel during the monsoon period. This is achieved by jointly studying abundances of the stable water isotopologues H₂O and HDO (denoted as standardized ratio product δD), as these contain fundamental information about moisture pathways and processes.

In a three-step concept, this thesis exploits observational and model-based {H₂O, δD } pair distributions for identifying effects of transport and moisture processes associated with the WAM. The first step includes the generation of a novel, global and daily dataset of tropospheric {H₂O, δD } pairs based on remote sensing measurements from the satellite sensor Metop/IASI. For making data from isotope-enabled models comparable to these remotely sensed results, a so-called retrieval simulator is developed. As second step, a model-based framework using Lagrangian trajectory information is built for allowing meaningful interpretations of {H₂O, δD } pairs over the Sahel. Based on this framework, the third step performs an extensive and multi-scale analysis of IASI {H₂O, δD } pairs over the Sahel. This includes various simulations from the isotope-enabled models ICON-ART_{iso} and COSMO_{iso} as well as additional rainfall data from GPM IMERG. This synthesis allows to identify characteristic {H₂O, δD } signals during the WAM and to attribute those to clear effects of air mass mixing and microphysical processes such as condensation, evaporation and equilibration of rain drops.

Overall, this study reveals the potential of using {H₂O, δD } pair data from IASI together with high-resolution modeling and process attribution methods for investigating the tropospheric moisture budget and its dominant control mechanisms. Such a joint analysis holds great promise for evaluating and improving numerical weather and climate models in the long run. Ultimately, this thesis underlines the general potential of isotopic analyses to study the hydrological cycle.

¹Further abbreviations are explained in Appendix C.

Zusammenfassung

Der westafrikanische Monsun ist einer der markantesten Bestandteile des Klimas über Westafrika. Er ist die Hauptquelle für Regen über der semi-ariden Sahelzone und hat damit einen entscheidenden Einfluss auf die Sozioökonomie dieser Regionen, die größtenteils auf Landwirtschaft angewiesen sind. Allerdings stellt das komplexe Zusammenspiel der großräumigen Zirkulation mit kleinskaligen Wolken- und Regenprozessen eine große Herausforderung dar, die atmosphärischen Zweige des Wasserkreislaufs während des Monsuns verlässlich zu quantifizieren.

Der Fokus der zugrundeliegenden Arbeit liegt daher auf der Analyse von troposphärischem Wasserdampf und seiner Einflussfaktoren während des westafrikanischen Monsuns. Für diesen Zweck werden Verteilungen der Wasserdampfisotopologie H_2O and HDO (standardmäßig als δD angegeben) gemeinsam betrachtet, da diese fundamentale Rückschlüsse über Feuchtepfade und -prozesse ermöglichen.

In einem dreistufigen Konzept nutzt diese Arbeit beobachtungs- und modellgestützte $\{\text{H}_2\text{O}, \delta\text{D}\}$ Verteilungen zur Analyse von Effekten von Transport- und Feuchteprozessen im Zusammenhang mit dem Monsun. Im ersten Schritt wird ein neuartiger Datensatz von troposphärischen $\{\text{H}_2\text{O}, \delta\text{D}\}$ -Paaren auf der Basis von globalen and täglichen Fernerkundungsdaten des Satellitensensors Metop/IASI erzeugt. Um Isotopendaten aus Modellanalysen mit diesem neuen Datensatz vergleichbar zu machen, wird ein sogenannter Retrieval Simulator entwickelt. Der zweite Schritt zielt auf eine modellgestützte Prozessanalyse unter Verwendung von Luftmassen-Trajektorien, um eine aussagekräftige Interpretation von $\{\text{H}_2\text{O}, \delta\text{D}\}$ -Verteilungen zu ermöglichen. Darauf aufbauend wird im dritten Schritt eine umfangreiche und mehrskalige Analyse der IASI $\{\text{H}_2\text{O}, \delta\text{D}\}$ -Verteilungen über der Sahelzone mit Fokus auf den Monsun durchgeführt. Dies beinhaltet Simulationen mit den numerischen Wettermodellen $\text{ICON-ART}_{\text{iso}}$ und $\text{COSMO}_{\text{iso}}$ sowie Niederschlagsbeobachtungen von der NASA (GPM IMERG). Die Synthese dieser Datensätze und Methoden ermöglicht, charakteristische $\{\text{H}_2\text{O}, \delta\text{D}\}$ -Signale über der Sahelzone zu identifizieren und diese auf eindeutige Effekte der Luftmassenmischung und mikrophysikalischen Prozesse wie Kondensation, Verdunstung und Äquilibration von Regentropfen zurückzuführen.

Insgesamt verdeutlicht diese Studie das Potential von $\{\text{H}_2\text{O}, \delta\text{D}\}$ -Verteilungen von IASI zusammen mit hoch aufgelösten Modellrechnungen sowie detaillierten Prozessanalysen zur Untersuchung des troposphärischen Feuchtebudgets. Damit unterstreicht diese Arbeit den grundsätzlichen Wert der Wasserdampfisotopologie für die Analyse des hydrologischen Kreislaufs.

Contents

Abstract	iii
Zusammenfassung	v
I. Introduction	1
1. Introduction	3
II. Theoretical Background	7
2. Characterization of the West African Monsoon	9
2.1. General Meteorology	9
2.2. Dominant Moisture Transport Pathways	11
2.3. Structure of Squall Lines	13
2.4. The Monsoon Season 2016	14
3. Stable Water Isotopologues in the Troposphere	17
3.1. Theory for Interpreting Paired $\{H_2O, \delta D\}$ Distributions	18
3.1.1. Air Mass Mixing	19
3.1.2. Isotopic Fractionation	20
3.1.3. The Rayleigh Model	21
3.1.4. The Super-Rayleigh Regime	21
3.1.5. Paired Interpretation Using Frequency Contours	22
3.2. Tropospheric $\{H_2O, \delta D\}$ Pair Datasets	23
3.2.1. Remote Sensing Products	24
3.2.2. Isotope-Enabled Modeling	26
3.3. Isotopic Characterization of the West African Monsoon	29
3.4. Thesis Objectives	30
3.4.1. Research Goals	30
3.4.2. Overview of Used Datasets	32

III. Analyses	37
4. MUSICA IASI {H₂O, δD} Pair Processing	39
4.1. Main Sensor Characteristics	40
4.2. General Retrieval Processing	42
4.2.1. Preselection of Observations	42
4.2.2. PROFFIT-nadir Retrieval	44
4.2.3. Vertical Representativeness of Retrieval Results	47
4.3. Optimal Estimation {H ₂ O, δD} Pair Product	48
4.3.1. Retrieval of H ₂ O and δD	48
4.3.2. Post-Processing for Optimal {H ₂ O, δD} Pair Information	49
4.3.3. Spatio-temporal Data Characterization	59
4.4. Retrieval Simulations for {H ₂ O, δD} Pairs	67
4.4.1. Methodology	68
4.4.2. Evaluation of Retrieval Simulations	73
4.5. Summary	78
5. Lagrangian Process Attribution of {H₂O, δD} Pairs over West Africa	79
5.1. Data and Methodology	80
5.1.1. Trajectory Sorting	80
5.1.2. Trajectory-Based Process Attribution of {H ₂ O, δD} Pairs	81
5.2. Average {H ₂ O, δD} Development Along the West African Monsoon	81
5.3. Isotopic Process Attribution Along Single Trajectories	82
5.4. Isotopic Process Attribution Along Trajectory Ensembles	85
5.4.1. Identification of Transport Clusters	85
5.4.2. Mixing Processes	89
5.4.3. Rayleigh Processes	90
5.4.4. Super-Rayleigh Processes	92
5.5. Understanding {H ₂ O, δD} Pairs over the Sahel	95
5.6. Summary	97
6. Analysis of {H₂O, δD} Pair Data During the West African Monsoon	99
6.1. Annual Cycle over West Africa	100
6.1.1. Remote Sensing-Based Analysis	100
6.1.2. Model-Based Analysis	104
6.2. Seasonal Monsoon Development	109
6.2.1. Preparing the Model Data	110
6.2.2. Analysis of Monsoon Season 2016	111

6.3. Convective-Scale Variability	117
6.3.1. Clustering Approach	117
6.3.2. Clustering of Monsoon Season 2016	118
6.4. Summary	120
IV. Conclusions	123
7. Conclusions	125
7.1. Research Highlights	125
7.2. Perspectives	127
V. Appendix	129
A. Dust Analysis of {H₂O, δD} Pairs over Tenerife	131
B. Yearly {H₂O, δD} Pair Data and Precipitation Distributions	135
C. List of Acronyms	145
Bibliography	147

Part I.

Introduction

1 | Introduction

The meteorology and hydrology over West Africa is mainly governed by the complex West African Monsoon (WAM¹) system (Fink et al., 2017). Over the semi-arid Sahel zone, it causes a tremendous annual cycle from very dry winter months to a wet and intense rainy season, which has enormous socio-economic impacts on rural areas (Sultan et al., 2005, Berntell et al., 2018). From a climatological perspective, changes in the WAM system are critical for the degree of desiccation (Hulme, 2001) and the greenhouse effect (Spencer and Braswell, 1997) of West Africa. In addition, the WAM system is also found to affect weather systems in remote areas through dynamical teleconnections, such as over the North Atlantic and Europe (Bielli et al., 2010, Gaetani et al., 2011, Pante and Knippertz, 2019). However, the lack of observational networks (Parker et al., 2008) as well as the poor performance of numerical weather and climate models (Roehrig et al., 2013, Vogel et al., 2018) hamper a detailed analysis of the still poorly understood tropospheric moisture budget over the Sahel.

During the last decades, the analysis of stable water isotopologues in atmospheric water vapor and precipitation has been established as a powerful tool for investigating atmospheric moisture pathways. As each water isotopologue (hereafter referred to as water isotope) is associated with characteristic binding energies and diffusivities, the ratios of different isotopes are altered during phase changes. Therefore, the paired analysis of the light water isotope H₂O against its heavier HDO (given as standardized ratio δD) provides strong potential for evaluating effects of tropospheric moisture processes (e.g. Worden et al., 2007, Noone et al., 2011, Noone, 2012, Bolot et al., 2013, González et al., 2016, Schneider et al., 2016, Lacour et al., 2018). So far only few studies have investigated tropospheric distributions of stable water isotopes during the WAM. Observational and model-based studies underlined the strong influence of large-scale dynamics and microphysical processes on the isotopic composition of rain and water vapor (Risi et al., 2008b, 2010a, Tremoy et al., 2012, 2014). However, Risi et al. (2010b) concluded that quantifying the individual effects of the different moisture processes remains a major challenge.

The aim of the thesis is to advance the understanding of the variability of tropospheric moisture during the WAM. For the first time, paired distributions of H₂O and δD (referred to as {H₂O, δD } pairs) in Sahelian water vapor will be used to systematically investigate the hydrometeorological

¹Abbreviations are explained in Appendix C.

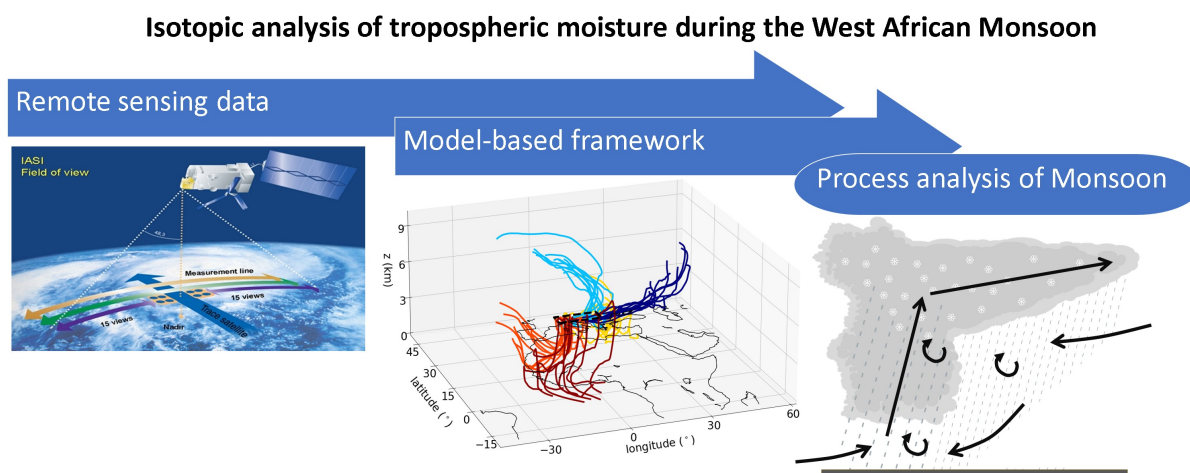


Figure 1.1. Illustrative scheme of the three-step concept, as followed throughout the thesis and described in Sect. 1.

control mechanisms of the WAM. This serves to further highlight the general value and potential of $\{H_2O, \delta D\}$ pair distributions for studying tropospheric moisture pathways.

This research goal is pursued within this thesis in a three-step concept (Fig. 1.1). As first step, the lack of continuous and dense water vapor measurements over West Africa is filled by generating a novel and comprehensive dataset of observational $\{H_2O, \delta D\}$ pair data for the mid-troposphere. For this purpose, radiance measurements from the thermal infrared sensor IASI are used, which offers a global scanning of the Earth's atmosphere multiple times per day. Optimal $\{H_2O, \delta D\}$ pair information is achieved by post-processing the individual H_2O and δD retrieval results for the aforementioned period, which are provided within Schneider et al. (2021c) and developed during the MUSICA project (Schneider et al., 2016). Additionally, a so-called retrieval simulator is developed and evaluated. Its purpose is to harmonize data from isotope-enabled models according to the vertical characteristics of the MUSICA IASI $\{H_2O, \delta D\}$ pair product, which is crucial for meaningful direct model-to-satellite comparisons.

The second step develops a model-based framework for improving the interpretation of observed and modeled variability in the paired $\{H_2O, \delta D\}$ phase space. The basis for this is the analysis of Lagrangian backward trajectories that indicate the pathway of an air parcel through the atmosphere. The approach is to track the isotopic evolution along these trajectories and to attribute characteristic $\{H_2O, \delta D\}$ signals to idealized effects of mixing and microphysical cloud and rain processes (Noone, 2012). This allows to interpret the Sahelian moisture by means of moisture processes that have occurred during the history of observed air masses.

Finally, as third step, the $\{H_2O, \delta D\}$ pair data from IASI as well as from the isotope-enabled models ICON-ART_{iso} and COSMO_{iso} are brought together in order to perform a multi-scale process analysis of the WAM. The aforementioned trajectory results in addition to global satellite-

based rainfall measurements from GPM IMERG support the attribution of the observed $\{H_2O, \delta D\}$ pair signals over the Sahel to underlying dynamical and microphysical processes. This allows to use the isotopic perspective for evaluating the impact of different model configurations to the simulated representation of the WAM system, such as horizontal resolution and convection treatment.

This work is structured as follows: Chapter 2 provides an overview on the WAM development and mechanisms on a climatological perspective as well as for the monsoon season 2016. This specific season is characterized in detail during the field campaign DACCIWA and will therefore serve here as a target period for detailed monsoon studies. Chapter 3 discusses the general physics and tropospheric datasets of stable water isotopes and concludes with emphasizing the research goals of this thesis (Sect. 3.3). The three-step analysis discussed above is realized in Chap. 4, which includes the description of the full MUSICA IASI $\{H_2O, \delta D\}$ pair processing, Chap. 5, which demonstrates the Lagrangian process attribution exemplarily for the WAM season 2016 and Chap. 6, where the WAM is analyzed on the annual, seasonal and convective scale. Chapter 7 concludes with a final summary of the main research highlights and discusses future perspectives.

Part II.

Theoretical Background

2 | **Characterization of the West African Monsoon**

This section provides an overview of the West African Monsoon system. The focus herein lies on its general meteorological features, the dominant transport patterns that affect the tropospheric moisture budget over the Sahel as well as on convective squall line systems. Furthermore, a particular focus will be set on the monsoon season 2016, which will be investigated in detail in the context of this work.

2.1. General Meteorology

During the monsoon period between June and September, the meteorology and hydrology of West Africa is dominated by the complex West African Monsoon (WAM) system. While the sun migrates smoothly northwards and shifts the zonal belt of tropical convection from the eastern Tropical Atlantic towards the West African coast, there is typically in the beginning of the boreal summer an abrupt jump of maximum precipitation from the Gulf of Guinea to the semi-arid Sahel zone (Sultan and Janicot, 2003, Hagos and Cook, 2007). The shift of convective activity up to 15–20° N marks the onset of the WAM and initiates intense precipitation events over the Sahel throughout the whole monsoon period.

This phenomenon is mainly driven by contrasts in near-surface temperature, pressure and moisture between the eastern Tropical Atlantic and the Sahara. Equatorial upwelling of cold water leads to relatively low sea surface temperatures, which is accompanied by tropospheric subsidence (Atlantic Cold Tongue (ACT); Caniaux et al., 2011, Druyan and Fulakeza, 2014). The main counterpart of the ACT is the so-called Saharan Heat Low (SHL), a lower tropospheric thermal depression in the Saharan desert west of 10° E, where high solar insolation and low surface evaporation lead to an intense surface heating (Lavaysse et al., 2009). As the SHL deepens, it further intensifies the thermal contrast between the Sahara and the eastern Tropical Atlantic and generates a weak cyclonic circulation near the Saharan surface, which is linked to dry convection (see arrow I in Fig. 2.1a). This circulation is a key driver for the northward shift of the monsoon precipitation and analogously the so-called Intertropical Discontinuity (ITD). The latter marks the northernmost position of the monsoon precipitation (Fitzpatrick et al.,

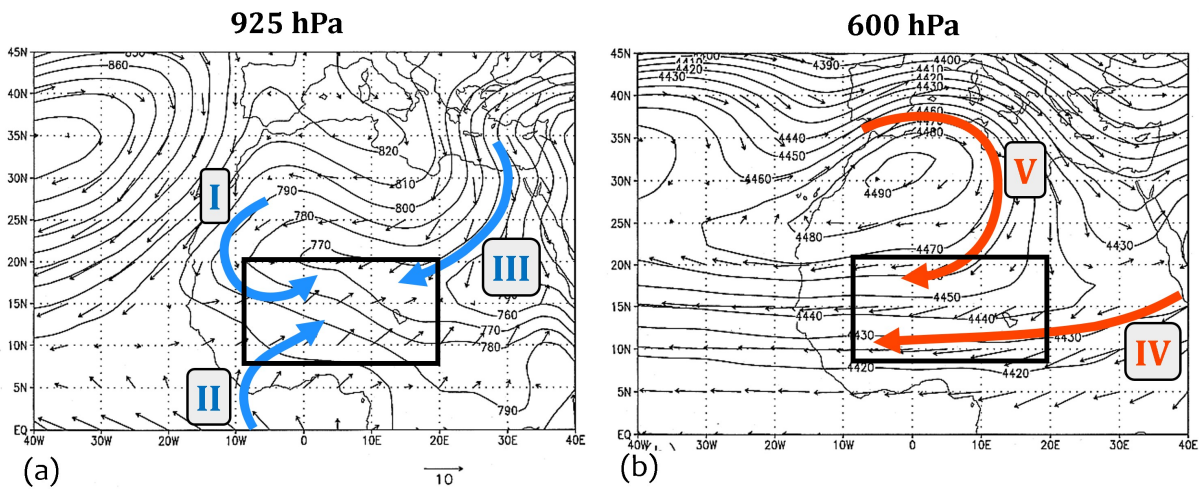


Figure 2.1. Overview of the dominant dynamical structures during the West African Monsoon (figure taken and adapted from Fig. 4 in Cook, 1999). The thin black arrows indicate the total wind vectors (the reference length of the black arrow represents 10 m s^{-1}) and the contour lines indicate the geopotential height (contour interval of 10 gpm), based on a reanalysis climatology for July at (a) 925 hPa and (b) 600 hPa. The large colored arrows point towards the cyclonic circulation due to the Saharan Heat Low (I, blue), the southwesterly monsoon flow (II, blue), the dry northeasterly Harmattan winds (III, blue), the African Easterly Jet (IV, red) and the anticyclonic circulation linked to the Saharan high (V, red). The black box covers the Sahel region, where the monsoon convection is strongest.

2015), as it represents the near-surface confluence of the moist southwesterly monsoon flow (arrow II in Fig. 2.1a) against the dry Harmattan, the dry northeasterly trade winds originating from the Mediterranean Sea (arrow III in Fig. 2.1a; Sultan and Janicot, 2003, Tyrlis and Lelieveld, 2013). Accordingly, the ITD is characterized by large gradients in humidity, temperature, vertical stability and dust content. As response to the SHL development, the ITD shifts from its winter position near the Guinea Coast to the Sahel with maximum latitudes up to 20° N (Lélé and Lamb, 2010).

The strong meridional temperature (and moisture) gradients in the lower troposphere cause a marked thermal wind at around 600 hPa and 15° N (arrow IV in Fig. 2.1b Cook, 1999). This so-called African Easterly Jet (AEJ) is a prominent zonal wind feature that dominates the mid-troposphere over West Africa during the monsoon period. Its intensity and altitude is strongly influenced by deep moist convection around the ITD and by dry convection within the SHL (Thorncroft and Blackburn, 1999). In turn, the AEJ is associated with horizontal and vertical wind shear and thereby creates an environment, which favours the development of so-called African Easterly Waves through barotropic and baroclinic instabilities (AEW; e.g. Thorncroft and Hoskins, 1994, Thorncroft et al., 2008, Leroux and Hall, 2009). AEWs are westward propagating

synoptic-scale wave disturbances in the lower troposphere of northern Africa with a period of 3–5 days (Carlson, 1969, Burpee, 1972). They are found to be in close interaction with organized deep convection over Central Africa (Berry and Thorncroft, 2005, Mekonnen et al., 2006, Hsieh and Cook, 2008, Berry and Thorncroft, 2012) and modulate the West African precipitation by fostering the development of highly organized, long-lived Mesoscale Convective Systems (MCS; e.g. Gu et al., 2004, Fink and Reiner, 2003, Skinner and Diffenbaugh, 2013, Crétat et al., 2015). Over West Africa, more than 40 % of all MCSs are related to AEWs (Fink and Reiner, 2003). This emphasizes the complex coupling of the mid-tropospheric AEJ with the development of the Sahelian precipitation (Sy et al., 2018, Niang et al., 2020). Further, the low-level SHL is associated with a mid-tropospheric high pressure system leading to a divergent flow within an anticyclonic circulation over North Africa (arrow V in Fig. 2.1b; Chen, 2005). This circulation feeds extra-tropical air masses into the Sahel, thereby it is found to contribute to the maintenance of the AEJ and to modulate its intensity (Thorncroft and Blackburn, 1999, Niang et al., 2020).

2.2. Dominant Moisture Transport Pathways

The complex mechanisms of the WAM are associated with distinct transport pathways, which lead to high levels of air mass mixing in the Sahelian troposphere (e.g. Niang et al., 2020). These large-scale atmospheric features of the WAM system are found to be in close connection with the precipitation and low-level moisture over West Africa (Lélé and Leslie, 2016). Water vapor can be transported from remote areas along the pathways marked in Fig. 2.1, which either moisten the otherwise dry Sahel directly, or lead to a recycling of the transported water vapor through rainout and continental evapotranspiration over West Africa. The latter is particularly essential for determining the low-level moisture over the Sahel (e.g. Cadet and Nnoli, 1987, Gong and Eltahir, 1996, Fontaine et al., 2003, Koster et al., 2004, Nieto et al., 2006). The following paragraphs provide an overview of studies that address the large-scale moisture transport along the pathways from Fig. 2.1. A more detailed and systematic moisture analysis of the individual pathways is performed in Chap. 5.

The main sources of water vapor in the Sahelian troposphere are represented by the near-surface features marked in Fig. 2.1a. A simple yet powerful method for estimating robust source-sink relationships of atmospheric moisture is the analysis of net change due to the combined effects of evaporation and precipitation (Stohl and James, 2004). For instance, during the monsoon period, when the southwesterlies reach the coastal zones of West Africa, the dominant source of Sahelian moisture is the eastern Tropical Atlantic (Cadet and Nnoli, 1987, Fontaine et al., 2003). The subsidence associated with the cold surface waters of the ACT suppress the convective activity over the eastern Tropical Atlantic, such that the monsoon southwesterlies can get largely moistened through local evaporation over the Gulf of Guinea and transport the water vapor

towards the Guinea Coast. As response to the SHL, this moisture can then be advected to the Sahel zone (Thorncroft et al., 2011). A further moisture source for the monsoon system is the North Atlantic (e.g. Fontaine et al., 2003, Grams et al., 2010). Nieto et al. (2006) identified a band of evaporation along the North Atlantic from the Sahel latitudes up to the Iberian coasts. The northerly trade winds transporting the evaporated water over the North Atlantic are then pulled towards the Sahel by the cyclonic circulation of the SHL. Another evaporative source of lower tropospheric moisture over the Sahel is the Mediterranean Sea (e.g. Fontaine et al., 2003, Nieto et al., 2006). Here, the increase of sea surface temperatures is found to be positively correlated with the Sahelian moisture and precipitation, which is in agreement with the fact that warmer sea surfaces lead to enhanced local evaporation (Rowell, 2003). The evaporated water from the Mediterranean Sea is then transported by the northeasterly Harmattan trade winds towards the Sahel without experiencing any significant precipitation.

In contrast, the impact of the mid- and upper-tropospheric features marked in Fig. 2.1b to the Sahelian moisture is not as clear as for the low-level transport pathways. For instance, the AEJ appears rather dry, as it remains constantly around 600 hPa, and yet it is associated with an easterly moisture flux in the mid-troposphere (Cadet and Nnoli, 1987). On the one hand, upper tropospheric water vapor that was lifted within the Indian Monsoon can be mixed into the mid-tropospheric AEJ (Sy et al., 2018, Niang et al., 2020). This water vapor is then transported to the Sahel, where it might represent only a minor contribution compared to the lower tropospheric moisture fluxes, but appears to be correlated with the Sahelian precipitation (Sy et al., 2018). On the other hand, the monsoon southwesterlies from the Gulf of Guinea are found to feed moisture into the AEJ, when they are lifted over the Sahel through convection to altitudes of the AEJ (Cadet and Nnoli, 1987, Niang et al., 2020). The other feature shown in Fig. 2.1b is the transport pathway as response to the divergent anticyclonic flow at the Saharan high. This flow is associated with relatively dry air from the extra-tropical upper troposphere. Roca et al. (2005) documents that through occasional Rossby wave breaking the polar jet can feed on its southern side at 50° N and 250 hPa very dry air masses into this anticyclonic flow over North Africa. Along its propagation over the Saharan regions, this flow can receive an input of low-level moisture as result of enhanced boundary layer mixing, as during summer the top of the Saharan boundary layer can reach altitudes up to 5–6 km (Hall and Peyrillé, 2006). But still the anticyclonic flow over North Africa does not constitute a significant moisture source for the Sahel. Nevertheless, both discussed mid-tropospheric flows together with their dry and unsaturated air masses can be essential for the development of MCSs, what will be discussed in the following section.

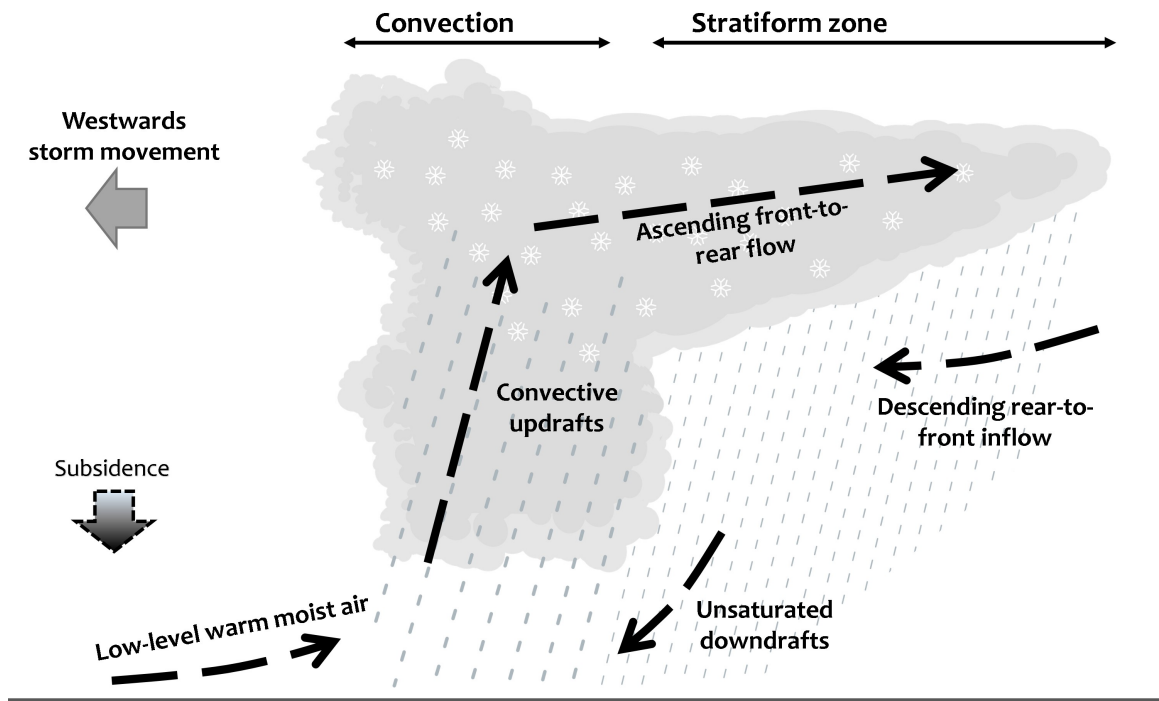


Figure 2.2. Schematic cross-section of a tropical squall line (adapted from Fig. 1 in Houze et al., 1989). The black dashed arrows indicate the dominant flow patterns within the convective system.

2.3. Structure of Squall Lines

As previously discussed, the Sahel represents a region, where air masses with markedly different pathways and moisture contents come together and hence induce high levels of convective instability and vertical wind shear. A prominent example of MCSs that develop in this environment are so-called squall lines with a characteristic north-south orientation and a fast westwards propagation (Houze, 1977, Houze et al., 1989, Fink and Reiner, 2003). Such squall line systems are found to generate up to 80 % of the annual rainfall over the Sahel (Dhonneur, 1981, Fink et al., 2006) and are therefore a key driver of monsoon precipitation.

Following the conceptual model of Houze et al. (1989), Fig. 2.2 provides a schematic cross-section through a squall line, which typically consists of a leading convective zone followed by a stratiform cloud shield. Near-surface moist air masses (e.g. the monsoon southerlies) inject moisture into the system and foster the development of convective updrafts. Here, the moist convection induces the intense formation of precipitation and clouds reaching up to the tropopause. The trailing stratiform part is characterized with a slowly ascending front-to-rear flow that emerges from the upper portion of the convective part and advects detrained ice particles rearward. During the ascend, these particles grow through vapor deposition and eventually fall out of the cloud shield. This results in moderate stratiform precipitation in the stratiform zone of the

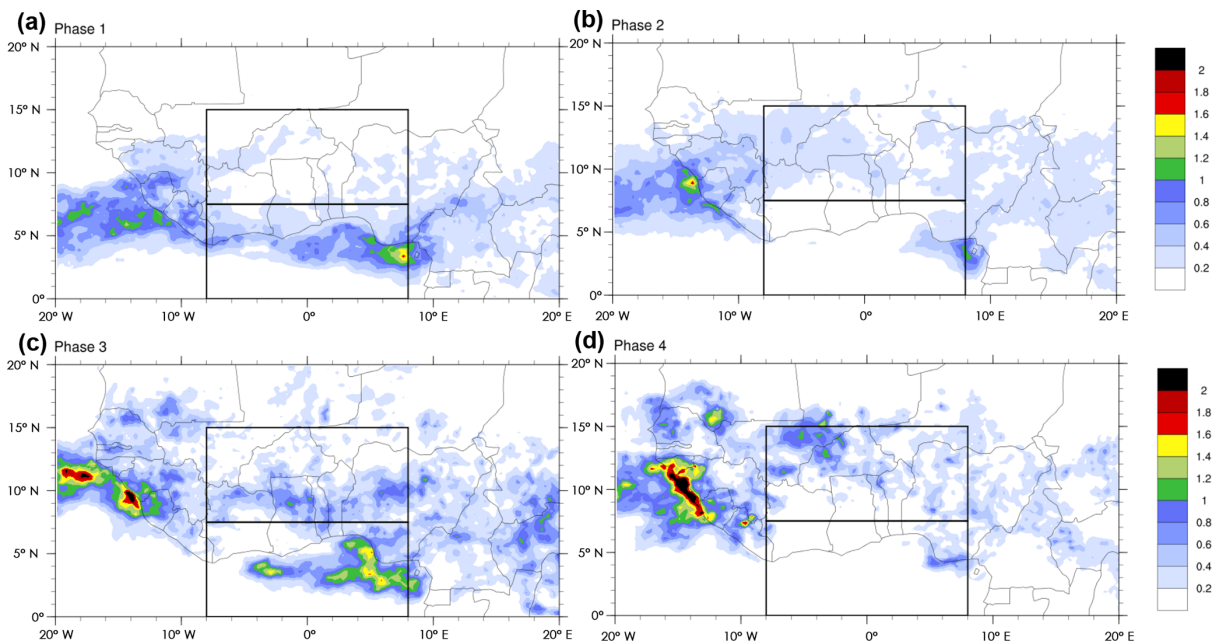


Figure 2.3. Horizontal distribution of mean precipitation during the four DACCIWA phases (a–d), as described in Sect. 2.4. The underlying data are based on the precipitation product TRMM and given in mm per hour. Figure is taken from Knippertz et al. (Fig. 6, 2017).

squall line, accounting for 40 % of the total rain from the squall line system (Houze, 1977). An additional rear-to-front flow (e.g. the AEJ) penetrates dry and unsaturated air masses into the rear of the system below the stratiform cloud shield. This fosters the evaporation of falling rain and the development of unsaturated downdrafts through the evaporative cooling of surrounding air masses. These downdrafts generate a near-surface cold pool and can be injected into the convective updraft, thereby fostering the formation of new convective cells.

2.4. The Monsoon Season 2016

As the West African area lacks of dense and operational measurement networks, various large international field campaigns have been organized in the past in order to advance the general understanding of the WAM and its controlling factors. For instance, the African Monsoon Multidisciplinary Analysis (AMMA) program was a comprehensive field campaign that was carried out in the Sahel in 2006 aiming at the documentation of the WAM climate system (Redelsperger et al., 2006). Details about related water isotope studies performed in this context will be presented in Sect. 3.3.

In this work, a main focus will be on the WAM season 2016, which was studied and documented in detail during the Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa (DACCIWA) project (Knippertz et al., 2017). This included extensive field measurements in the coastal West

Africa as well as radiosonde and aircraft measurements (Flamant et al., 2018). As part of this project, Knippertz et al. (2017) performed an in-depths characterization of the meteorological conditions during the early WAM season 2016 and divided the temporal WAM evolution in four different phases (see Fig. 2.3), which will be the basis for the monsoon analyses in Chap. 5 and in Sect. 6.2 and 6.3.

During the pre-onset stage (Phase 1, 01–21 June 2016, Fig. 2.3a) the rainfall maximum lied over the Guinea Coast and was characterized by several rainy episodes associated with synoptic-scale disturbances (e.g. see Maranan et al., 2019). Just before the monsoon onset, the northern African regions were affected by a substantial extra-tropical disturbance inducing strong northerlies towards the Sahel zone and thereby suppressing the Sahelian precipitation. This feature was accompanied with an unusually strong and eastward shifted SHL, until an abrupt ventilation from the Tropical Atlantic weakened the SHL. After a subsequent gradual intensification of the SHL, a shift of maximum rainfall to the Sahel took place and initiated the post-onset stage (Phase 2, 22 June–20 July 2016, Fig. 2.3b), with more or less climatological, undisturbed monsoon conditions. That is, the monsoon onset was associated with the development of strong southerly monsoon winds that transported moisture to the Sahel and thereby created a northward shift of the ITD up to 20° N. In contrast, the Guinea Coast was rain-free and anomalously dry during this period. In Phase 3 (21–26 July 2016, Fig. 2.3c) again an SHL breakdown occurred and subsequently an unusual wet westerly wind regime formed. This caused widespread precipitation over large parts of West Africa, shifting the maximum rainfall southwards to the Guinea Coast. The return of the circulation to undisturbed monsoon conditions took place in Phase 4 (27–31 July 2016, Fig. 2.3d), with the SHL and maximum rainfall returning to conditions similar to Phase 2.

3 | Stable Water Isotopologues in the Troposphere

An atom of a given element is uniquely defined based on the amount of protons and electrons, however, the number of neutrons and therefore the mass may differ for different atoms of the same element. These variants of a given element are called isotopes, and a molecule that consists of different isotope combinations is referred to as isotopologue. For instance, the naturally most abundant isotopes of hydrogen are ^1H (in the following denoted as H) and ^2H (deuterium, denoted as D) and of oxygen ^{16}O (denoted as O), ^{17}O and ^{18}O . Together these isotopes can form different isotopologues of water (in the following denoted as water isotopes) with relative abundances given in Table 3.1. The most frequent water isotope is the light H_2^{16}O (denoted as H_2O) with making up more than 99 % of all water isotopes, whereas the heavier water isotopes show maximum occurrences of up to 0.2 %. In general, the water isotope occurrences decrease with increasing molecular masses. For describing distributions of water isotopes the δ -notation (in ‰) is commonly used, for instance between H_2O and the heavier HDO, with both given as volume mixing ratios:

$$\delta\text{D} = \left(\frac{\text{HDO}/\text{H}_2\text{O}}{R_{\text{vsmow}}} - 1 \right) \cdot 1000 \quad (3.1)$$

R_{vsmow} is the isotopic ratio of the Vienna Standard Mean Ocean Water as defined by the International Atomic Energy Agency (IAEA) (see Table 3.1; Hagemann et al., 1970). Analogous to Eqn. 3.1, $\delta^{18}\text{O}$ refers to the ratio of H_2^{18}O to H_2O . Even though HDO is not the most abundant heavy water isotope, the focus of the following study will be on ratios of H_2O against δD , as these two water isotopes are best retrievable from remote sensing observations (see Chap. 4).

The subtle differences in the molecular structure cause that different water isotopes react slightly different to atmospheric processes that depend on the individual atomic masses, e.g. phase changing processes. This behaviour makes the analysis of different water isotopes extremely useful for investigating the hydrological cycle, as each phase changing process leaves individual imprints on the ratio of different water isotopes. To emphasize the potential impact of such an analysis, Fig. 3.1 extends the schematic overview of a squall line system as given in Fig. 2.2 with those processes that alter the ratio of light and heavy water isotopes. The dominant processes

Table 3.1. Natural relative occurrences of the most abundant stable water isotopes in the atmosphere (Galewsky et al., 2016).

Water Isotope	Occurrence (%)
H_2^{16}O	99.73
H_2^{18}O	0.199
H_2^{17}O	0.038
HD^{16}O	0.031

are surface and rain evaporation as well as rain condensation, which are in close connection with the dynamical features of the WAM. It underlines the variety and complexity of processes that affect abundances of water isotopes and thus requires careful analyses for disentangling the individual effects of each process to the isotopic composition of water vapor. Throughout this work, this figure will be discussed in detail.

The following sections provide insights into common theories describing the behaviour of different water isotopes by emphasizing the paired analysis of light and heavy isotopes (Sect. 3.1). Furthermore, the observation of tropospheric water isotopes from space-based remote sensing and different modeling approaches will be motivated (Sect. 3.2). Section 3.3 presents results from previous studies investigating abundances of stable water isotopes during the West African Monsoon in order to motivate and highlight the key objectives of the thesis as well as the datasets considered for addressing the presented research targets.

3.1. Theory for Interpreting Paired $\{\text{H}_2\text{O}, \delta\text{D}\}$ Distributions

Throughout the last decades, the paired analysis of stable water isotopes has proven highly valuable for retrieving information about atmospheric moisture processes (e.g. Worden et al., 2007, Noone et al., 2011, González et al., 2016, Schneider et al., 2016, Lacour et al., 2018, Eckstein et al., 2018, Graf et al., 2019). As a theoretical basis for interpreting paired distributions of H_2O (given in volume mixing ratios) and δD in water vapor, Noone (2012) compiled a set of idealized process curves to describe how different tropospheric moisture processes are reflected in the $\{\text{H}_2\text{O}, \delta\text{D}\}$ phase space. In this work, the focus is on those process curves that govern the isotopic variability in the Sahelian troposphere during the West African Monsoon, as summarized in Fig. 3.2. This section provides the theoretical background for the individual process curves, for which observational and model-based evidence will be given in the course of this study.

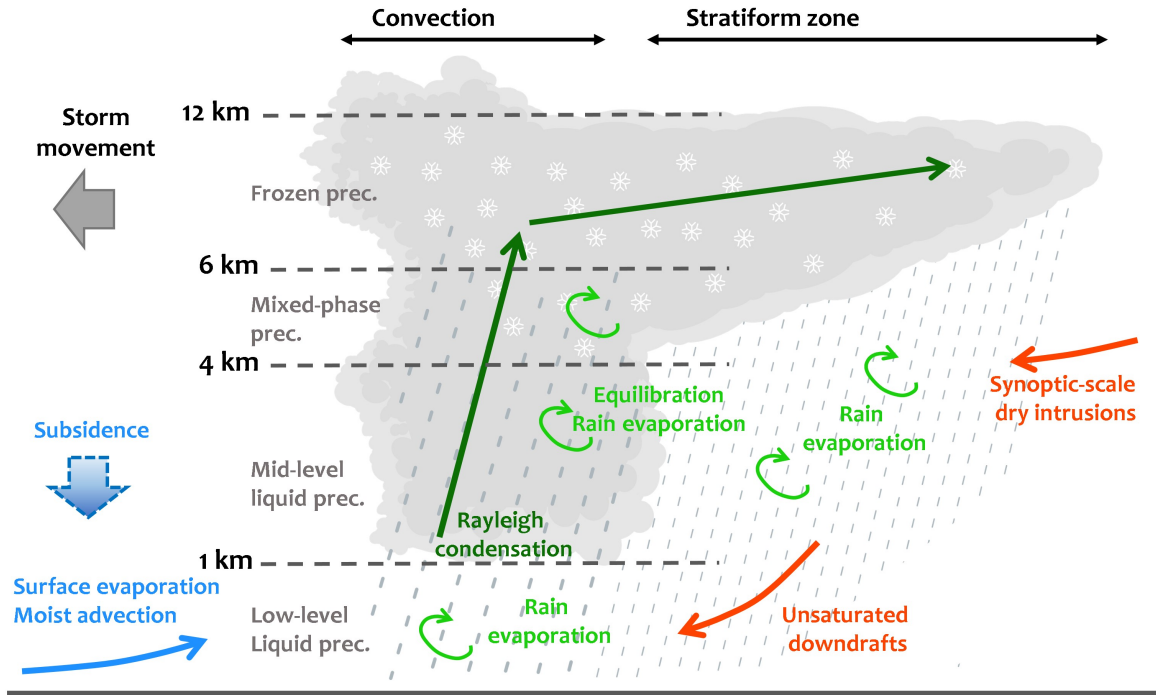


Figure 3.1. Schematic cross-section of a tropical squall line (adapted from Fig. 4a in Risi et al., 2010a). The blue arrows denote dry air intrusions into and within the system and the red one the moist contributions (analogous to Fig. 2.1). The dark and light green arrows illustrate the occurrences of microphysical rain processes for the different precipitation zones (described in grey).

3.1.1. Air Mass Mixing

First, let's assume that two air parcels with the specific moisture contents q_{v0} and q_{v1} and the isotopic compositions δD_0 and δD_1 mix without fractionation (i.e., no phase changes during the mixing). The mixed moisture composition q_v and δD then result as follows:

$$q_v = f q_{v0} + (f - 1) q_{v1} \quad (3.2)$$

$$\delta D = \frac{f q_{v0} \delta D_0 + (f - 1) q_{v1} \delta D_1}{q_v} \quad (3.3)$$

with f indicating the relative contribution of the two air masses (Noone et al., 2011). Note that the subscript v is used for the vapor phase and c for the condensate throughout this work (the only exception is made for H_2O and δD that both refer to the vapor phase, if no further subscript is given). Later, the distinction will be made for different categories of condensate, namely rain (r), snow (s), liquid (l) and ice (i) clouds. While q_v exhibits linear mixing, δD follows a hyperbolic curve, as the ratio between light and heavy isotopes is dominated by the moister air mass (Noone, 2012). The position of the hyperbolic curve in the $\{H_2O, \delta D\}$ diagram is determined

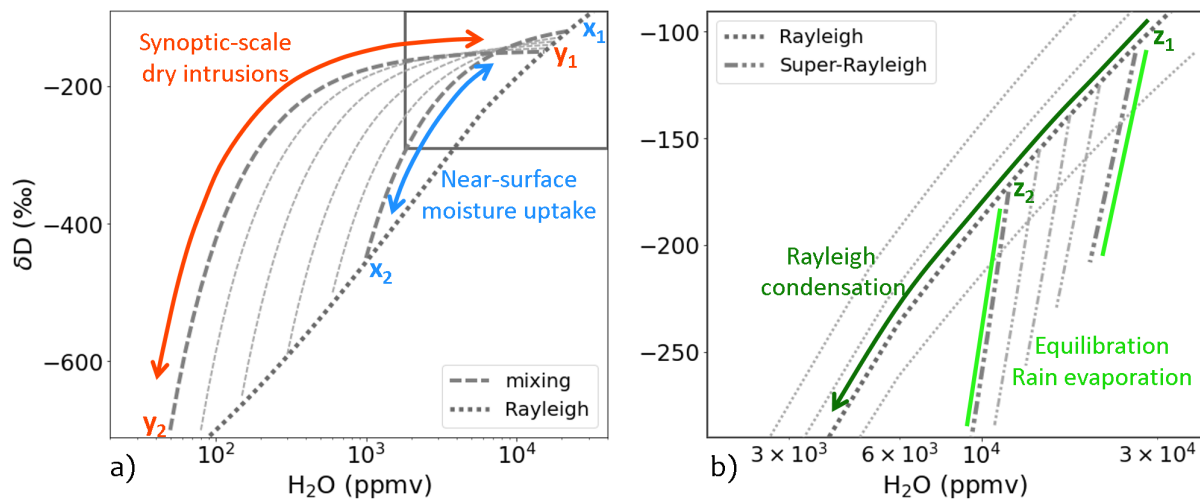


Figure 3.2. Overview of idealized process curves for paired distributions of H_2O and δD according to Noone (2012). The colored arrows indicate process curves that are found to represent the isotopic composition of the Sahelian troposphere during the WAM season 2016. (a) Effects of air mass mixing, where the blue (orange) curve indicate mixing between the end members x_1 and x_2 (y_1 and y_2). (b) Effects of microphysical processes, where the dark green line marks a Rayleigh process with initial conditions of $T_0 = 30^\circ\text{C}$, $RH_0 = 90\%$, and $\delta\text{D}_0 = -80\text{‰}$ and the light green lines Super-Rayleigh signals starting at two different positions of the Rayleigh line (z_1 and z_2). The colors are chosen in agreement with Fig. 3.1. Note that (b) only shows a subset of (a) marked by the black box in the top right corner of (a).

by the isotopic composition of the moist and dry end members. For instance, Fig. 3.2a shows the mixing curves for a dry mixing process, where the dry end member is located at very low H_2O and δD (orange curve), and for a moist mixing process with both end members (x_1 and x_2) being relatively moist (blue curve). For instance, the dry mixing curve is representative for air masses that originate from dry regions of the upper troposphere (mixing member y_2) and become more humid while they are subsiding into lower altitudes (mixing member y_1), whereas the moist mixing curve illustrates a near-surface moistening due to surface evaporation (blue and red arrows in Fig. 3.1, discussed in Sect. 5.3 and 5.4.2).

3.1.2. Isotopic Fractionation

Second, if microphysical processes take place and induce a phase change of atmospheric water, fractionation between H_2O and HDO occurs. Since the HDO molecule has higher binding energies in the condensed phase than H_2O , the two isotopes have different saturation vapor pressures leading to equilibrium fractionation (Urey, 1947, Bigeleisen, 1961). Additionally, differences in molecular mass leads to non-equilibrium fractionation due to differing diffusivities. The former

refers to a reversible isotope separation under thermodynamic equilibrium between a condensate (isotopic ratio R_c) and the ambient vapor (isotopic ratio R_v) according to the fractionation factor:

$$\alpha_{eq,c} = \frac{R_c}{R_v} \quad (3.4)$$

The values for $\alpha_{eq,c}$ vary with temperature as also the saturation vapor pressure does and were determined for liquid and ice condensation in various laboratory studies (e.g. Merlivat and Nief, 1967, Majoube, 1971, Horita and Wesolowski, 1994). Non-equilibrium fractionation is assumed to occur in addition to equilibrium fractionation for processes that enforce a fast isotope flux between vapor and liquid, for instance in case of ventilation or undersaturation.

3.1.3. The Rayleigh Model

A simple framework for the isotopic fractionation in a precipitating air parcel is the Rayleigh distillation process (Rayleigh, 1902, Dansgaard, 1964). In this model, a moist adiabatic ascent is assumed with immediate removal of the condensate (Johnson et al., 2001). As soon as the dew point temperature is reached, condensation begins and condensate forms from ambient vapor under equilibrium conditions. While this process enriches the condensate with heavy isotopes, the ambient vapor gets depleted according to

$$\ln \left(\frac{R_v}{R_{v0}} \right) = (\alpha_{eq} - 1) \ln \left(\frac{q_v}{q_{v0}} \right) \quad (3.5)$$

The conditions at the starting point of the ascent are defined by q_{v0} and the isotopic ratio R_{v0} . For condensation above the frost point (263 K, according to Noone (2012) and Ciais and Jouzel (1994)) the fractionation factor for liquid water is used (Eqn. (3.4)) and at colder temperatures, the factor over ice is applied. A typical Rayleigh line for convective condensation over West Africa is shown in dark green in Fig. 3.2b (dark green arrows in Fig. 3.1, discussed in Sect. 5.3 and 5.4.3).

3.1.4. The Super-Rayleigh Regime

If a liquid hydrometeor falls into unsaturated air, evaporation takes place and acts as a reversed distillation process (Bony et al., 2008), i.e. in case of the rain drop fully evaporating the isotopic composition of the vapor resembles that of the rain (Risi et al., 2008a). During partial evaporation of rainfall the lighter isotopes evaporate preferentially (Lee and Fung, 2008, Risi et al., 2008a, Noone, 2012). While enriching the rain water, the ambient vapor content increases due to the input of relatively more depleted evaporated rainfall water. In the $\{H_2O, \delta D\}$ space this leads to a drop below the Rayleigh curve (Dansgaard, 1964, Rozanski et al., 1992) and creates a so-called Super-Rayleigh signal, representing a Rayleigh process with an increased fractionation factor (Noone, 2012):

$$\alpha > \alpha_{eq} \quad (3.6)$$

Starting at different positions of the Rayleigh curve, this creates signals indicated by light green arrows in Fig. 3.2b. As will be shown in Chap. 5, these curves correspond to rain evaporation in mid-levels (4–6 km) and near-surface sub-cloud areas (0–1.5 km) over the Sahel (light green arrows in Fig. 3.1).

In saturated conditions (RH = 100 %), equilibrium exchange of water molecules between vapor and liquid may affect their isotopic compositions, because saturation does not necessarily imply that an equilibrium between liquid and vapor is reached immediately for HDO as well. Particularly in the case of a fast process such as during the fall of a droplet, the isotope composition of the vapor can be altered by equilibrium exchange with the falling droplet. If an isotopic disequilibrium between the ambient air (δD) and the rain drop (δD_r) exists, an HDO flux works towards equilibrating both phases (Stewart, 1975, Lawrence et al., 2004). For instance, if rain drops with low δD fall through saturated areas with a relatively more enriched vapor, a net isotopic flux from the vapor to the condensate occurs, while q_v remains constant. This leads to a lowering of δD at constant q_v . Thereby, $\{H_2O, \delta D\}$ signals develop, which lie below the Rayleigh curve (light green lines in Fig. 3.2, light green arrows in Fig. 3.1, discussed in Sect. 5.4.4).

The disequilibrium between rain and vapor can be approximated as follows (e.g. Tremoy et al., 2014, Aemisegger et al., 2015, Graf et al., 2019):

$$\delta D_{v,deq} = \delta D - \delta D_{v,eq} = \delta D - [\alpha_{eq}(\delta D_r + 1000) - 1000] \quad (3.7)$$

$\delta D_{v,eq}$ is the isotopic composition that the vapor would have if it were in isotopic equilibrium with the rain drop (subscript r stands for rain). In a saturated and equilibrated state, $\delta D_{v,deq}$ tends towards 0 ‰, whereas partial rain evaporation (both equilibrium and non-equilibrium fractionation) generates a negative disequilibrium.

3.1.5. Paired Interpretation Using Frequency Contours

For an improved, paired interpretation of different water isotopes, various studies have proposed the use of two-dimensional distribution functions that statistically summarize given point clouds, for instance, in the paired $\{H_2O, \delta D\}$ phase space (e.g. Worden et al., 2007, Noone, 2012, Schneider et al., 2017). This is particularly useful, if large datasets are investigated, where a traditional scatter plot would become easily cluttered.

The thesis relies on the framework of Christner et al. (2018) and Eckstein et al. (2018), where the $\{H_2O, \delta D\}$ pair data are summarized in relative normalized frequency distributions. To compute these, the H_2O and δD distributions are binned on a chosen grid and evaluated with a two-dimensional histogram. The histogram counts are smoothed with a Gaussian filter with a chosen standard deviation (in this work, typically a value of 15–25 bins, depending on the data ensemble size). Afterwards, the smoothed histogram counts are normalized according to the

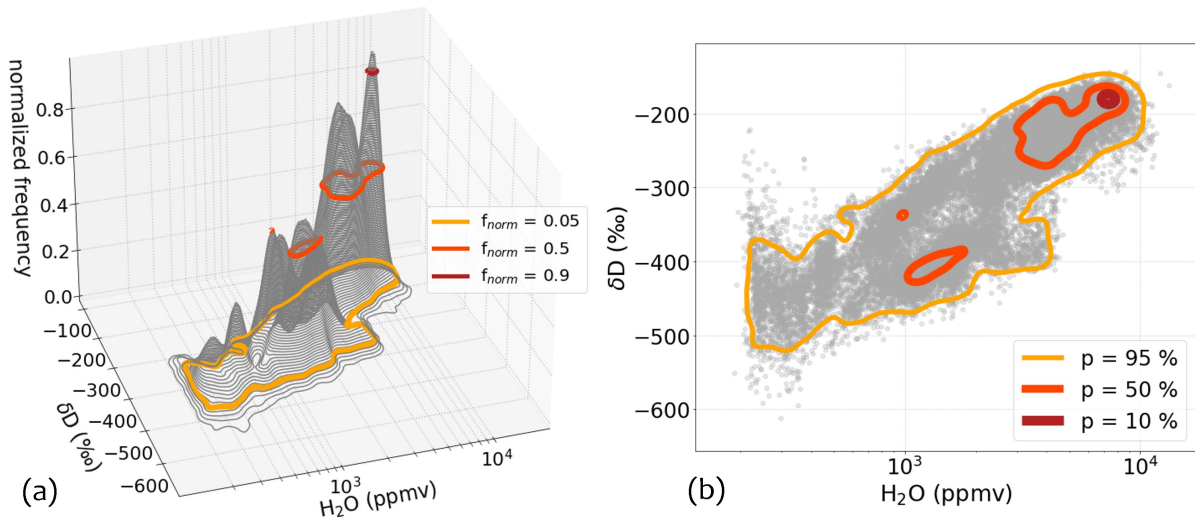


Figure 3.3. Relative normalized frequency distributions describing an exemplary $\{H_2O, \delta D\}$ pair dataset from the Metop/IASI sensor (see Sect. 4.3). (a) Contour lines describing a two-dimensional histogram that indicates the relative normalized frequencies of the $\{H_2O, \delta D\}$ data points binned on a uniform grid (50 bins between the min-max-values of H_2O and 100 bins for δD). The colored lines correspond to the normalized frequencies of 0.05 (orange), 0.5 (red) and 0.9 (dark red). (b) Use of the colored relative normalized frequency contours from (a) to interpret the two-dimensional $\{H_2O, \delta D\}$ data. For this purpose, the normalized frequencies are transformed to the corresponding data fractions that are summarized by each contour: The orange contour summarizes 95 % of all data, the red one 50 % and the dark red contour the main 10 %.

sum of the histogram and its maximum value (Fig. 3.3a).

An example of how these histogram diagnostics can be used is given in Fig. 3.3b, where a point cloud in the $\{H_2O, \delta D\}$ phase space is summarized by three different relative normalized frequency distributions (also denoted as two-dimensional histogram contours). The contour corresponding to the normalized frequency of 0.05 (0.5, 0.9) is used to indicate the distribution of the main 95 % (50 %, 10 %) of all considered data points. The 95 %-contour displays the overall shape of the full point cloud and captures also robust, outlying structures, whereas the 50 %-contour represents the median amount of the paired $\{H_2O, \delta D\}$ data points. The weighted maximum of the frequency distribution is indicated by the 10 %-contour.

3.2. Tropospheric $\{H_2O, \delta D\}$ Pair Datasets

Nowadays, a variety of observational platforms as well as atmospheric models are capable of providing reliable information about isotope abundances in tropospheric water vapor and

precipitation. Examples for the former are paleoclimatic records of δD and $\delta^{18}\text{O}$ from archives such as glaciers, ground ice, speleothems, corals, trees and marine sediments (e.g. Past Global Changes (PAGES) Iso2k project; Konecky et al., 2020), global station-based measurements of isotopes in precipitation (Global Network of Isotopes in Precipitation (GNIP) by IAEA and the World Meteorological Organization), a global laser-based network of isotope measurements in water vapor (Stable Water Vapour Isotope Database (SWVID); Wei et al., 2019), campaign-based and long-term aircraft and ship measurements (e.g. Brenninkmeijer et al., 2007, Dyroff et al., 2015, Renfrew et al., 2019) as well as infrared remote sensing platforms capturing isotopes in water vapor from ground (Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON); e.g. Rokotyan et al., 2014, Schneider et al., 2016) and from space (e.g. Worden et al., 2006, Frankenberg et al., 2009, Boesch et al., 2013, Schneider et al., 2016). In the meantime, various Global Circulation Models (GCM; e.g. Yoshimura et al., 2008, Risi et al., 2010c, Werner et al., 2011, Eckstein et al., 2018) as well as Regional Climate Models (RCM; e.g. Sturm et al., 2005, Pfahl et al., 2012) have been successfully equipped with the fractionation physics of stable water isotopes. Joint efforts for creating unified and merged model products were realized within the project Stable Water Isotope Intercomparison Group (SWING; Risi et al., 2012). Additionally, work has been done towards developing so-called retrieval simulators that merge model output according to characteristics of space-based remote sensing products (e.g. Field et al., 2012, Schneider et al., 2017). Another powerful tool for investigating the sources and pathways of air masses and thus to better understand its isotopic composition is the use of Lagrangian models. In contrast to Eulerian models that simulate the atmosphere for fixed, discrete grid points and volumes (such as GCMs and RCMs), Lagrangian models are used to calculate the propagation of infinitesimal small air parcels through the atmosphere (e.g. Sodemann et al., 2008, Dütsch et al., 2018).

As this thesis aims at cross-comparisons between modeled and remotely sensed $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair data, the following section will provide further information for related data platforms and introduce the considered models.

3.2.1. Remote Sensing Products

During the last decades the space-based remote sensing of tropospheric water isotopologues has progressed considerably in terms of retrieval development, quality and application. Even though HDO is not the most abundant heavy water isotope (see Table 3.1), it is nevertheless, of course after the light H_2O , the one that is currently best retrievable from spectral measurements. This is because HDO exhibits some strong spectroscopic features in the infrared spectral window, which are distinct from those of H_2O (see Sect. 4.2.2; Toth, 1999, Clerbaux et al., 2009).

On the one hand, cloud-free observations from short-wave infrared sensors were used to generate

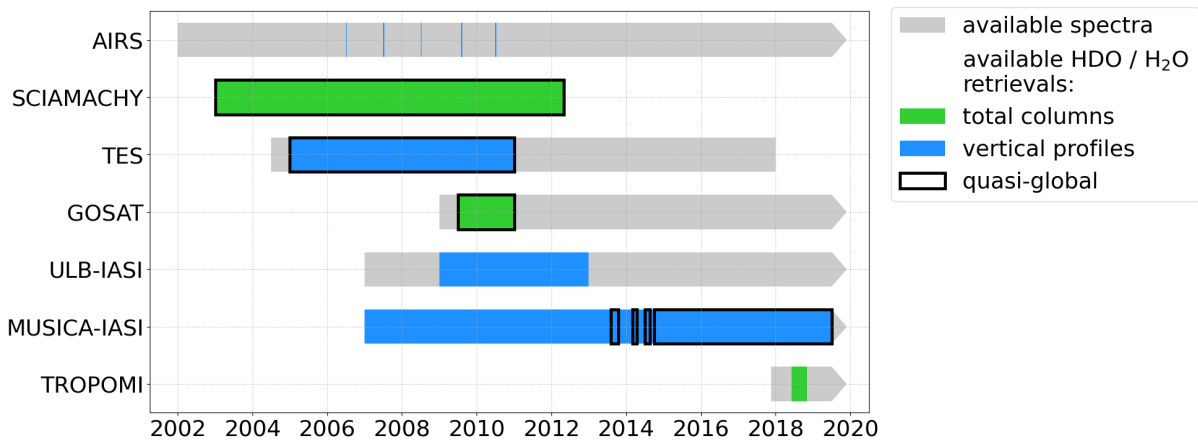


Figure 3.4. Overview of tropospheric water isotope retrievals from satellite sensors. The grey bars indicate the time periods with available radiance measurements for each sensor. The periods for which water isotope information has already been retrieved are shown in green for total column products and in blue for vertically resolved profiles. In case of quasi-global retrievals, the bars are framed in black.

total columns of the ratio HDO/H₂O, specifically for SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography; Frankenberg et al., 2009, Scheepmaker et al., 2013), GOSAT (Greenhouse Observing Satellite; Boesch et al., 2013, Frankenberg et al., 2013) and TROPOMI (Tropospheric Monitoring Instrument; Schneider et al., 2020). As these sensors measure sun light that is reflected on the Earth’s surface, current HDO/H₂O retrievals are only available for land observations. Typically, such tropospheric HDO/H₂O total column products provide highest sensitivity near the ground between 0–2 km.

On the other hand, thermal infrared sensors allow for retrieving HDO/H₂O ratios with weak vertical profile information for land as well as ocean observations. Measurements from TES (Tropospheric Emission Sounder; Worden et al., 2006, 2012), AIRS (Atmospheric Infrared Sounder; Worden et al., 2019) and IASI (Infrared Atmospheric Sounding Interferometer; Herbin et al., 2009, Schneider and Hase, 2011, Lacour et al., 2012, Schneider et al., 2016, 2021c) were used for retrieving H₂O and HDO/H₂O with focus on the mid-troposphere (approx. between 750–350 hPa). To ensure coherence in the vertical sensitivities of remotely sensed H₂O and δD , which is needed for a combined interpretation, a further post-processing that creates optimal $\{H_2O, \delta D\}$ pair information is proposed by Schneider et al. (2012) (see Sect. 4.3).

Figure 3.4 provides an overview of the referenced tropospheric water isotope retrievals for the troposphere. This figure reveals the fact that most of the aforementioned retrievals were performed for case studies limited in space and time. Up to now, the only global and long-time referenced space-borne datasets of tropospheric HDO/H₂O are provided for the sensors TES (Worden et al., 2012) and SCIAMACHY (Schneider et al., 2018), but ended with the respective

satellite missions in 2011 and 2012.

The thesis will focus on the IASI retrieval results from Schneider et al. (2021c) and, by adjusting the $\{H_2O, \delta D\}$ pair post-processing from Schneider et al. (2012), generate the multi-annual and global water isotope dataset as indicated in Fig. 3.4 (marked with *MUSICA IASI*). Detailed descriptions of the IASI sensor, the corresponding processing and an in-depth discussion of the resulting dataset are found in Chap. 4.

3.2.2. Isotope-Enabled Modeling

Various techniques of modeling water isotopes have been developed throughout the last decades and established as a powerful tool for studying atmospheric moisture transport and processes. The formalisms describing equilibrium and non-equilibrium fractionation processes of different water isotopes are well-known and can be used to extend an existing hydrological cycle in a model by the additional simulation of stable water isotopes like HDO. This requires to take into account the effects of fractionation and transport on the isotopic composition during every branch of the whole prognostic cycle, such as evaporation from land and ocean surfaces, microphysical cloud and rain processes and atmospheric transport and mixing. Originating in paleo-climatic applications, the isotopologue-enabled atmospheric modeling has evolved along many different atmospheric scales, ranging from large-scale and long-term climate studies to cloud-resolving large-eddy-simulations (see, for instance, the overviews given in Yoshimura (2015) and Galewsky et al. (2016)).

Within this study, different types of models will be used to investigate atmospheric distributions of H_2O against δD . In the following, the focus will be on simulations of the isotope-enabled RCM COSMO_{iso}, the isotope-enabled GCM ICON-ART_{iso} as well as on Lagrangian backward trajectories using the tool LAGRANTO. In order to make such model data directly comparable to satellite data, a short overview will be given about current approaches that aim to adjust model data according to characteristics of remotely sensed observations.

3.2.2.1. The regional model COSMO_{iso}

COSMO_{iso} is the isotope-enabled version of the non-hydrostatic limited-area weather and climate model COSMO (Steppeler et al., 2003) and is documented in detail in Pfahl et al. (2012). As a regional model that is fed by boundary data from a global model, it efficiently enables simulations with high spatio-temporal resolutions at convective-resolving scales. It incorporates the fractionating processes of HDO and $H_2^{18}O$ within its whole hydrological cycle (Pfahl et al., 2012). Fractionation is assumed whenever phase changes occur that involve the vapor phase. A one-moment microphysical scheme is used, which calculates the isotopic composition for water vapor, liquid and ice clouds as well as rain and snow. For this purpose, it includes the

fractionation schemes of Stewart (1975) for rain evaporation and Jouzel and Merlivat (1984) for snow formation. Further, it uses the isotope-enabled multi-layer soil moisture scheme $TERRA_{iso}$ for fractionating soil evaporation and non-fractionating plant transpiration (Christner et al., 2018). Fractionation during ocean evaporation is represented by the Craig-Gordon-model (Craig and Gordon, 1965).

3.2.2.2. The global model $ICON-ART_{iso}$

A very recent isotopically equipped global circulation model is $ICON-ART_{iso}$ (Eckstein et al., 2018) that builds on the numerical weather prediction packages of the non-hydrostatic global forecast model $ICON$ (Icosahedral Nonhydrostatic; Zängl et al., 2015) and on the flexible tracer framework from the extension ART (Aerosols and Reactive Trace Gases; Rieger et al., 2015, Schröter et al., 2018). This tracer framework facilitates a flexible set up of artificial tracers tagging chosen air masses and moisture sources within a single model run.

In contrast to $COSMO$, $ICON$ is based on an unstructured triangular grid that enables high scalability on massively parallel high performance computing systems. Further, it allows for incorporating local grid refinements (nesting) within the global model domain and with a two-way-interaction, whose impact towards improving the representation of the atmospheric state has been shown for instance in Pante and Knippertz (2019) and Weimer et al. (2020). The exchange of information between the global and nested domain is realized with a relaxation-based interpolation for the prognostic variables of $ICON$, i.e. horizontal and vertical wind fields, virtual potential temperature, density, specific humidity and liquid and frozen cloud water (Pante and Knippertz, 2019, Reinert et al., 2019). As part of the thesis, the two-way-nesting functionality has been extended to also considering the water isotope state from $ICON-ART_{iso}$.

The fractionation physics of stable water isotopes are adapted from $COSMO_{iso}$. In its current version, $ICON-ART_{iso}$ does not yet consider fractionation within the soil moisture scheme $TERRA$. Instead, climatological mean values are used for setting the isotopic composition of water evaporated from land (Terzer et al., 2013) and ocean (LeGrande and Schmidt, 2006) surfaces. A comprehensive and detailed documentation of the model structure and the considered fractionation schemes is given in Eckstein (2017).

3.2.2.3. The Lagrangian analysis tool $LAGRANTO$

The Lagrangian analysis tool $LAGRANTO$ allows for calculating backward and forward air trajectories based on atmospheric data of Eulerian models, i.e. models that simulate the atmosphere on fixed grid points (Wernli and Davies, 1997). $LAGRANTO$ is not restricted to a pre-defined, fixed spatial grid, but interpolates the individual pathway of a single air parcel through the

atmosphere by using the three-dimensional Eulerian wind fields. Additionally, it provides the functionality of interpolating meteorological variables that are included in the model output onto the computed trajectory pathways. In its current version, LAGRANTO is capable of processing data from COSMO simulations (Miltenberger et al., 2013) as well as from ECMWF (European Centre for Medium-Range Weather Forecasts) reanalyses for the atmosphere (Sprenger and Wernli, 2015) and for the ocean (Schemm et al., 2017).

3.2.2.4. Retrieval Simulation Techniques

In order to achieve reliable cross-comparisons of remotely sensed water vapour observations against data from isotope-enabled models, it is necessary to take into account that their vertical data structure may be fundamentally different. While an atmospheric model provides distinct values on a discrete vertical grid, remote sensing products typically have a limited vertical resolution and sensitivity. The vertical smoothing of a remote sensing product is described by so-called averaging kernels that relate the retrieved product with the true atmospheric state (detailed discussions in Sect. 4.2.3).

To overcome the discrepancies of the vertical data structure, the general approach is to convolve the model state with the averaging kernels of the sensor used for comparison. However, for highly variable atmospheric gases such as water vapour also the averaging kernels are often very variable, as they highly depend on the atmospheric temperature and water vapour profiles. Therefore, the averaging kernels need to be specifically determined for the individual meteorological conditions as present in the model (Schneider et al., 2017).

For instance, first model comparisons with the sensor TES were performed based on monthly mean maps for the TES averaging kernels. After collocating the modeled atmosphere with the actual meteorology from the observations, the corresponding TES averaging kernels were then applied to the model data (Risi et al., 2010a, 2012). Yoshimura et al. (2011) applied a very similar technique, but instead of monthly mean kernels they used the actual TES kernels.

A first approach to compute the kernels explicitly for the modeled atmosphere is realized in Field et al. (2012). Here, statistical relations between remote sensing characteristics of TES and some selected categorical parameters serve to predict the averaging kernels for a limited amount of combinations of surface and atmospheric conditions. This method offers the advantage of the kernel calculations being very fast, however, it is mainly restricted to atmospheric conditions of low latitudes and over oceans.

Further, Schneider et al. (2017) present an approach for simulating the characteristics of the sensor IASI based on a modeled atmosphere by relying on a simplified radiative transfer model. Despite the strong simplifications of the radiative transfer assumptions, already a reliable prediction of

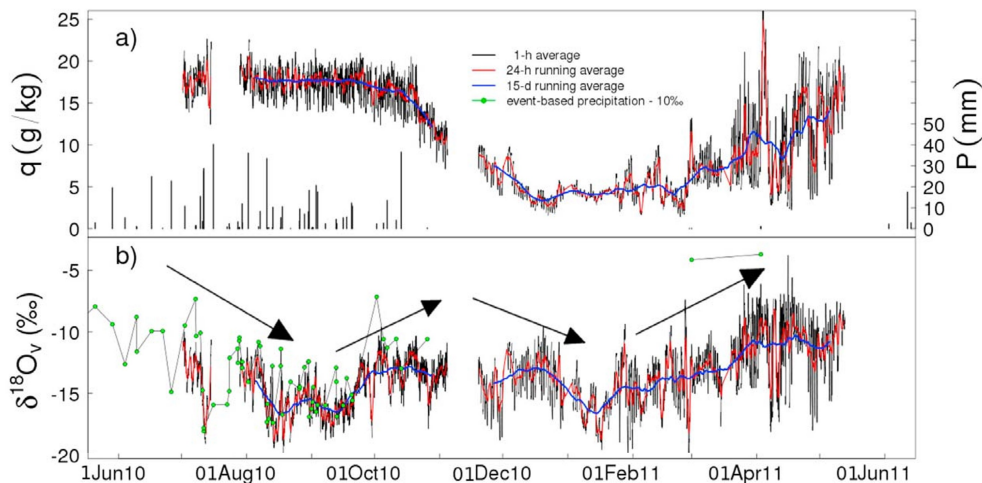


Figure 3.5. Temporal evolution from June 2010 to May 2011 of (a) specific humidity along with event-based precipitation amount P (black bars) and (b) $\delta^{18}\text{O}$ along with event-based values (green dotted line). The 1h (black), 24h (red) and 5-day (blue) averages are shown, respectively. The figure is taken from Tremoy et al. (Fig. 1, 2012).

the complex remote sensing characteristics is achieved with improving the model-to-satellite cross-comparison (Eckstein et al., 2018).

3.3. Isotopic Characterization of the West African Monsoon

So far only few studies have investigated tropospheric distributions of stable water isotopes during the WAM. In the context of the AMMA campaign (see Sect. 2.4) the isotopic composition of rainfall samples ($\delta^{18}\text{O}_r$) collected in the Sahel during the monsoon development were analyzed (Risi et al., 2008b) and along individual squall line events (Risi et al., 2010a). The former led to the conclusions that during the pre-onset stage the isotopic composition of precipitation reflects the intensity and organization of individual convective systems, whereas in the post-onset stage the abrupt increase of convective activity over the Sahel goes along with a persistent and abrupt drop in $\delta^{18}\text{O}_r$. Similar effects were identified during the passage of squall lines and were attributed to dynamical mixing (downward advection of air at the rear of the convective systems) and microphysical processes (rain evaporation and equilibration), with both also depleting $\delta^{18}\text{O}$ in the vapor. Supplementary model analyses from Risi et al. (2010b) underlined the importance of convection as controlling factor for isotope abundances in water vapor and precipitation in wet periods over the Sahel, whereas during dry periods the isotopic composition is dominated by dry air mass mixing. In further studies, Tremoy et al. (2012) supported these findings by investigating a 1-year record of laser-based $\delta^{18}\text{O}$ measurements (Fig. 3.5b) and demonstrated

the potential of a paired analysis with q_v (Fig. 3.5a). In contrast to q_v , they identified a W-shape in the annual cycle of $\delta^{18}\text{O}$ with minimum isotope values during summer, when convection and q_v are maximum, and during winter as response to extra-tropical dry intrusions (with a minimum in q_v). In a follow-up study, Tremoy et al. (2014) performed a systematic classification of convective rain systems along different WAM seasons and again found that most rain events correspond to a strong isotopic depletion in lower tropospheric water vapor. By using GNIP rainfall data, Nlend et al. (2020) identified a similar W-shape over the coast of Cameroon at the Gulf of Guinea as Tremoy et al. (2012) found over the Sahel, with the difference that the coastal $\delta^{18}\text{O}$ minima appear during spring and autumn, when convection is maximum.

However, Risi et al. (2010b) concluded that the quantification of convective processes for explaining isotopic distributions in water vapor still remains a key challenge. As emphasized in Sect. 3.1, different processes may lead to similar signals in individual isotope distributions, making a direct attribution of a single moist process to a specific isotope composition very challenging. This issue applies particularly to regions like the Sahel, where a variety of dynamical and microphysical processes superpose and lead to highly variable moisture distributions in the troposphere (see Sect. 2.1). Therefore long-term and extended networks of isotopic data over West Africa (Risi et al., 2008b) in addition to precise model simulations (Risi et al., 2010b) are highly useful in order to better understand the West African Monsoon and its water cycle.

3.4. Thesis Objectives

As discussed in Sect. 3.3, the analysis of stable water isotopes over West Africa during the monsoon period is still a fairly young research field with great potential for investigating the hydrological cycle in the Tropics. The purpose of this section is to emphasize the general motivation of this thesis by outlining the research goals, which address current challenges and issues highlighted in Sect. 3.3. Further, this section provides an overview about the datasets considered during this thesis in order to achieve the formulated research goals.

3.4.1. Research Goals

In contrast to the isotopic studies of the WAM discussed in Sect. 3.3, this thesis aims at disentangling the complex moisture pathways of the WAM system by considering, for the first time so far, the paired $\{\text{H}_2\text{O}, \delta\text{D}\}$ phase space for mid-tropospheric water vapor over the Sahel. To achieve this aim, three analysis steps with the following research goals are formulated in this context, what serves as an overall guideline for the thesis:

Table 3.2. Datasets used within this study. Each dataset is described within Sect. 3.4.

Platform	ID	Specification	Hor. Res.	Region	Time Period
IASI	–	{H ₂ O, δ D} proc.	12 km	global	Oct 2014 – Jun 2019
COSMO _{iso}	CISO1	param. convection	14 km	Africa	Jun – Jul 2016,
	CISO2	expl. convection	“	“	“
	CISO3	“	7 km	“	“
ICON-ART _{iso}	IISO1	param. convection	160 km	global	Jan 2017 – Dec 2019
	IISO2	“	80 km	“	“
	IISO3	“, 2-way to IISO4	“	“	“
	IISO4	“, 2-way to IISO3	40 km	Africa	“
LAGRANTO	–	based on CISO2	–	Sahel	Jun – Jul 2016
GPM IMERG	RGPM1	half-hourly prec.	0.1°	Sahel	Jun – Jul 2016
	RGPM2	monthly mean prec.	“	global	Jan 2015 – Dec 2018

1. To address the need of large-scale and long-time observations of tropospheric moisture over West Africa, a novel satellite-based global and multi-annual dataset of {H₂O, δ D} pair data is created.
2. To allow sophisticated interpretations of paired {H₂O, δ D} distributions in the Sahelian mid-troposphere, a model-based framework is developed and applied.
3. By bringing the satellite- and model-based {H₂O, δ D} pairs together, the isotopic perspective is used to disentangle the underlying dynamical and microphysical control mechanisms of tropospheric moisture during the West African Monsoon.

The first research goal will be addressed in Chap. 4. By exploiting the IASI retrieval results that are provided by Schneider et al. (2021c) and are globally available for October 2014 to June 2019, the thesis develops a post-processing for generating an improved {H₂O, δ D} pair product with focus on the mid-troposphere. The respective processing will be presented and the resulting dataset characterized in terms of spatio-temporal representativeness. In addition, a new IASI retrieval simulator based on the full radiative transfer assumptions from the IASI retrieval as used in Schneider et al. (2021c) will be developed and evaluated. This lays the foundation for allowing reliable model-to-satellite comparisons as anticipated within the third research goal.

For addressing the second research goal, a Lagrangian moisture attribution procedure for paired $\{H_2O, \delta D\}$ distributions will be developed in order to disentangle the complex moisture pathways during the WAM (Chap. 5). For this purpose, the LAGRANTO tool will be used to compute backward air trajectories based on COSMO_{iso} simulations for the Sahelian mid-troposphere. The interpolation of moisture diagnostics along the individual trajectories allows to track the Lagrangian evolution in the paired $\{H_2O, \delta D\}$ phase space. Characteristic signals will be interpreted based on the idealized process curves for effects of air mass mixing and microphysical cloud and rain processes (see Fig. 3.2; Noone, 2012) as well as depending on the underlying hydrometeorological conditions. This provides knowledge about process occurrences along the backward trajectories, so that the isotopic composition at the Sahel can be interpreted by means of the respective air mass histories.

Chapter 6 targets on the third research goal. It builds on the results of Chap. 4 and 5 and performs an extensive documentation of the isotopic composition in mid-tropospheric water vapor during the WAM across different scales (intra-annual, seasonal and convective scale). This includes the joint analysis of $\{H_2O, \delta D\}$ pair distributions from remote sensing (IASI) and isotope-enabled models (ICON-ART_{iso} and COSMO_{iso}) as well as the trajectory results discussed in Chap. 5. The aim is to qualitatively examine to which extent the observed $\{H_2O, \delta D\}$ variability is linked to dynamical and microphysical processes associated with the WAM. Additional information about hourly and monthly precipitation distributions taken from the GPM IMERG datasets will support this analysis.

3.4.2. Overview of Used Datasets

As discussed above, this study will make use of various datasets that provide information about $\{H_2O, \delta D\}$ pair distributions in the West African troposphere. This is complemented by precipitation datasets as well as backward trajectories for the Sahelian troposphere. Table 3.2 provides an overview of all considered datasets, whose technical details will be presented in the following section. The datasets used consist of the following:

MUSICA IASI $\{H_2O, \delta D\}$ pair data The global and daily remotely sensed $\{H_2O, \delta D\}$ pair dataset based on Metop/IASI measurements serves as a common research item throughout the whole thesis. By using the MUSICA IASI retrieval results from Schneider et al. (2021c), an optimal estimation $\{H_2O, \delta D\}$ pair product with main focus on the mid-troposphere is generated (vertical grid spacing shown in Fig. 3.6). This product is available for all cloud-free retrieval data from October 2014 to June 2019 in form of two global maps of $\{H_2O, \delta D\}$ pair data per day (according to the overpass times of 09.30 and 21.30 local time).

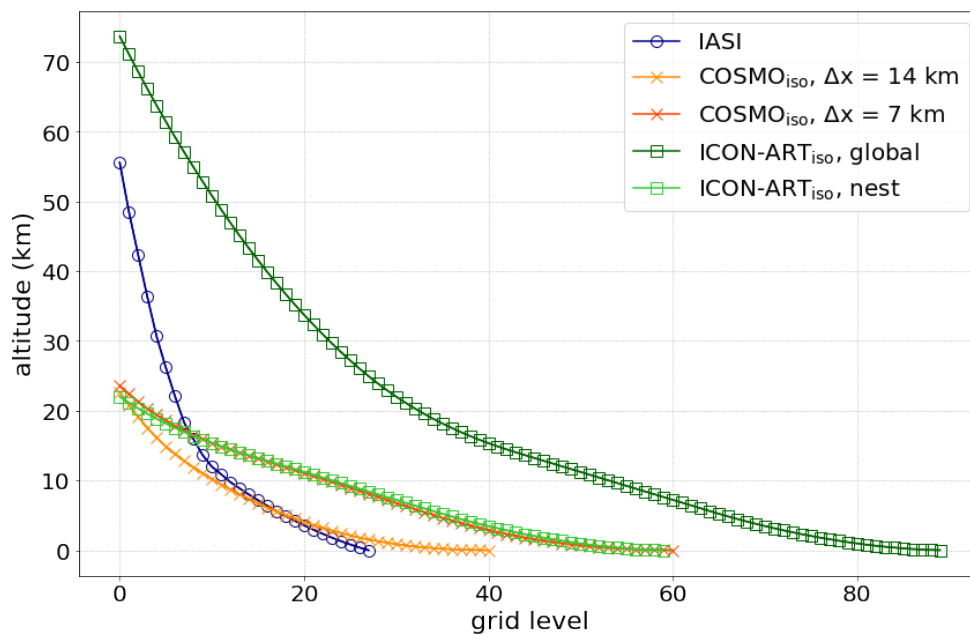


Figure 3.6. Vertical grid levels of the datasets from Table 3.2.

A detailed documentation of the post-processing and the resulting data availability is given in Chap. 4.

COSMO_{iso} simulations This study includes high-resolution, nudged simulations of the regional isotope-enabled model COSMO_{iso} (CISO1, CISO2 and CISO3, see Table 3.2) with a focus on West Africa during the WAM season 2016. These simulations were performed and provided by Andries de Vries (ETH, Zurich). The simulation periods are chosen to match the DACCIWA campaign (01 June – 31 July 2016; Knippertz et al., 2017) and the model output frequency was set to 1 h. Data provided by the global isotope-enabled model ECHAM5_{wiso} (Werner et al., 2011) are used as initial and boundary conditions as well as for a spectral nudging of the horizontal wind fields above 850 hPa. This serves to keep the meteorology close to reality, as the ECHAM5_{wiso} simulation was nudged to ERA-interim reanalyses provided by ECMWF. The model domain of the COSMO_{iso} simulations is chosen such that it covers the dominant moisture source regions of the WAM (see Fig. 3.7). For CISO1 and CISO2, the model has 40 vertical hybrid levels between the surface and 22.7 km (Fig. 3.6) and a horizontal grid spacing of 14 km (similar to the horizontal pixel size of Metop/IASI data, see Sect. 4.1). CISO3 employs 60 vertical levels up to 23.6 km and a horizontal resolution of 7 km (Fig. 3.6). Vergara-Temprado et al. (2020) stated that for a horizontal grid spacing below 25 km switching off the parameterization of deep convection leads to overall better results than increasing the horizontal resolution. Specifically for the WAM, various studies reported significant improvements when using explicit convection (Marshall et al., 2013,

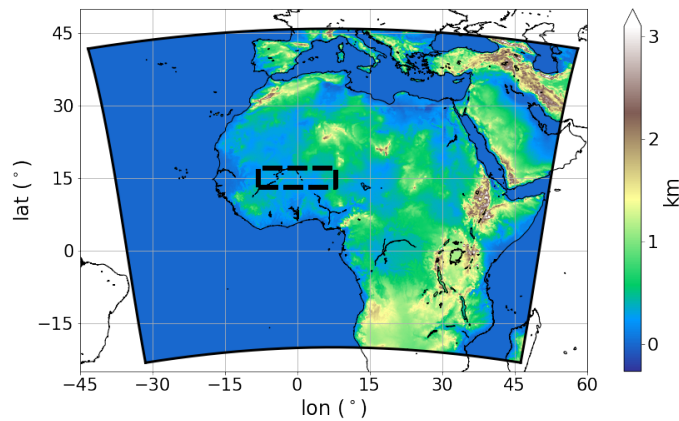


Figure 3.7. Geographical area of the model domain used for the simulation with COSMO_{iso}. The colors show the topography as considered in the COSMO_{iso} simulation. The black dashed box frames the target region over the Sahel, which was used to initialize the backward trajectories within LAGRANTO.

Maurer et al., 2017, Martínez and Chaboureau, 2018, Berthou et al., 2019, Crook et al., 2019, Pante and Knippertz, 2019). Based on these results, the parameterization schemes for deep and shallow convection are switched off for the simulations CISO2 and CISO3.

LAGRANTO trajectories Using the COSMO_{iso} dataset CISO2 as input for the LAGRANTO trajectory tool, backward trajectories are calculated as part of this study for the Sahelian mid-troposphere during the WAM season 2016. In accordance with the typical residence time of atmospheric water the trajectory length is set to 7 days (Sodemann, 2020). Trajectories are started daily at 09 and 21 UTC from 08 June to 30 July 2016, at 575 and 625 hPa and for approximately every 1° within the domain from 13° to 17° N, 8° W to 8° E (see Fig. 3.7). In total, this results in 12,720 trajectories. In addition to various meteorological variables, the specific contents q_x of H₂O and HDO in vapor (v), sedimenting (rain (r), and snow (s)) and non-sedimenting condensates (liquid (l) and ice (i) clouds) are interpolated onto the three-dimensional trajectory pathways. As the trajectory setup is chosen to match the characteristics of the remotely sensed {H₂O, δ D} pair data from Metop/IASI (see Sect. 4.1), q_v for H₂O and HDO is converted into volume mixing ratios (ppmv) and δ D is calculated along each trajectory. In the following, it will be referred to the starting point (first calculation step, day 0) as the target time and to the last calculation step (day -7) as the trajectory origin.

ICON-ART_{iso} simulations Additionally, the global isotope-enabled ICON-ART_{iso} is employed for providing multi-annual model-based {H₂O, δ D} pair data globally and with a focus on West Africa. As part of this work, ICON-ART_{iso} simulations are performed for different model configurations (IISO1–4). IISO1 employs a global grid with a horizontal grid size of

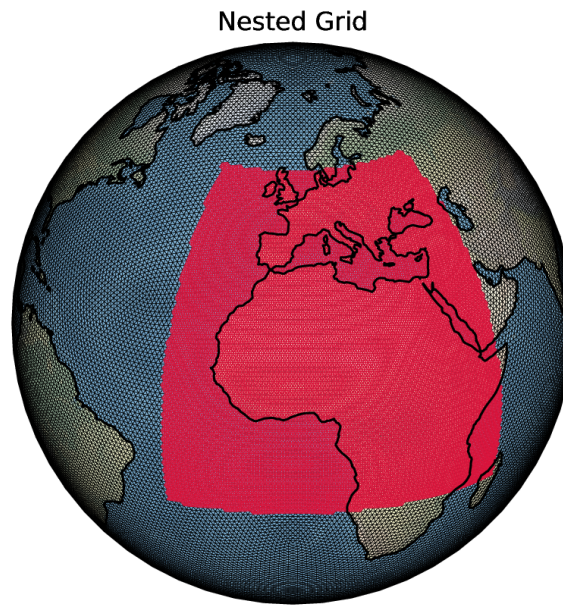


Figure 3.8. The global unstructured triangular grid of ICON for a horizontal grid size of 80 km (black grid lines) and the refined nested domain with a grid size of 40 km (red grid lines).

160 km and with 90 vertical hybrid levels from the surface up to 75 km (Fig. 3.6). The only difference to IISO2 is the grid size reduction to 80 km. IISO3 builds also on a global grid with 80 km grid size, but incorporates a refined and two-way-interacting nest that covers West Africa (see Fig. 3.8). The simulation for the nested domain IISO4 uses a 40 km grid with 60 vertical hybrid levels up to 22 km (Fig. 3.6). All four are free-running simulations, with a 3-year target period between 2017–2019 and model output generated twice per day (09 and 21 UTC, in agreement with the local overpass times of Metop/IASI over West Africa). Operational reanalysis data of the Integrated Forecasting System (IFS) from ECMWF are used for initializing the meteorological state, including water vapor, liquid and frozen cloud and rain water. The initialization of the moist variables specific for HDO is coupled to the initial H₂O state by considering climatological values of δD and applying those to the ICON fields. This includes three global mean δD values for water vapor (troposphere, tropopause and stratosphere) and one for the hydrometeors (Eckstein et al., 2018). To minimize the influence of this rough δD initialization, a spin-up time of 6 months is considered, i.e. each simulation is initialized on the 22 June 2016. As the horizontal resolution for all four simulations are far from resolving convection explicitly, the convection is computed via the parameterized schemes for deep and shallow convection (Zängl et al., 2015). Sea surface temperatures and sea ice cover are interpolated on a daily basis from monthly climatologies provided by the AMIP II project (Eckstein et al., 2018).

GPM IMERG precipitation data For obtaining information about actual precipitation distributions over West Africa, data are used from the Global Precipitation Measurement Mission (GPM) launched by NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace and Exploration Agency) (Huffman et al., 2014). GPM is an international satellite network that provides global precipitation estimates on an half-hourly base. Currently, it comprises data from twelve different microwave imagers and sounders as well as infrared measurements from six geostationary satellite series, all from different international space agencies. The Integrated Multi-satellitE Retrievals for GPM (IMERG) algorithm combines the precipitation measurements of all considered satellites and creates a calibrated precipitation estimate with a horizontal resolution of 0.1° every 30 minutes from 60°S to 60°N (for higher latitudes only the microwave data are available), starting from 2014 up to present. The resulting GPM IMERG precipitation data are distributed as half-hourly datasets (RGPM1). In addition, monthly averages (RGPM2) are derived and combined with the monthly surface precipitation gauge analysis of the Global Precipitation Climatology Centre (GPCC). For a full data description please refer to Huffman et al. (2019).

Part III.

Analyses

4 | MUSICA IASI {H₂O, δD} Pair Processing

As shown in Fig. 3.4, there is no publicly referenced global and continuous dataset including current measurements (after 2012) of water vapor and its isotopic composition in the troposphere from satellite-based remote sensing systems. However, such datasets are urgently needed for allowing meaningful analyses of the isotopic composition of atmospheric moisture. Satellite platforms can provide continuous measurements on the global scale with a high spatio-temporal resolution, while ground-based measurements are fixed to one location.

Therefore, in this section the aim is to present a novel dataset of tropospheric water isotopes by using measurements from the thermal infrared sensor IASI onboard the Metop satellites (introduced in Sect. 3.2.1). Currently, IASI is carried by the three satellites Metop-A, Metop-B and Metop-C as part of the EUMETSAT Polar System (EPS) programme and its successor IASI-NG (IASI Next Generation) will be part of the second generation of the EPS, which aims to operate follow-up Metop satellites for another 20 years (see Fig. 4.1). Based on the full swath width of the IASI sensors, this mission is able to provide a global scan of the atmosphere multiple times per day. As part of the MUSICA project (MUlti-platform remote Sensing of Isotopologues for investigating the Cycle of Atmospheric water; Schneider et al., 2016), a retrieval processor was developed that individually retrieves vertical profiles of H₂O and δD as well as additional interfering trace gases (Schneider and Hase, 2011). This processor was further improved and applied to all cloud-free IASI observations from October 2014 to June 2019 (processor and product description in Schneider et al., 2021c). However, due to the individual retrieval processing of H₂O and δD, their vertical sensitivities can still differ. For this purpose, Schneider et al. (2012) suggested a further post-processing of the H₂O and δD products. This post-processing generates {H₂O, δD} pairs, i.e. H₂O and δD data products having the same vertical sensitivities.

Within the thesis, the post-processing method of Schneider et al. (2012) is adapted and slightly modified in order to increase the sensitivity for dry atmospheric conditions. Together with the supply of diagnostic flag variables, this modified post-processing for creating optimal {H₂O, δD} pair information is applied to all MUSICA IASI retrieval results from Schneider et al. (2021c). This section serves to provide details about the corresponding processing as well as to document the resulting dataset in terms of spatio-temporal data coverage. Additionally, it includes the

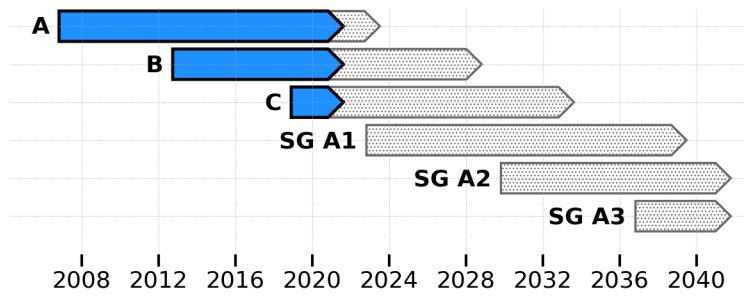


Figure 4.1. Timeline of the Metop satellites, which are currently orbiting (A,B and C) and which are planned in the near future (SG A1, SG A2 and SG A3). The blue arrows indicate the periods during which the satellites have successfully been operated and the gray dotted arrows mark the expected operating periods. The satellites Metop-A, Metop-B and Metop-C represent the first generation of the EUMETSAT Polar System (EPS) program, while its successor program EPS Second Generation (EPS SG) comprises the Metop-SG A1, Metop-SG A2 and Metop-SG A3 satellites.

development of a new MUSICA IASI retrieval simulator and the evaluation of its effects for direct model-to-satellite comparisons of COSMO_{iso} data with the new MUSICA IASI $\{H_2O, \delta D\}$ pair product.

This section is structured as follows: Section 4.1 provides some general information about the IASI sensor. The general structure and radiative transfer algorithm of the MUSICA IASI retrieval is described in Sect. 4.2. Section 4.3 shares details of the processing and results specific for the optimal estimation $\{H_2O, \delta D\}$ pair product. Finally, Sect. 4.4 discusses the new MUSICA IASI retrieval simulator.

4.1. Main Sensor Characteristics

IASI is a Fourier transform spectrometer measuring the thermal infrared upwelling radiation that is affected by atmospheric emission, absorption and scattering processes. Its spectral resolution is 0.5 cm^{-1} . The polar sun-synchronous orbits are designed such that the satellites overpass the equator at 09.30 and 21.30 local time. With around 14–15 orbits per satellite per day and a full swath width of 2200 km, each IASI sensor achieves a global scan of the atmosphere twice daily. The launches of Metop-A, -B and -C took place in 2006, 2012 and 2018, respectively. The expected lifetime of each satellite is 5 years. Currently, all three Metop satellites are successfully operating in orbit. An overview of the sensor characteristics is given in Table 4.1 (Clerbaux et al., 2009).

Table 4.1. Details about the Metop mission, the sensor IASI (Clerbaux et al., 2009) and the MUSICA IASI $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair processing.

Type	Specification
Launch dates	Metop A: October 19, 2006 Metop B: September 17, 2012 Metop C: November 7, 2018
Altitude	~ 817 km
Orbit type	polar, sun-synchronous
Local overpass times	descending orbit: ~ 09.30 local time ascending orbit: ~ 21.30 local time
IASI sensor	Fourier transform spectrometer
Spectral range of sensor	645–2760 cm^{-1}
Spectral range of MUSICA retrieval	1190–1400 cm^{-1}
Spectral resolution	0.5 cm^{-1}
Horizontal resolution	12 km at nadir
Full swath width	~ 2200 km
Orbit rate	~ 14 orbits per day and sensor
Global earth coverage	2 per day and sensor
MUSICA retrieval rate	~ 350 000 retrievals per day and sensor
MUSICA target species	$\{\text{H}_2\text{O}, \delta\text{D}\}$ pairs, CH_4 , N_2O , HNO_3
MUSICA retrieval grid	29 altitude levels between 0 and 55.6 km

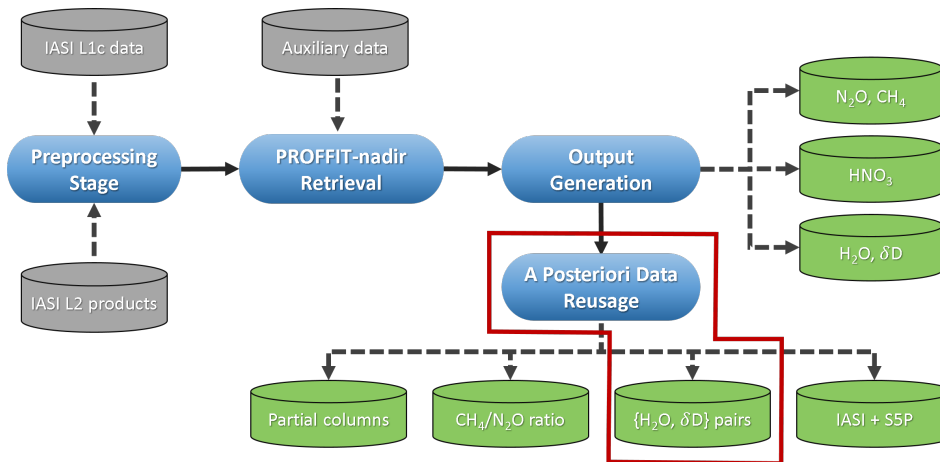


Figure 4.2. Overview of the full MUSICA IASI processing chain and its output products. The red frame indicates the part that is developed within the thesis. Further information about this processing chain can be found in Schneider et al. (2021c).

4.2. General Retrieval Processing

This section presents the main structure and the general radiative transfer concept of the MUSICA IASI retrieval (see Fig. 4.2). The retrieval properties specific to H_2O and δD will be discussed in Sect. 4.3.

During the pre-processing stage the input data required for the retrieval are collected, e.g. the measured spectra (EUMETSAT IASI L1C) and auxiliary products (EUMETSAT IASI L2 products), and filtered according to EUMETSAT data flags (Sect. 4.2.1). The used retrieval algorithm PROFFIT-nadir is the core of the retrieval, where an optimal estimation method retrieves the atmospheric state (Sect. 4.2.2 and Sect. 4.2.3). The final retrieval results are then stored as NetCDF4 files that are in agreement with the CF metadata naming conventions (Version 1.7).

4.2.1. Preselection of Observations

Within the IASI L2 products, EUMETSAT shares a set of diagnostic flags that allow for a user-friendly sorting of the corresponding IASI L1C spectra. The pre-processing (see Fig. 4.2) stage makes use of some of these in order to perform an initial filtering prior to the actual MUSICA IASI retrieval processing (Schneider et al., 2021c).

In its current version, the MUSICA IASI retrieval requires the measured scenes to be cloud-free. Therefore, a corresponding cloudiness assessment flag provided by EUMETSAT in the IASI L2 products is used to select observations, where the instrumental field of view is clear or includes only small cloud contaminations.

Further, the IASI L2 dataset also includes temperature profiles that are generated by EUMETSAT

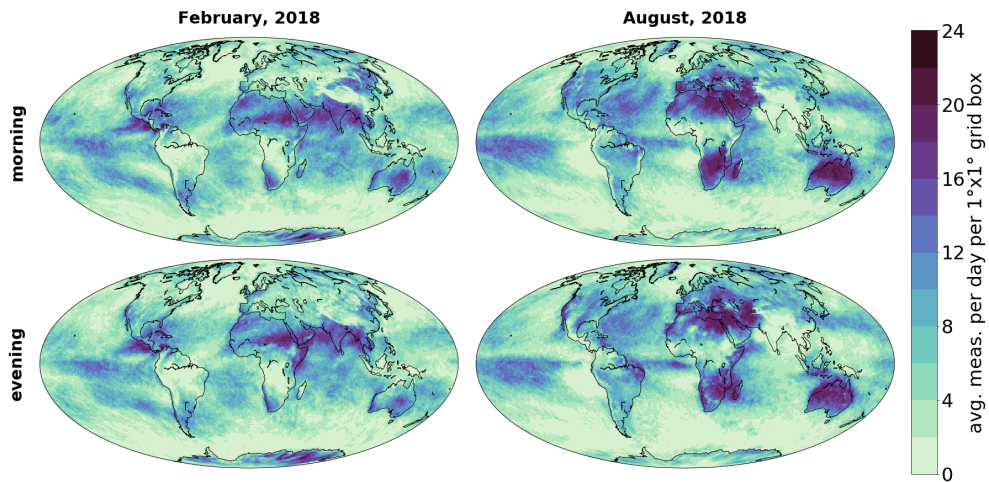


Figure 4.3. Monthly averaged amount of available IASI observations that pass the required filter conditions prior to the actual MUSICA IASI retrieval processing (see Sect. 4.2.2). Data are shown for February and August 2018 and evaluated on a $1^\circ \times 1^\circ$ grid.

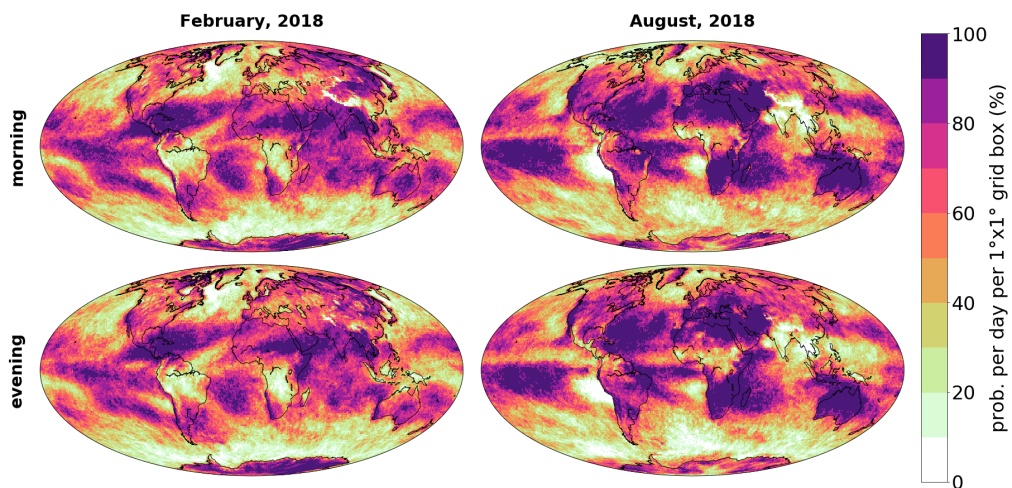


Figure 4.4. Similar to Fig. 4.3, but here showing the probability for obtaining at least one valid IASI measurement per day and per $1^\circ \times 1^\circ$ grid box after the initial filtering (see Sect. 4.2.2).

with an optimal estimation retrieval scheme. An additional IASI L2 flag indicating the quality of the respective temperature retrieval serves to filter those observations where the retrieval converged or the sounding was accepted by EUMETSAT.

An overview about the availability of valid IASI observations that results after the pre-selection is given in Fig. 4.3 (averaged amount) and Fig. 4.4 (probability for having at least one observation per day and $1^\circ \times 1^\circ$ grid box). The horizontal coverage mainly reflects the mean cloud distributions. Regions that are mostly cloud-free exhibit the highest numbers of valid observations, in particular for sub-tropical land areas (e.g. Australia, North and South Africa during boreal summer). In contrast, if cloudy conditions prevail, this leads to low numbers of valid observations, for instance for monsoon conditions (West African Monsoon and Indian Monsoon during August), and over cold sea surfaces (Southern Oceans, Eastern Pacific, Eastern Atlantic and Eastern Indian Ocean during respective winter times).

4.2.2. PROFFIT-nadir Retrieval

The MUSICA IASI processor uses the thermal infrared nadir retrieval algorithm PROFFIT-nadir (Schneider and Hase, 2011, Wiegeler et al., 2014) that is an extension of the PROFFIT algorithm (PROFile FIT, Hase et al., 2004). For the MUSICA IASI processing, it performs an optimal estimation of the trace gases H_2O , δD , CH_4 , N_2O , CO_2 and HNO_3 as well as atmospheric and surface temperature (Wiegeler et al., 2014, Schneider et al., 2016, Borger et al., 2018, García et al., 2018, 2020, Schneider et al., 2021c). By using the line-by-line radiative transfer model PRFFWD (PRoFit ForWarD model; Hase et al., 2004, Schneider and Hase, 2009), it simulates how the thermal infrared upwelling radiation passes through the atmosphere and is affected by emission and absorption (scattering processes are not relevant in the considered spectral window and is therefore not considered in PRFFWD). The aim is to estimate the vertical profiles of the considered absorbers such that the resulting simulated spectrum agrees with the measured IASI spectrum (see Fig. 4.5a). This procedure is based on well-established retrieval methods as described in Rodgers (2000) and will be shortly presented below.

The forward relation between the atmospheric states \mathbf{x} (atmospheric trace gases and temperature profiles) and the measured IASI spectrum \mathbf{y} can be written as follows:

$$\mathbf{y} = F(\mathbf{x}, \mathbf{b}) \quad (4.1)$$

with \mathbf{b} comprising auxiliary parameters such as the surface emissivity and instrumental characteristics. The forward function F represents the radiative transfer model (here PRFFWD) and simulates the spectrally depending interactions between each trace gas and the upwelling radiation. This is based on spectroscopic line parameters that indicate, according to individual line strengths, the spectroscopic properties (e.g. molecular absorption) of each trace gas at different

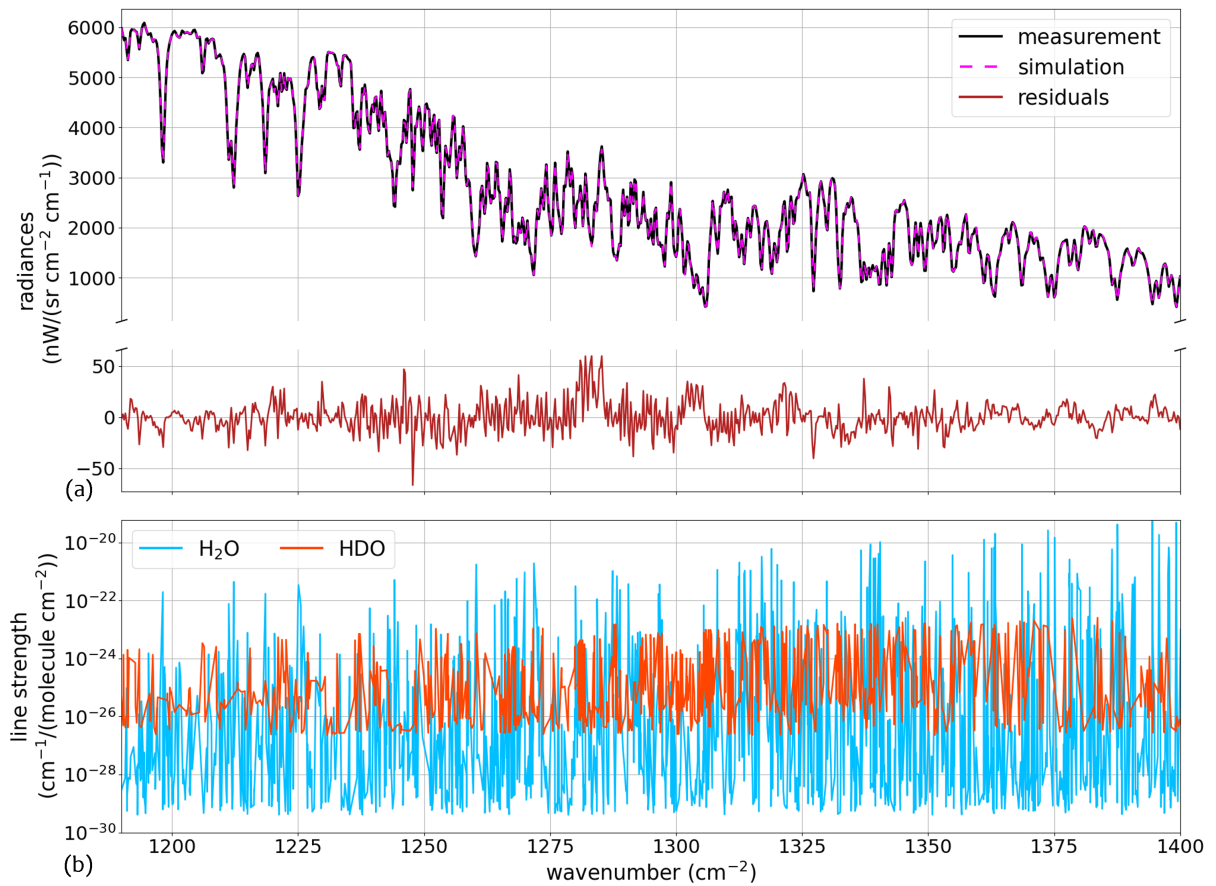


Figure 4.5. (a) Example of a spectrum measured by IASI over a tropical ocean (black line) and the simulated spectrum of the MUSICA IASI retrieval (magenta dashed line). The residual between the measured and simulated spectra is shown red. (b) Line strengths of H_2O and HDO lines within the spectral window of the MUSICA IASI retrieval, taken from the HITRAN2016 database.

wavelengths. During the MUSICA IASI retrieval, the spectroscopic lines of the HITRAN2016 molecular spectroscopic database are used (Gordon et al., 2017), with a modification of the HDO line strengths according to Schneider et al. (2021c). An example is given in Fig. 4.5b for the spectroscopic lines of H_2O and HDO , as together they are found to most strongly affect the radiances in the spectral window of the MUSICA IASI retrieval (see Table 4.1; Schneider et al., 2021c). In particular for wave numbers between $1210\text{--}1255 \text{ cm}^{-1}$, HDO exhibits spectroscopic features that are stronger than H_2O and lead to a distinct impact on the spectrum. Also around 1300 cm^{-1} there are various HDO lines with stronger line strengths than H_2O , but this spectral region contains also prevailing spectroscopic CH_4 and N_2O signatures.

Further, PRFFWD considers the spectrally depending efficiency with which the Earth's surface emits its stored energy as thermal radiation, the so-called surface emissivity (Njoku, 2014). Over land, the emissivity values are provided within the EUMETSAT IASI L2 products, whereas for ocean observations the emissivity values of Masuda et al. (1988) and for observations over ice

surfaces the ASTER spectral library (Baldrige et al., 2009) are both used.

For solving Eqn. 4.1 with respect to the targeted absorber profiles \mathbf{x} , an inversion of F is required. However, as different atmospheric states \mathbf{x} can lead to similar spectra \mathbf{y} , this is an under-determined problem and the direct calculation of F^{-1} is generally not possible.

For this purpose, the MUSICA IASI retrieval restricts the solution space by defining a cost function J that considers additional a priori information of the atmospheric state (Schneider and Hase, 2011, Borger et al., 2018):

$$J = [\mathbf{y} - F(\mathbf{x}, \mathbf{b})]^T \mathbf{S}_{\mathbf{y}, \text{noise}} [\mathbf{y} - F(\mathbf{x}, \mathbf{b})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{R} [\mathbf{x} - \mathbf{x}_a] \quad (4.2)$$

The first term indicates the difference between the measured (\mathbf{y}) and simulated ($F(\mathbf{x}, \mathbf{b})$) spectrum with taking into account the actual measurement noise covariance $\mathbf{S}_{\mathbf{y}, \text{noise}}$. The right term of Eqn. (4.2) is the regularization term which constraints the retrieved state \mathbf{x} towards an a priori assumed state \mathbf{x}_a . The kind and strength of the constraint is given by the corresponding constraint matrix \mathbf{R} , for which the MUSICA IASI retrieval uses an approximated inversion of the a priori covariance matrix \mathbf{S}_a (more details in Schneider et al., 2021c):

$$\mathbf{R} \approx \mathbf{S}_a^{-1} \quad (4.3)$$

Minimizing the cost function (4.2) creates the constrained solution $\hat{\mathbf{x}}$. However, as Eqn. (4.2) is non-linear, its minimization has to be approximated iteratively. For the $(i+1)$ th iteration this yields

$$\mathbf{x}_{i+1} = \mathbf{x}_a + \mathbf{G}_i [\mathbf{y} - F(\mathbf{x}_i, \mathbf{b}) + \mathbf{K}_i (\mathbf{x}_i - \mathbf{x}_a)] \quad (4.4)$$

with the gain matrix

$$\mathbf{G} = \left(\mathbf{K}^T \mathbf{S}_{\mathbf{y}, \text{noise}}^{-1} \mathbf{K} + \mathbf{R} \right)^{-1} \mathbf{K}^T \mathbf{S}_{\mathbf{y}, \text{noise}}^{-1}. \quad (4.5)$$

Here, \mathbf{K} is the Jacobian matrix, i.e. the derivatives indicating how the spectrum \mathbf{y} varies due to changes of the absorber \mathbf{x} . \mathbf{G} provides the derivatives that indicate how the new state \mathbf{x}_{i+1} varies due to changes in the measured spectrum \mathbf{y} . Then, by applying the forward model F on the new absorber state \mathbf{x}_{i+1} this creates the updated simulated spectrum, which can be used for the next iteration of Eqn. (4.4). This cycle is repeated, until the simulated spectrum converges against the measured spectrum. Convergence means in this context that the residuals between the simulated and measured spectrum remain stable for further iterations (see Fig. 4.5). The corresponding final state $\hat{\mathbf{x}}$ is the solution of the optimal estimation method.

For the retrieval processing a priori profiles (\mathbf{x}_a in Eqn. (4.2)) have to be assumed. While the EUMETSAT IASI L2 temperature profiles are used for setting up the a priori assumptions for the atmospheric and surface temperature, a simulation of the coupled chemistry-climate model CESM1/WACCM for the period 1979–2014 (Marsh et al., 2013) serves as reference for the a priori assumptions of the target trace gases. Specifically, for the vertical a priori profiles of H_2O ,

N_2O , CH_4 and HNO_3 a mean latitudinal dependence together with seasonal cycles is assumed. Additionally, ground-based (González et al., 2016, Christner et al., 2018) and airborne (Dyroff et al., 2015) in-situ measurements were used to generate a single global relation between H_2O and δD and thereby constructing the δD a priori assumptions based on those of H_2O (Schneider et al., 2021c).

4.2.3. Vertical Representativeness of Retrieval Results

Using the gain matrix \mathbf{G} and the Jacobian matrix \mathbf{K} , the so-called averaging kernel matrix \mathbf{A} is defined as

$$\mathbf{A} = \mathbf{G}\mathbf{K} \quad (4.6)$$

and characterizes the relation between the retrieved state $\hat{\mathbf{x}}$, the true atmospheric state \mathbf{x}_{atm} and the a priori state \mathbf{x}_a :

$$\hat{\mathbf{x}} = \mathbf{A}(\mathbf{x}_{\text{atm}} - \mathbf{x}_a) + \mathbf{x}_a \quad (4.7)$$

In general, an averaging kernel (rows of \mathbf{A}) indicates the vertical sensitivity of the retrieved state $\hat{\mathbf{x}}$. Following the definition from Rodgers (2000), the averaging kernel depicts the fraction of the retrieved result coming from the retrieval itself and not from the a priori assumption. In case of a perfect retrieval the kernel matrix would equal the identity matrix, expressing total independence of the retrieval results from the chosen a priori state. Thus, the degree of deviation from unity quantifies the vertical information content of a remotely sensed observation.

For instance, a common metric for describing the vertical information content is the degree of freedom for signal (DOFS). It is defined as the trace of the averaging kernel matrix. The value of DOFS indicates the number of vertical structures that can be independently determined from an observation (Rodgers, 2000).

Further, the sum of the values along an individual averaging kernel is called measurement response (Eriksson, 2000, Baron et al., 2002). A measurement response of 1 implies that the retrieved state is a smoothed but unbiased image of the true atmospheric profile, whereas values deviating from unity are obtained, if the retrieval constraint deviates from pure smoothing (von Clarmann et al., 2020).

To examine the vertical resolution of a retrieved profile (i.e. the capability to detect vertical structures), the MUSICA IASI product provides metrics that characterize the vertical sensitivity, i.e. the shape of the averaging kernels. First, the relative position of the sensitivity weighted altitude compared to its nominal altitude is termed information displacement. For this, the centroid offset as defined by Backus and Gilbert (1970) is used in a discretized form (Keppens et al., 2015). And second, to describe the vertical smoothing of the retrieved state, the MUSICA

IASI retrieval provides two different diagnostics. The definition of Backus and Gilbert (1970) is used to create a kernel-weighted spread around the kernel centroid, while the data density reciprocal of Purser and Huang (1993) serves to indicate the layer width that covers a DOFS of 1. Discussions of these two metrics can be found in Keppens et al. (2015) and von Clarmann et al. (2020), and in the context of the MUSICA IASI full retrieval product in Schneider et al. (2021c).

4.3. Optimal Estimation $\{H_2O, \delta D\}$ Pair Product

This section provides an overview about the MUSICA IASI water vapor state vector as retrieved during the PROFFIT-nadir retrieval (Sect. 4.3.1) and the a posteriori generated $\{H_2O, \delta D\}$ pair product, as developed and improved within this study (Sect. 4.3.2). For the latter, this includes a detailed description of the underlying processing as well as a spatio-temporal characterization of the newly generated MUSICA IASI $\{H_2O, \delta D\}$ pair dataset.

4.3.1. Retrieval of H_2O and δD

The PROFFIT-nadir retrieval considers the spectral region between $1190\text{--}1400\text{ cm}^{-1}$. In this region, the water vapor isotopologues H_2O and HDO reveal strong spectroscopic lines (see Fig. 4.5). Besides those, there are interfering spectroscopic features, such as of CH_4 , N_2O and HNO_3 , in this spectral window (Clerbaux et al., 2009, Herbin et al., 2009). The vertical retrieval grid consists of 29 altitude levels between the surface and 55.6 km (Fig. 3.6). In the lower troposphere the vertical distance of the different grid points is around 400 m and increases beyond 5 km in the stratosphere.

By considering the logarithmic scale for the individual target species, the individual retrieval of H_2O and HDO can be approximated to a δD retrieval (Schneider et al., 2006, Worden et al., 2006). Since $\ln[HDO] - \ln[H_2O] = \ln[HDO/H_2O]$, additional constraints for the ratio HDO/H_2O can be considered. Using this relation, the MUSICA IASI retrieval performs actually an optimal estimation for the proxy base $\{(\ln[H_2O] + \ln[HDO])/2, (\ln[HDO] - \ln[H_2O])\}$ (Schneider and Hase, 2011). Schneider et al. (2012) showed that these two state vectors constitute reliable proxies for variations in H_2O and δD .

Thus, the full MUSICA IASI state vector is defined as

$$\mathbf{x} = \left(\mathbf{x}_{wv}^T \quad \mathbf{x}_{n2o}^T \quad \mathbf{x}_{ch4}^T \quad \mathbf{x}_{hno3}^T \quad \mathbf{x}_t^T \right)^T \quad (4.8)$$

with \mathbf{x}_{wv} representing the two water vapor proxy base vectors $(\ln[H_2O] + \ln[HDO])/2$ and $(\ln[HDO] - \ln[H_2O])$. \mathbf{x}_t is the atmospheric temperature state.

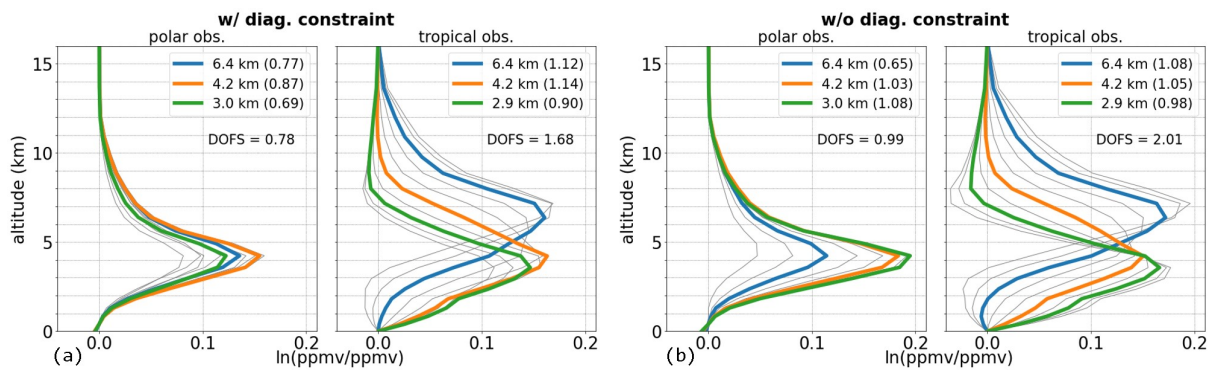


Figure 4.6. Row averaging kernels for the $\{H_2O, \delta D\}$ pair product with diagonal constraint (a) and without diagonal constraint (b). The kernels are shown for a polar observation ($83.5^\circ N, 147.0^\circ W$, left panels in a and b) and for a tropical observation ($4.8^\circ N, 45.8^\circ W$, right panels in a and b), both above oceans and measured on the 01 July 2017. The measurement response values for the kernels at 3.0, 4.2 and 6.4 km above sea level are given in the respective parentheses in the legends.

4.3.2. Post-Processing for Optimal $\{H_2O, \delta D\}$ Pair Information

By considering the individual MUSICA IASI retrieval results for H_2O and δD , this study develops and applies a further post-processing that aims at generating an optimal estimation $\{H_2O, \delta D\}$ pair product. The following sections provide details about the corresponding processing, including information about the error treatment and data selection according to data quality.

4.3.2.1. Generation of the $\{H_2O, \delta D\}$ Pair Product

Due to its high variability in the troposphere, H_2O can be detected very well in contrast to δD , which varies only weakly. As the MUSICA IASI retrieval produces an individual optimal estimation for the different target states, this generally results in different averaging kernels for H_2O and δD . Therefore, the vertical sensitivity of H_2O is much higher than for δD (Schneider et al., 2021c). However, an optimal analysis of $\{H_2O, \delta D\}$ pairs would require similar characteristics in the vertical sensitivities to ensure that the retrieved H_2O and δD refer to the same vertical structures. Therefore, Schneider et al. (2012) proposed a post-processing for harmonizing the vertical information contents of individually retrieved H_2O and δD . This is achieved by reducing the strength of the averaging kernels of H_2O with respect to those from δD . The resulting improvement has been demonstrated for IASI results against long-term datasets from ground-based remote sensing stations and in-situ aircraft measurements (Wiegele et al., 2014, Schneider et al., 2015).

As a first step towards harmonizing the sensitivities of H_2O and δD , the water vapor state vector $\hat{\mathbf{x}}_{wv}$ as well as the corresponding averaging kernel block matrix \mathbf{A}_{wv}

$$\hat{\mathbf{x}}_{wv} = \begin{pmatrix} \hat{x}_{wv,1} \\ \hat{x}_{wv,2} \end{pmatrix}, \mathbf{A}_{wv} = \begin{pmatrix} \mathbf{A}_{wv,11} & \mathbf{A}_{wv,12} \\ \mathbf{A}_{wv,21} & \mathbf{A}_{wv,22} \end{pmatrix} \quad (4.9)$$

are transformed from the $\{\ln[H_2O], \ln[HDO]\}$ basis system to the water vapor proxy base $\{(\ln[H_2O] + \ln[HDO])/2, (\ln[HDO] - \ln[H_2O])\}$:

$$\hat{\mathbf{x}}'_{wv,a} = \mathbf{P}\hat{\mathbf{x}}_{wv,a} \quad (4.10)$$

$$\hat{\mathbf{x}}'_{wv} = \mathbf{P}\hat{\mathbf{x}}_{wv} \quad (4.11)$$

$$\mathbf{A}'_{wv} = \mathbf{P}\mathbf{A}_{wv}\mathbf{P}^{-1} \quad (4.12)$$

In the following, primed variables are consistently referring to the water vapor proxy state base. Detailed information about the transformation operator \mathbf{P} can be found in Schneider et al. (2012), Wiegele et al. (2014) and Barthlott et al. (2017). $\hat{\mathbf{x}}'_{wv}$ is in fact the state that is optimally estimated by the MUSICA IASI retrieval and it represents proxies for H_2O and δD .

The following step harmonizes the differing sensitivities of the water vapor proxy states by reducing the sensitivity of the H_2O proxy to the sensitivity of the δD proxy. The a posteriori correction operator \mathbf{C}

$$\mathbf{C}' = \begin{pmatrix} \mathbf{A}'_{wv,22} & \mathbf{0} \\ -\mathbf{A}'_{wv,21} & \mathbf{I} \end{pmatrix} \quad (4.13)$$

serves to create the harmonized product

$$\mathbf{A}^*_{wv} = \mathbf{C}'\mathbf{A}'_{wv} \quad (4.14)$$

$$\hat{\mathbf{x}}^*_{wv} = \mathbf{C}'(\hat{\mathbf{x}}'_{wv} - \hat{\mathbf{x}}'_{wv,a}) + \hat{\mathbf{x}}'_{wv,a} \quad (4.15)$$

which is called Type 2 product in Schneider et al. (2012). The main property of $\hat{\mathbf{x}}^*_{wv}$ is that it provides profiles for H_2O and δD having practically the same averaging kernels. This allows meaningful analyses of paired $\{H_2O, \delta D\}$ distributions.

Figure 4.6a illustrates the row kernels from \mathbf{A}^*_{wv} for two exemplary $\{H_2O, \delta D\}$ pair results above a polar and a tropical site. The main sensitivity lies in the free troposphere, with peaks at ~ 4 km above sea level (a.s.l.). The tropical observation shows additional sensitivity at 6–7 km a.s.l. The kernels for the polar observation exhibit values of measurement response and DOFS falling below 1 by up to 30 % (see values in the parentheses in the legend), meaning there is low vertical sensitivity. In contrast, the $\{H_2O, \delta D\}$ pair kernels for the tropical observation exhibit a DOFS of 1.68, though its measurement response still deviates from 1 by about 10–14 %. The next section introduces a method that increases the DOFS and the measurement response a posteriori (the corresponding kernels are shown in Fig. 4.6b, but will be discussed throughout the following section).

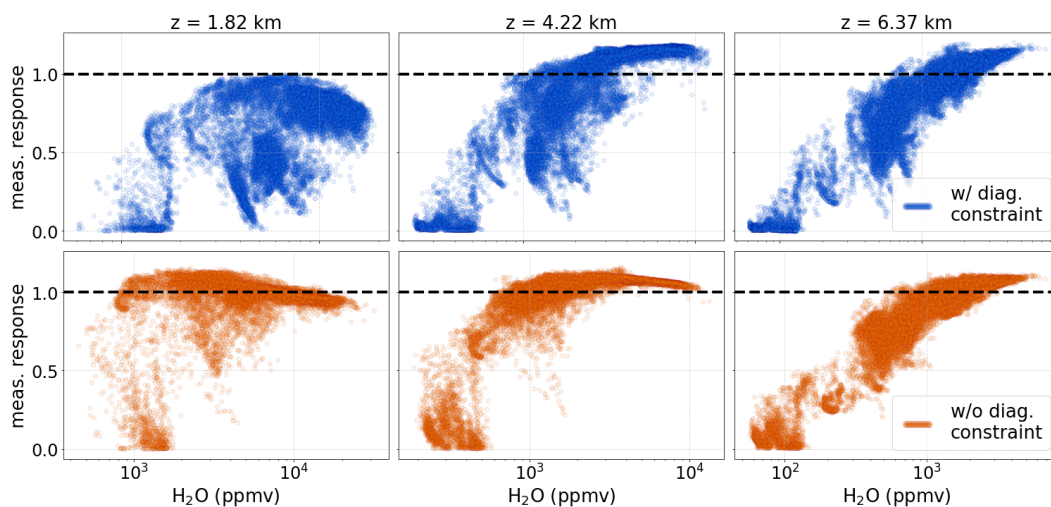


Figure 4.7. Measurement response for the MUSICA IASI $\{H_2O, \delta D\}$ pair product along Metop-A orbit 55524 on 01 July 2017, for the original product (blue scatter) and the improved one (orange scatter). The former includes the diagonal retrieval constraint, while it is removed for the latter. Results are shown for 1.8, 4.2 and 6.4 km above sea level.

4.3.2.2. Optimization of Retrieval Regularization

By inspecting the row kernels in \mathbf{A}_{wv}^* for a full orbit, it becomes apparent that there is a general and non-negligible deficit in the measurement response of the $\{H_2O, \delta D\}$ pairs. The blue dots in Fig. 4.7 show that, for instance, at 1.8, 4.2 and 6.4 km a.s.l. and for moisture contents below 10^4 ppmv a large amount of data contain measurement response values far below 1.

As emphasized in Sect. 4.2.3, the measurement response is a metric for the influence of the a priori assumptions on the retrieval result. Thus, a too low measurement response can be an indicator for a too strong retrieval constraint that excessively pulls the retrieved states towards the a priori profiles (Rodgers, 2000, von Clarmann et al., 2020). Therefore, to reduce the observed lack of sensitivity in the $\{H_2O, \delta D\}$ pairs, a method for a posteriori modifying and reducing the underlying retrieval constraint is applied. For this purpose, the following section adapts a linear optimal estimation method from Rodgers and Connor (2003) that creates a best estimate of a given retrieval result with regards to a new constraint:

$$\mathbf{M}' = \mathbf{R}'_d{}^{-1} \mathbf{A}'^T (\mathbf{A}' \mathbf{R}'_d{}^{-1} \mathbf{A}'^T + \mathbf{S}'_{\hat{\mathbf{x}}, \text{noise}})^{-1} \quad (4.16)$$

The purpose of the operator \mathbf{M}' is that it allows a modification of the retrieval solution $\hat{\mathbf{x}}$ and its kernels \mathbf{A}' according to a weaker constraint. \mathbf{R}'_d is the regularization matrix chosen according to the desired constraint.

In general, a regularization restricts the variability of a retrieved state vector in order to keep the retrieval solution within the range of physically realistic profiles (Phillips, 1962, Tikhonov, 1963, Rodgers, 2000). By reducing the regularization strength, it increases the allowed variability

for the retrieved states. As a consequence, the information content increases. On the downside, a weaker constraint causes larger noise and can produce information that is not provided by the measurement (Rodgers, 2000). Therefore, the remainder of this section discusses the optimal definition of the input matrices R'_d , A' and $S'_{\hat{x},\text{noise}}$ as well as the correct usage of M' in order to enhance the sensitivity of the $\{H_2O, \delta D\}$ pairs.

First of all, to achieve the full benefit from the matrix operations in Eq. (4.16), the full MUSICA IASI state for the kernel matrix A' is considered as this is also used during the original retrieval processing. This means that the non-harmonized water vapor proxy state A'_{wv} from Eq. (4.12) and the interfering effects of the other retrieval state vectors are taken into account. Since the retrieval output from Schneider et al. (2021c) provides only the dominant averaging kernels and cross-correlations, A' is composed as follows:

$$A' = \begin{pmatrix} A'_{wv,11} & A'_{wv,12} & 0 & 0 & 0 & A'_{t,wv,1} \\ A'_{wv,21} & A'_{wv,22} & 0 & 0 & 0 & A'_{t,wv,2} \\ 0 & 0 & A_{ghg,11} & A_{ghg,12} & 0 & A_{t,ghg,1} \\ 0 & 0 & A_{ghg,21} & A_{ghg,22} & 0 & A_{t,ghg,2} \\ 0 & 0 & 0 & 0 & A_{hno3} & A_{t,hno3} \\ 0 & 0 & 0 & 0 & 0 & A_t \end{pmatrix} \quad (4.17)$$

The kernels of N_2O and CH_4 are denoted as $A_{ghg,11}$ $A_{ghg,22}$, respectively. The indices 21 and 12 indicate the respective cross-dependencies. The cross-dependency of the temperature retrieval to the water vapor proxy states is marked with $A'_{t,wv,1}$ and $A'_{t,wv,2}$. The entries, for which the corresponding kernel matrix are not provided, are filled with the null matrices 0 .

Further, the retrieval noise error covariance $S'_{\hat{x},\text{noise}}$ is derived, i.e. the variability in the measured radiances that was not explained during the retrieval processing (corresponding to the spectral residuals in Fig. 4.5a). $S'_{\hat{x},\text{noise}}$ can be calculated from A' and the regularization matrix R' that was originally applied during the retrieval (Rodgers, 2000, Schneider et al., 2021c):

$$S'_{\hat{x},\text{noise}} = A'(I - A')R'^{-1} \quad (4.18)$$

Now the question arises about the choice of a new meaningful regularization matrix R'_d . For this purpose, it is useful to first recapitulate the original MUSICA approach for setting up the retrieval constraint. For each target species an individual covariance matrix S'_a is given that describes the potential departure of the retrieval solution from the a priori state. This depends on the choice of the height-dependent correlation length, the a priori assumed size of vertical structures that may be resolvable (Schneider et al., 2021c). By inverting S'_a this yields the corresponding

regularization matrix \mathbf{R}' . During the MUSICA IASI retrieval the inversion of \mathbf{S}'_a is realized by a decomposition into its diagonal and derivative values (Hase et al., 2004, Schneider et al., 2021c):

$$\begin{aligned} \mathbf{R}' = & (\boldsymbol{\alpha}_0 \mathbf{L}_0)^T \boldsymbol{\alpha}_0 \mathbf{L}_0 + (\boldsymbol{\alpha}_1 \mathbf{L}_1)^T \boldsymbol{\alpha}_1 \mathbf{L}_1 \\ & + (\boldsymbol{\alpha}_2 \mathbf{L}_2)^T \boldsymbol{\alpha}_2 \mathbf{L}_2 \end{aligned} \quad (4.19)$$

with α_i as the strength of the individual constraining terms and \mathbf{L}_i as the constraint operators. The diagonal matrices α_i are derived from \mathbf{S}'_a and are provided for each target state as output variables from the MUSICA IASI retrieval (Schneider et al., 2021c). \mathbf{L}_0 represents the diagonal constraint operator and equals the identity matrix \mathbf{I} . Its effect is to shift the retrieved profile towards the a priori profile. \mathbf{L}_1 and \mathbf{L}_2 are the first- and second-order derivative operators and constrain the retrieved profile according to the shape of the a priori profile, thereby representing smoothing constraints. For the retrieval of atmospheric trace gases with weak spectroscopic signatures smoothing constraints can be advantageous over \mathbf{L}_0 , because a diagonal constraint tightens the retrieval by means of the absolute a priori values, with potentially inducing a bias in the retrieval (e.g. Steck, 2002). Therefore, the suggestion is that the consideration of the diagonal constraint in Eq. (4.19) causes the observed sensitivity lack in the $\{H_2O, \delta D\}$ pair data for dry conditions. Following this hypothesis, the herein applied approach is to remove the diagonal constraint operators for the water vapor states and create the new weaker regularization matrix $\mathbf{R}'_{\text{wv,d}}$ as follows:

$$\mathbf{R}'_{\text{wv,d}} = (\boldsymbol{\alpha}_1 \mathbf{L}_1)^T \boldsymbol{\alpha}_1 \mathbf{L}_1 + (\boldsymbol{\alpha}_2 \mathbf{L}_2)^T \boldsymbol{\alpha}_2 \mathbf{L}_2 \quad (4.20)$$

Keeping the regularization matrices of the other target states unchanged, the new regularization matrix \mathbf{R}'_d can be built for the full MUSICA state. With that, Eq. (4.16) is fully determined and now \mathbf{M}' can be used to adjust the kernel matrix \mathbf{A}' according to the new constraint \mathbf{R}'_d :

$$\mathbf{A}'_m = \mathbf{M}' \mathbf{A}' \quad (4.21)$$

Based on the modified kernel matrix \mathbf{A}'_m , the optimal estimation $\{H_2O, \delta D\}$ pair product can now be created for the constraint reduced state. By extracting $\mathbf{A}'_{\text{wv,m}}$ as the first 2×2 block from \mathbf{A}'_m , the new a posteriori operator \mathbf{C}'_m is calculated analogous to Eqn. (4.13). This allows for deriving the constraint reduced pair product:

$$\mathbf{A}^*_{\text{wv,m}} = \mathbf{C}'_m \mathbf{A}'_{\text{wv,m}} \quad (4.22)$$

$$\hat{\mathbf{x}}^*_{\text{wv,m}} = \mathbf{C}'_m \mathbf{M}' (\hat{\mathbf{x}}'_{\text{wv}} - \hat{\mathbf{x}}'_{\text{wv,a}}) + \hat{\mathbf{x}}'_{\text{wv,a}} \quad (4.23)$$

This product $\mathbf{A}^*_{\text{wv,m}}$ with reduced constraint shows a clear increase of the measurement response (see lower panels of Figure 4.7). While the improvements are rather small for 6.4 km a.s.l., the results at 1.8 and 4.2 km a.s.l. have a much better measurement response for moisture contents

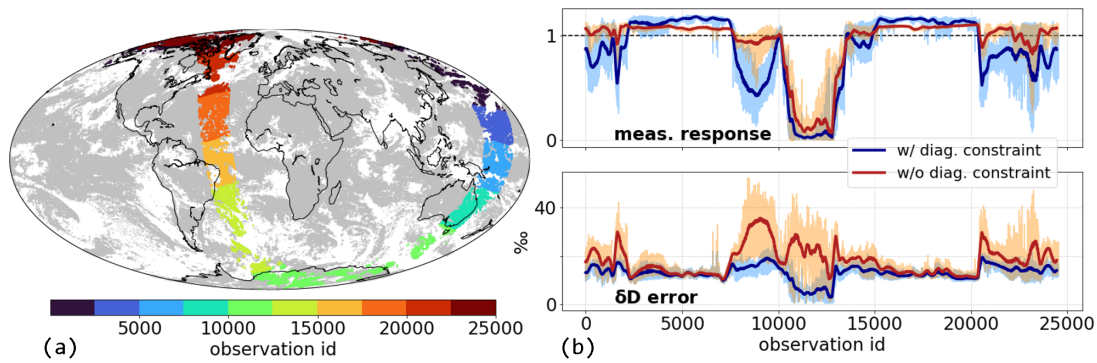


Figure 4.8. Time-series of different retrieval metrics of observations along the Metop-A orbit 55524. (a) shows the location of the observations along the orbit and colour-coded with the chronological observation numbering for this orbit (observation IDs). The gray scatter illustrate all cloud-free Metop-A and -B observations for 01 July 2017 that were considered for the $\{H_2O, \delta D\}$ pair post-processing. Panel (b) indicates the time-series of the measurement response for the $\{H_2O, \delta D\}$ pairs (upper panel) and the total δD error (lower panel), both for 4.2 km above sea level. Results for the original product and the product after reducing the retrieval constraint are shown in blue and red, respectively. The solid lines are running means over 200 observations.

above 700 ppmv.

The time-series of the measurement response along the orbit used in Fig. 4.7 is shown in Fig. 4.8b (upper panel). It is found that the constraint reduction leads to a general decrease of the deviation from 1. Over the Pacific and Atlantic Ocean (observation IDs of 2500–7500 and 15 000–20 000) there is a shift of the slightly over-estimated measurement response towards 1. In contrast, for higher latitudes its values are originally below 1, but increase significantly due to the constraint reduction. This is in particular pronounced for observations above Australia (observation IDs of 7500–10 000), where an averaged increase in the measurement response of up to 0.5 is apparent. Also for polar observations in the Northern Hemisphere (observation IDs of 0–2500 and 20 000–25 000) the measurement strongly improves.

Analogous improvements become apparent for the individual row kernels in Fig. 4.6 (compare Fig. 4.6a and b). The measurement response increases for the dry polar data at 3.0 and 4.2 km a.s.l. by 56 and 18 %, respectively. Also for the tropical site the measurement response values approach unity. These improvements are not at the expense of vertical resolution, instead they go along with respective improvements in the maximum amplitudes of the individual kernels as well as in the DOFS. For instance, the DOFS over the tropical site increases from 1.68 to 2.01, indicating that now information at two different altitude layers can be estimated independently.

4.3.2.3. Error Treatment

Several studies have intensively discussed the error treatment for satellite observations in general (Rodgers, 2000, von Clarmann et al., 2020) and with a focus on MUSICA IASI retrieval data (Schneider and Hase, 2011, Borger et al., 2018). Schneider et al. (2021c) provided an overview of the errors that result for the most recent MUSICA IASI retrieval. Along with the kernel modifications for reducing the diagonal constraint for water vapor, a respective processing is required for the dominant MUSICA IASI error covariances.

Given the error covariance $S'_{\hat{x}}$ in the proxy state base, the modified a posteriori operator C_m is used to transform $S'_{\hat{x}}$ according to the reduced constraint:

$$S'_{\hat{x},m} = C'_m M' S'_{\hat{x}} M'^T C'^T_m \quad (4.24)$$

This processing is performed for the retrieval noise error covariance $S'_{\hat{x},noise}$ from Eqn. (4.18) and for the temperature cross-covariance $S'_{\hat{x},temp.}$:

$$S'_{\hat{x},temp.} = A'_{t,wv} S'_{a,temp.} A'^T_{t,wv} \quad (4.25)$$

This strongly depends on the choice of the assumed a priori uncertainty covariance $S'_{a,temp.}$ (Schneider et al., 2021c).

As these two are the dominant errors for the MUSICA IASI δD product (Schneider et al., 2021c), their sum constitute an estimate of the total error covariance for the improved H_2O and δD states:

$$S'_{\hat{x},tot,m} = S'_{\hat{x},noise,m} + S'_{\hat{x},temp.,m} \quad (4.26)$$

The bottom panel in Fig. 4.8b illustrates how the total δD error changes due to the a posteriori constraint reduction. In general, with relaxing the regularization strength the retrieval noise will increase (e.g. Rodgers, 2000). Following this behaviour, the δD error exhibits a strong increase for areas where the impact of the regularization optimization is large and the measurement response increases. For instance, the strong improvements of the measurement response over the dry Australian desert are at the expense of increasing the averaged δD error by 20 ‰ with single peaks up to 50 ‰. An increase of the noise is also observed for high latitudes in the Northern hemisphere, whereas for observations above the Pacific and Atlantic Ocean the noise is only slightly affected.

4.3.2.4. Data Filtering

Supplementary to the raw IASI L1C measurements, EUMETSAT distributes auxiliary L2 diagnostics, such as cloud cover and surface type. Utilizing these diagnostics, Schneider et al. (2021c) provide the MUSICA IASI retrieval results for (almost) cloud-free conditions over land,

Table 4.2. Diagnostic flag variables and their recommended values for selecting MUSICA IASI $\{H_2O, \delta D\}$ pair data with high quality. The flags indicating the vertical sensitivity and uncertainty of the $\{H_2O, \delta D\}$ pair product are individually set for each altitude level, while the flags for the cloud cover and the retrieval fit quality are not height-dependent (Schneider et al., 2021c). The vertical sensitivity flag depends on the measurement response, information displacement and vertical resolution of the $\{H_2O, \delta D\}$ pair kernels and the total error flag depends on the sum of the temperature and noise error of δD .

Filter type	Flag variable	Values
cloud cover	eumetsat_cloud_summary_flag	1, 2
retrieval fit quality	musica_fit_quality_flag	2, 3
vertical sensitivity	musica_wvp_kernel_flag	1
total error	musica_deltad_error_flag	1

oceans and sea ice, with small cloud contaminations being possible. They supply an additional diagnostic flag variable containing information about the quality of the MUSICA IASI spectral fit, i.e. the root mean square errors of the residuals of the spectral fit (see Fig. 4.5a). Observations where the simulated spectrum did not converge against the measured spectrum are sorted out from the outset.

As part of the water isotopologue post-processing, an additional data selection is recommended according to the quality of the retrieved $\{H_2O, \delta D\}$ pair results and further height-dependent flags are shared for an intuitive and user-friendly data handling.

First, the flag variable `musica_wvp_kernel_flag` serves for filtering the $\{H_2O, \delta D\}$ pairs according to their vertical sensitivity, i.e. to obtain retrieval results that are actually sensitive to the true atmospheric state rather than to the a priori state (see Sect. 4.2.3). Therefore, this flag is defined based on the sensitivity metrics of the kernel matrix $\mathbf{A}_{wv,m}^*$. The value range of acceptable measurement response is set between 0.8 and 1.2. To limit the information displacement at an altitude level $z(i)$, the following criterion is considered:

$$\frac{|c(i) - z(i)|}{z_{cl}(i)} \leq 0.5 \quad (4.27)$$

with $c(i)$ being the centroid of the corresponding averaging kernel (Keppens et al., 2015) and $z_{cl}(i)$ the a priori correlation length at the respective altitude level (Schneider et al., 2021c). This criterion ensures that the deviation of the centroid from the nominal height is less than half of the a priori correlation length. As filter condition for the vertical resolution the relation

$$\frac{r_{LW}(i)}{z_{cl}(i)} \leq 3.5 \quad (4.28)$$

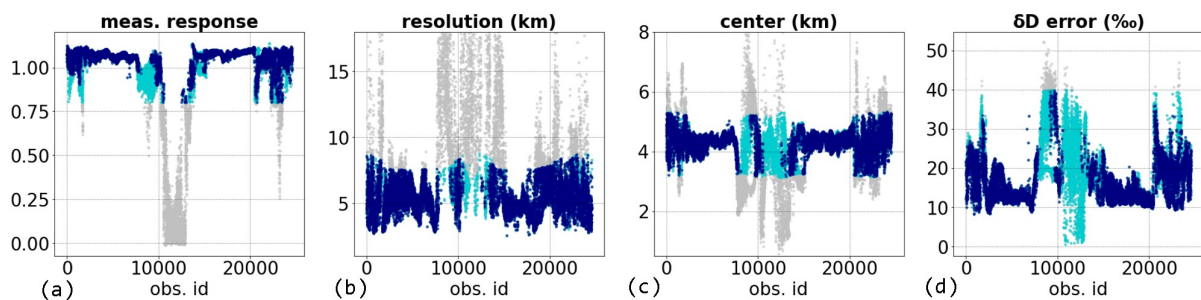


Figure 4.9. Effects of individual filter criteria on the corresponding metrics, along the Metop-A orbit 55524, as also shown in Fig. 4.8. The grey scatter show all observations along the specific orbit, the cyan scatter show the available data for each variable according to its individual filter criterion and the dark blue scatter are the available data when simultaneously filtering for all four metrics. The metrics (a) – (c) are used to define the flag `musica_wvp_kernel_flag`, and the noise in (d) is used for the flag `musica_deltad_error_flag` (compare with Table 4.2).

is applied. $r_{LW}(i)$ is the layer width per one DOFS from Purser and Huang (1993) (see Sect. 4.2.3) as a proxy for the vertical resolution of an averaging kernel. By considering the kernel properties relative to the correlation length, also kernels with larger values in their metrics can pass the aforementioned filters if larger values in the corresponding correlation length are assumed. Second, the error flag `musica_deltad_error_flag` is provided, which identifies data points with too high uncertainties in the δD retrieval results, namely errors due to measurement noise and atmospheric temperature uncertainties. The corresponding height-dependent flag displays retrieval results with a total δD error below 40 %.

The aforementioned filter conditions are visualized in Fig. 4.9 and the respective flags and their recommended values are summarized in Table 4.2. The flag variables `musica_wvp_kernel_flag` and `musica_deltad_error_flag` are binary, i.e. they only consist of the values 1 (for indicating high quality) and 0 (for low quality). Even though the recommended filter conditions are chosen somewhat arbitrary, they efficiently remove recognisable outliers in terms of kernel properties (see Fig. 4.9a–c) and data uncertainties (see Fig. 4.9d) of the retrieved $\{H_2O, \delta D\}$ pairs. Therefore, the simultaneous application of the corresponding quality flags `musica_wvp_kernel_flag` and `musica_deltad_error_flag` serves for a convenient and meaningful selection of reliable $\{H_2O, \delta D\}$ pair data. However, to enable a flexible adjustment of the individual filter conditions for individual purposes, the output datasets contain the filtered and unfiltered $\{H_2O, \delta D\}$ pair data together with the flag and filter variables.

4.3.2.5. Matrix Compression

Analogous to Schneider et al. (2021c), the averaging kernel matrices for the $\{H_2O, \delta D\}$ pairs are stored in a decomposed and compressed format in order to reduce the required storage volume.

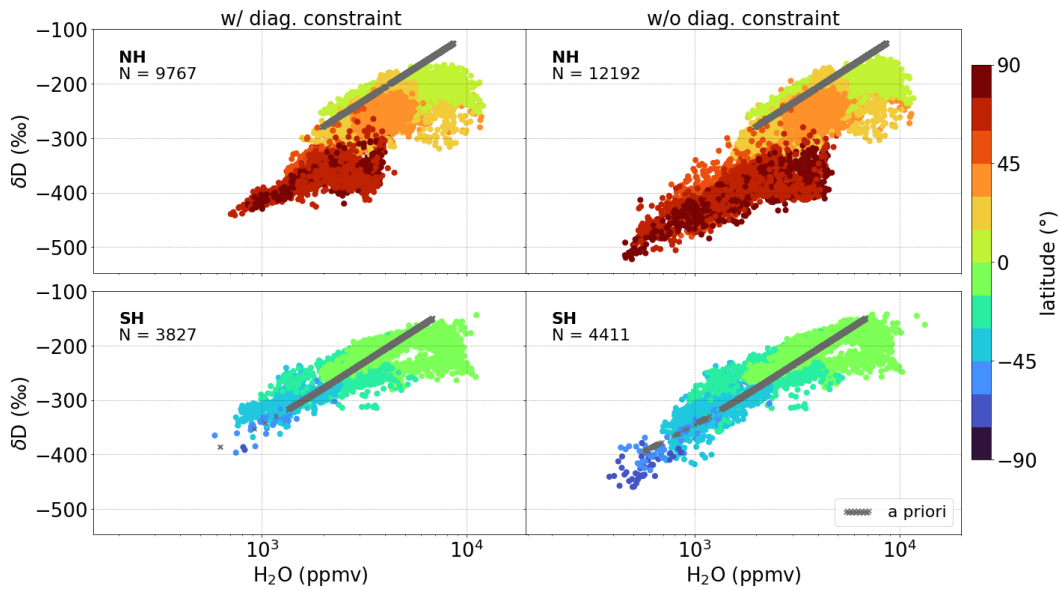


Figure 4.10. Quality filtered $\{H_2O, \delta D\}$ pairs (according to Table 4.2) for the original (with diagonal constraint) and the improved product (without diagonal constraint) at 4.2 km a.s.l. along the Metop-A orbit 55524 during boreal summer (orbit also shown in Fig. 4.8 and 4.9). The upper (lower) row shows the scatter for data of the northern (southern) hemisphere, color-coded with the corresponding latitude values. The grey scatter show the a priori values of the individual observations at the nominal altitude. The value of N indicates the respective number of plotted data points.

For this purpose, a singular value decomposition is applied for the matrices $A_{wv,m}^*$ and $A'_{t,wv,m}$ into the components U , D and V that decompose the kernel matrix through

$$A = UDV^T \quad (4.29)$$

The length of the singular value vector D is called rank. The actual compression is achieved by cutting off the lowest singular values in D and thereby reducing the rank. Consequently, also the number of singular vectors U and V are reduced. The optimal limit of the singular values for balancing the compression error against the effective storage reduction is discussed in Weber (2019). Based on that, singular values that are less than 0.1 % of the maximum singular value in D will be neglected.

4.3.2.6. The Final $\{H_2O, \delta D\}$ Pair Product

Performing the aforementioned post-processing and filtering results in the final $\{H_2O, \delta D\}$ pair product, as shown in Fig. 4.10 for a full Metop orbit. The consideration of the sensitivity optimization leads to a substantial increase of variability in the $\{H_2O, \delta D\}$ pairs at 4.2 km a.s.l. for dry regions. For instance, over polar areas the weaker constraints allow larger deviations from the corresponding a priori values, such that lower values in H_2O and δD can be observed.

This is analogous to the increase in the measurement response that is most pronounced for dry conditions (see Fig. 4.6, 4.7 and 4.8). As the measurement response is considered during the quality filtering for reliable $\{H_2O, \delta D\}$ pairs (see Table 4.2), its increase yields a higher number of observations passing the recommended data filter (see data amount in Fig. 4.10).

In summary, the MUSICA IASI water isotopologue post-processing provides an optimal estimation $\{H_2O, \delta D\}$ pair product in the troposphere with a substantial increase of sensitivity for dry conditions. Together with the recommended quality flags indicating observations with meaningful averaging kernels and low errors for δD , this is the main product provided freely to the scientific community.

4.3.3. Spatio-temporal Data Characterization

The following section gives an impression of the spatial and temporal coverage of the full optimal estimation $\{H_2O, \delta D\}$ pair dataset.

4.3.3.1. Output Volume of $\{H_2O, \delta D\}$ Pair Dataset

By using the retrieval output from Schneider et al. (2021c), the water isotopologue post-processing is performed as part of the thesis for all MUSICA IASI results between October 2014 and June 2019 (MUSICA IASI retrieval dataset referenced in Schneider et al., 2021a, Schneider et al.). With on average 350 000 cloud-free observations per sensor (Metop-A and Metop-B) and per day, this results in around 1200 million observations processed for mid-tropospheric $\{H_2O, \delta D\}$ pair information.

According to the local overpass times of the Metop satellites, the orbits are split into morning (~ 09.30 local time) and evening (~ 21.30 local time) overpasses and the respective observations are concatenated for all overpasses within a single day into an individual global NetCDF4 file. That is, two files per day emerge with each having a size of around 2 GB, resulting in about 1.5 TB per year. The full output volume is approximately 7 TB.

4.3.3.2. Degree of Freedom for Signal

Figure 4.11 shows the DOFS values of δD as monthly means for February and August 2018, for morning and evening observations, respectively. To consider the full quality range of the $\{H_2O, \delta D\}$ pair results, the δD distributions are filtered only for cloud-free scenes and acceptable retrieval fit quality, but not for the $\{H_2O, \delta D\}$ pair quality (only filtered for `musica_fit_quality_flag` ≥ 2 and `eumetsat_cloud_summary_flag` ≤ 2).

Maximum values are around 2 and are found over the tropics (persistently throughout the entire

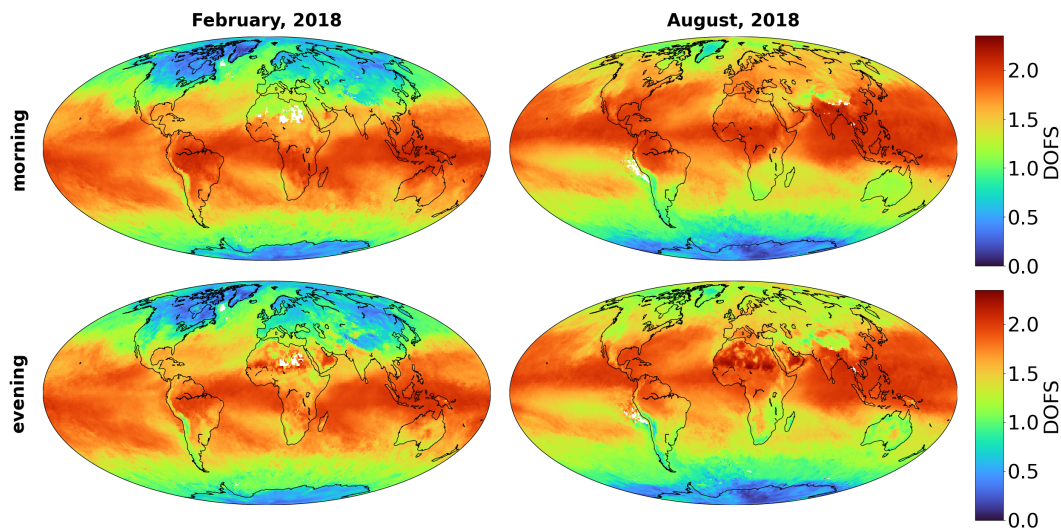


Figure 4.11. Monthly averages for February and August 2018 for the DOFS of the $\{H_2O, \delta D\}$ pair product without $\{H_2O, \delta D\}$ pair quality filtering, evaluated on a $1^\circ \times 1^\circ$ grid.

year) and the sub-tropics (during summer), indicating the capability of independently resolving signals in the lower and middle free troposphere (as indicated by the averaging kernels in Fig. 4.6). The DOFS minimum is located over the polar regions during winter times, as these regions are typically very dry (less water vapor signals) and cold (less thermal emissions).

Over oceans, the DOFS distribution roughly reflects the dominant sea-surface temperature patterns. For instance, the warm surface currents in the West Atlantic and West Pacific correlate with an increased sensitivity of the $\{H_2O, \delta D\}$ pair retrievals.

While the large-scale DOFS patterns show a strong intra-annual variability for all regions except the tropics, their diurnal variations are rather small. Instead, the latter becomes more pronounced for small-scale regional structures. In particular for land observations, thermal effects lead to a sensitivity maximum for morning times (Clerbaux et al., 2009), e.g. for Australia during February and for Europe and North America during August. Conversely, for the Sahara an inverted effect can be observed, i.e. an increase of DOFS from morning to evening. A reason for this could be the strongly enhanced boundary layer mixing that develops here throughout the day and mixes near-surface water vapour up to 5–6 km (Hall and Peyrillé, 2006). As a next step, data will be considered that have been additionally filtered for high sensitivity and low uncertainty in the $\{H_2O, \delta D\}$ pair product (see Table 4.2).

4.3.3.3. Vertical Distribution of Data Coverage

As discussed in Sect. 4.3.2, the MUSICA IASI water vapor retrieval is mainly sensitive to water vapor in the free troposphere. Figure 4.12 shows that this is reflected clearly on the vertical distribution of available $\{H_2O, \delta D\}$ pairs after applying the full recommended filters according

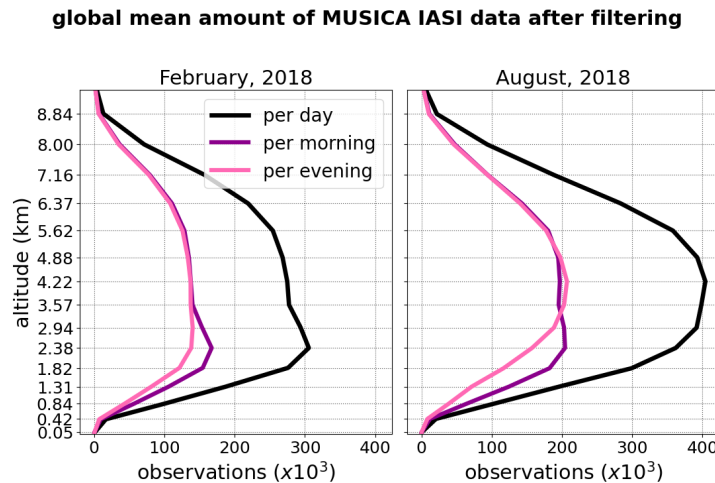


Figure 4.12. Averaged amount of quality filtered $\{H_2O, \delta D\}$ pairs for the tropospheric retrieval grid altitudes during February and August 2018. The black line indicates the global means that are further divided into the means for morning (violet) and evening (pink) overpasses.

to Table 4.2. Here, the amount of globally available observations per day and per morning and evening maps is averaged for February and August 2018 and is shown for each retrieval grid level between the surface and 9 km. The best data availability arises between 2–7 km a.s.l. On average, during boreal summer (at maximum over 400 000 data pairs per day) remarkably more observations are available than during austral summer (maximum 300 000 data pairs). This is in agreement with Fig. 4.3 and 4.4 and might be an effect of the greater amount of land masses and warm seas in the Northern Hemisphere.

In contrast, the diurnal variations are again rather small on the global scale. Only for altitudes below 3.5 km a.s.l. a slight decrease of data availability is observed during evening. This might be due to thermal heating that develops during the day and leads to vertical transport of low-level moisture, resulting in an upwards shift of the retrieval sensitivity. Such effects are stronger over land masses and during summer and probably lead to a larger morning-to-evening difference during boreal summer, as there are more land masses in the Northern than in the Southern Hemisphere.

4.3.3.4. Horizontal Distribution of Data Coverage

This section discusses the horizontal data coverage of $\{H_2O, \delta D\}$ pairs for different altitude regions after applying the respective quality filtering according to Table 4.2. To identify those retrieval results that provide vertical profile information, observations that simultaneously fulfil the respective filter criteria for two distinct altitudes will be considered.

Figure 4.13 shows the mean horizontal coverage of quality-filtered $\{H_2O, \delta D\}$ pairs for different

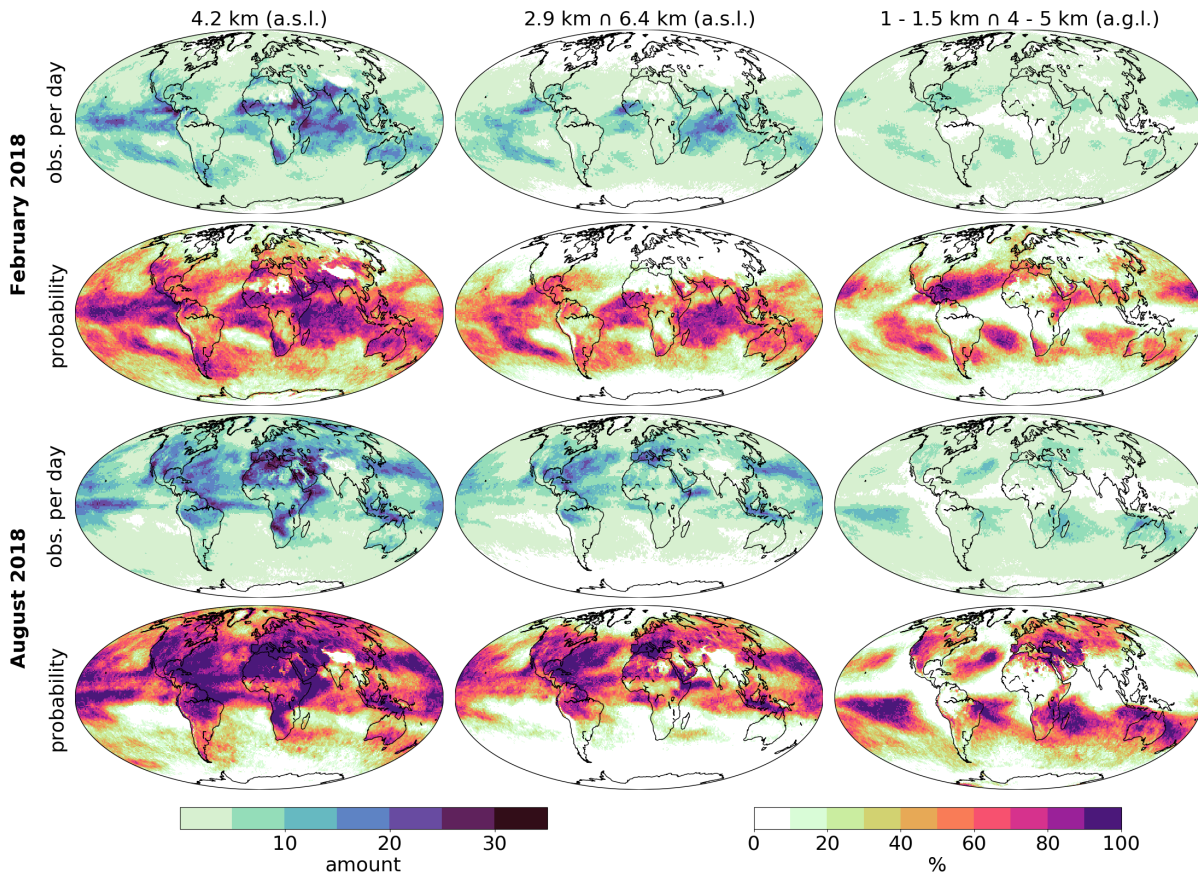


Figure 4.13. Monthly statistics for the horizontal availability of MUSICA IASI $\{H_2O, \delta D\}$ pair data for February 2018 (first and second row) and for August 2018 (third and fourth row). Data are filtered according to Table 4.2. The first row for each month, respectively, shows the averaged amount of available observations per $1^\circ \times 1^\circ$ grid box and per day, the respective second row gives the frequency of days with at least 1 reliable observation inside a single $1^\circ \times 1^\circ$ grid box. Shown are the results for observations at 4.2 km above sea level (a.s.l.) in the first column, for observations fulfilling the quality filter conditions simultaneously at 2.9 and 6.4 km a.s.l. (second column) and simultaneously at 1–1.5 and 4–5 km above ground level (a.g.l.; third column). For the latter, if more than one grid level falls inside the given altitude range, then the lower one is chosen.

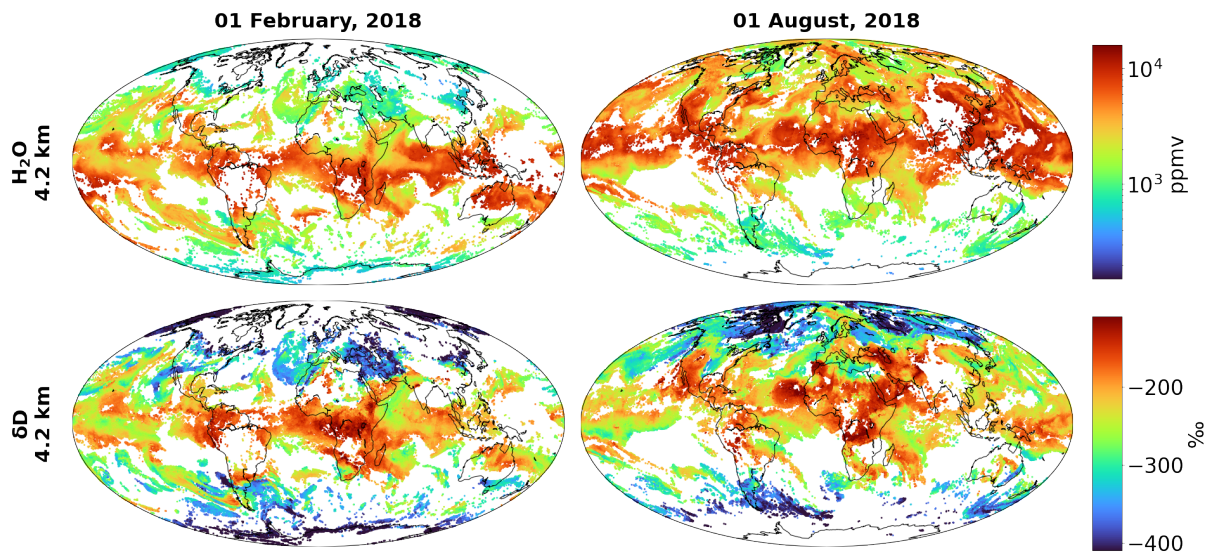


Figure 4.14. Horizontal distributions of H_2O and δD from the optimal estimation pair product, after filtering according to Table 4.2. Data are shown for 4.2 km a.s.l. and for 01 February and 01 August 2018 (including both morning and evening observations). The range of the colorbars is adjusted to Fig. 15 of Schneider et al. (2021c).

altitude regions during February and August 2018; the corresponding IASI observations are evaluated on a $1^\circ \times 1^\circ$ grid. The averaged number of daily available $\{H_2O, \delta D\}$ pairs and the fraction of days with at least one measurement are illustrated for each grid box. Additionally, Table 4.3 provides the total fractions of available $\{H_2O, \delta D\}$ pairs on each altitude region compared to all cloud-free IASI observations.

At 4.2 km a.s.l., up to 59 % of all cloud-free IASI observations provide reliable $\{H_2O, \delta D\}$ pair data, with best horizontal coverage over tropical and subtropical summer regions. Here, up to 35 observations are available per day and grid box and over wide areas there is a 100 % frequency of $1^\circ \times 1^\circ$ grid boxes with at least one reliable observation, especially in the tropics and the summertime sub-tropics. But also for high northern latitudes, where typically cold and dry conditions prevail, a satisfactory data availability is apparent. Furthermore, for about 22–30 % of the cloud-free observations the quality filter conditions are simultaneously fulfilled at 2.9 and 6.4 km a.s.l. Similar spatial patterns with lower values and less temporal coverage are observed, when compared to 4.2 km a.s.l. Even though the data coverage decreases significantly for areas with profile information at even lower altitudes (the quality filter conditions are simultaneously fulfilled at 1–1.5 and 4–5 km above ground level only for about 10–17 % of the cloud-free observations), interesting features emerge. The maximum availability of about 10 observations per grid box and per day shifts towards higher latitudes, such that over the tropics almost no data are found.

In this analysis, the morning and evening observations are jointly investigated. As can be

Table 4.3. Averaged fractions of available MUSICA IASI $\{H_2O, \delta D\}$ pair data after applying the quality filter according to Table 4.2, compared to the full (i.e. unfiltered) cloud-free IASI observations. The results are shown for 4.2 km above sea level (a.s.l.), for observations where the filter conditions are fulfilled simultaneously at 2.9 and 6.4 km a.s.l. and at 1–1.5 and 4–5 km above ground level (a.g.l.), respectively. For the latter, if more than one grid level falls inside the given altitude range, then the lower one is chosen.

Date	Overpass	4.2 km (a.s.l.)	2.9 \cap 6.4 km (a.s.l.)	1–1.5 \cap 4–5 km (a.g.l.)
Feb. 2018	morning	41.4 %	21.7 %	13.0 %
	evening	40.7 %	20.1 %	9.8 %
Aug. 2018	morning	57.5 %	30.1 %	23.0 %
	evening	58.8 %	27.1 %	11.2 %

deduced from Fig. 4.11 and 4.12, the differences between the morning and evening distributions will differ only little. For instance, Table 4.3 includes the fractions of available data after filtering according to Table 4.2 for the altitude regions from Fig. 4.13. The values do not differ significantly for the mid-troposphere during morning and evening times, but reduce for lower altitudes during the evening overpasses (analogous to Fig. 4.12).

To convey an impression of the actual horizontal data distribution of the $\{H_2O, \delta D\}$ pair product, Figure 4.14 depicts all data of H_2O and δD at 4.2 km a.s.l. for two days (01 February and 01 August 2018). The horizontal patterns of available data are in agreement with Fig. 4.13. Both H_2O and δD show highest values over tropical regions and decrease towards the polar areas. However, differences in their zonal distribution become apparent. For instance, while H_2O and δD show consistently high values over Northern Africa, large discrepancies appear at similar latitudes over the Pacific (high H_2O combined with decreased δD). Section 4.3.3.6 will give further insights into such relations between H_2O and δD .

4.3.3.5. Horizontal Distribution of Total Error

Figure 4.15 shows the horizontal distributions of the total errors of H_2O and δD at 4.2 km a.s.l., averaged on a $1^\circ \times 1^\circ$ grid for February and August 2018. Overall, an anti-correlation to the DOFS distributions in Fig. 4.11 may be identified. The lowest errors are found for warm and moist tropical and sub-tropical sites during summer, where the DOFS is maximum. Here, the minimum error values lie around 5 % and 10 ‰ for H_2O and δD , respectively. With decreased sensitivity during winter and for higher latitudes, an increase of the total errors becomes apparent, in particular over land areas. The errors can reach values up to $\sim 12\%$ and 30‰ for H_2O and δD ,

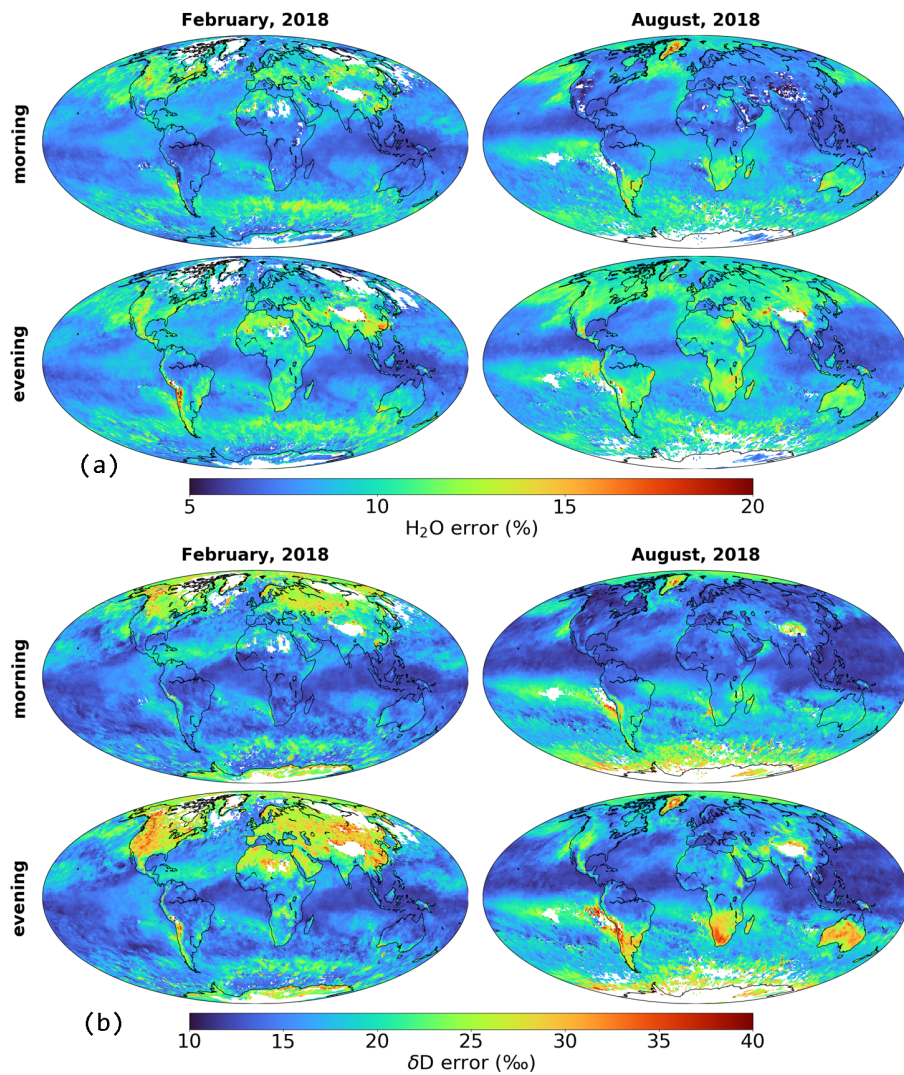


Figure 4.15. Monthly averages for morning and evening overpasses during February and August 2018 for the total errors of the filtered H_2O (a) and δD (b) product at 4.2 km a.s.l., evaluated on a $1^\circ \times 1^\circ$ grid. The filtering is performed according to Table 4.2.

but are still in the uncertainty range of other comparable remotely sensed products (Worden et al., 2006, 2019).

In particular over dry sub-tropical land surfaces (e.g. Sahara in February and South Africa and Australia during August), there is a remarkable increase of the total errors in δD from morning to evening times. This feature does not show any clear and consistent correlation with the DOFS development in Fig. 4.11 and thus requires further investigation in the future.

4.3.3.6. Data Example: Tropical Atlantic against Sahel

To convey an impression of the amount and scientific potential of the optimal estimation MUSICA IASI $\{H_2O, \delta D\}$ pair product, results are shown for two illustrative regions of interest, namely

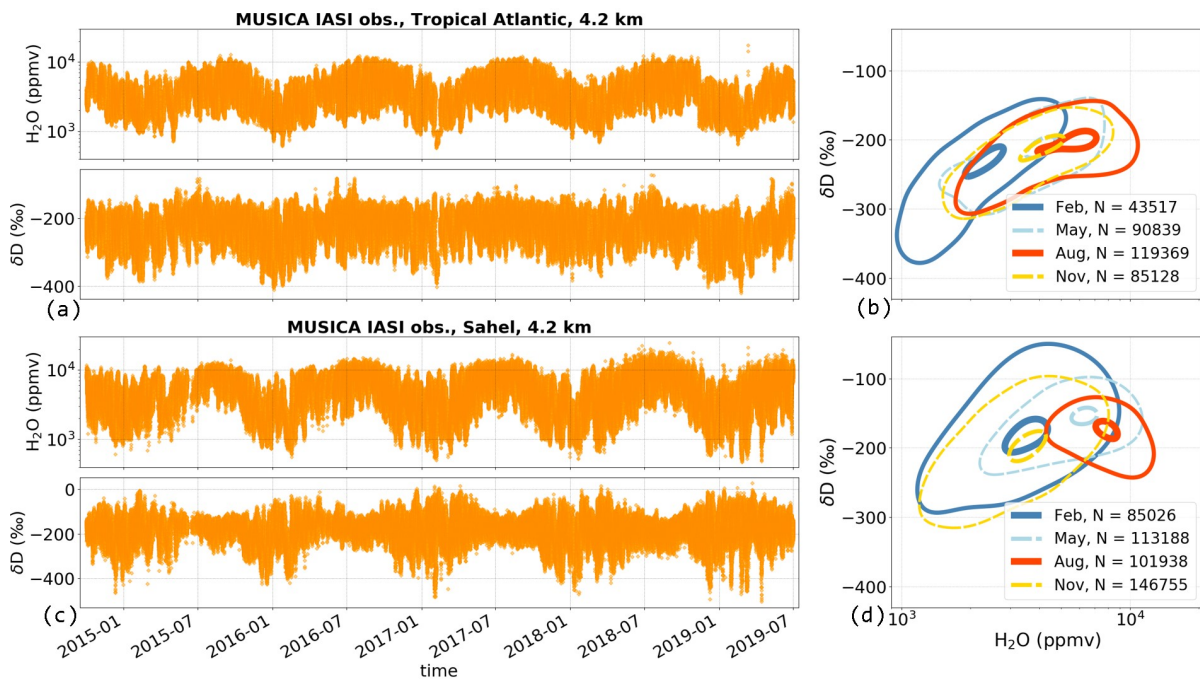


Figure 4.16. MUSICA IASI $\{H_2O, \delta D\}$ pair data for 4.2 km a.s.l. over the Tropical Atlantic ($13\text{--}17^\circ$ N, $46\text{--}30^\circ$ W) and the Sahel in West Africa ($13\text{--}17^\circ$ N, 8° W – 8° E), for the full MUSICA IASI period. The left plots show the time series of H_2O and δD . The right plots show the respective probability density functions of the two-dimensional $\{H_2O, \delta D\}$ distributions, indicating the location of the main 10 and 90 % scatter points. These contours are drawn for the data of February, May, August und November, summarized for all respective years.

the Tropical Atlantic ($13\text{--}17^\circ$ N, $46\text{--}30^\circ$ W) and the Sahel in West Africa ($13\text{--}17^\circ$ N, 8° W – 8° E). Figure 4.16 shows the time-series of the respective MUSICA IASI data for H_2O and δD at 4.2 km a.s.l. that have passed the full recommended filtering (according to Table 4.2) for the period October 2014 to June 2019. As discussed in Section 4.3.2, the harmonized retrieval results for H_2O and δD offer almost the same averaging kernels, thereby allowing for a meaningful interpretation of paired $\{H_2O, \delta D\}$ distributions. Based on that, Figure 4.16 also summarizes the mean monthly evolution (represented by February, May, August and November) of the $\{H_2O, \delta D\}$ pair distribution over the Tropical Atlantic and the Sahel. The data are illustrated with normalized two-dimensional histogram contours comprising the main 10 and 90 % of the scatter points (the calculation is described in Sect. 3.1.5).

Over the Tropical Atlantic, both H_2O and δD exhibit a similar annual cycle, even though it is weaker for δD . This can also be observed in the corresponding $\{H_2O, \delta D\}$ pairs, where the August data are on average moister and more enriched in δD than the February data. Despite some shifting and tilting, the overall contour shape does not change much from February to August. The contour for May reflects the transition from February to August and the one for November

vice versa.

In contrast, over the Sahel signs of an annual anti-correlation between H_2O and δD appear. Again, during boreal there is a minimum of H_2O and δD , even though it is slightly moister than over the Tropical Atlantic (for both the November and February contour). In the further course of the year, the variability of H_2O and δD decreases significantly, while the respective contours shift to higher H_2O (see May contour). However, this moistening is associated with a strong decrease of the maximum values in δD , leading to a remarkable tilting of the August contour, when compared to the February contour.

These regional differences highlight the benefit of adding information about δD when studying atmospheric moisture, because different moisture processes leave different impact on the shape and position of $\{H_2O, \delta D\}$ pair distributions. In the example of Fig. 4.16, it becomes clear that the Tropical Atlantic and the Sahel reveal significantly different structures in δD , while their H_2O distributions show clear and similar annual cycles. Therefore, this feature underlines that the tropospheric moisture over the two tropical regions is governed by structurally different processes. As δD is mainly affected during phase changes of water vapor, the hypothesis is that its observed anti-correlation to H_2O may be an effect of tropical convection that is exceptionally strong and highly organized on the mesoscale over the Sahel during the summertime monsoon period. Related dynamical changes in the contributing wind regimes might pose further contributing factors for changes in the $\{H_2O, \delta D\}$ phase space.

However, in order to robustly attribute such $\{H_2O, \delta D\}$ pair signals to underlying moisture processes, supplementary measurements and model analyses are required. As previous studies stated (e.g. Worden et al., 2007, Noone, 2012, Dyroff et al., 2015, González et al., 2016, Schneider et al., 2017, Christner et al., 2018, Eckstein et al., 2018, Lacour et al., 2018), such an analysis is then capable of providing a deeper understanding of atmospheric moisture pathways and will therefore be performed in-detailed in the further context of this thesis (see Chap. 5 and 6).

4.4. Retrieval Simulations for $\{H_2O, \delta D\}$ Pairs

In order to facilitate the direct comparison of the newly developed MUSICA IASI $\{H_2O, \delta D\}$ pair data to output from isotope-enabled models, the model data require an adjustment according to the IASI characteristics (see Sect. 3.2.2). The optimal approach is to simulate how the satellite sensor would observe the atmosphere given by an atmospheric model (Schneider et al., 2017). This simulated satellite observation can then be directly compared with the actual satellite observations.

Section 4.4.1 presents and evaluates a new method developed within the thesis for simulating the MUSICA IASI characteristics by using the full PRFFWD radiative transfer algorithm (see Sect. 4.2.2) and the post-processing for the optimal estimation $\{H_2O, \delta D\}$ pair product. The

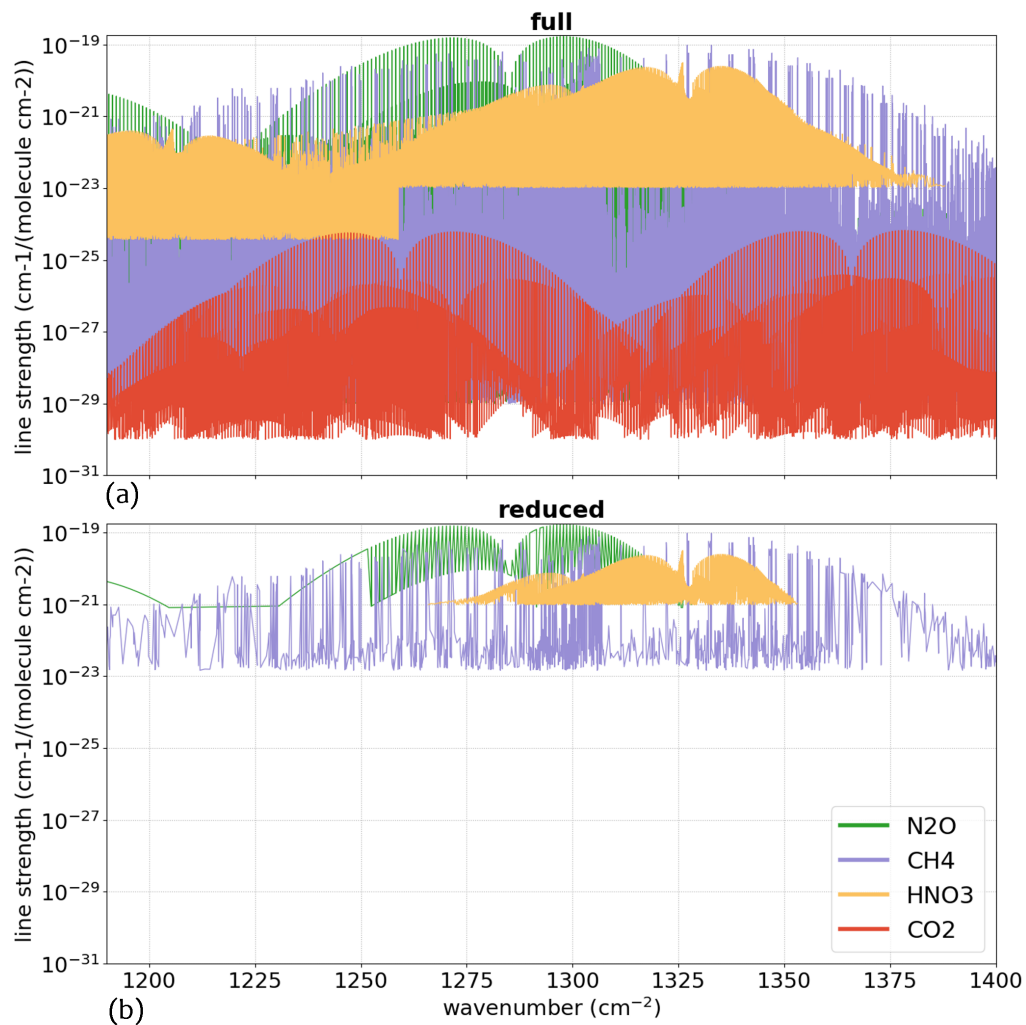


Figure 4.17. Spectroscopic line strengths for CH_4 , N_2O , HNO_3 and CO_2 within the MUSICA IASI spectral window, as used (a) for the original MUSICA IASI retrieval and (b) for the MUSICA IASI retrieval simulator.

retrieval simulator will be applied to data from $COSMO_{iso}$ and evaluated against the actual MUSICA IASI $\{H_2O, \delta D\}$ pair data (Sect. 4.4.2).

4.4.1. Methodology

The following two sections present the modifications regarding the radiative transfer assumptions during the application of PRFFWD in general and with additional focus on the data structure of $COSMO_{iso}$.

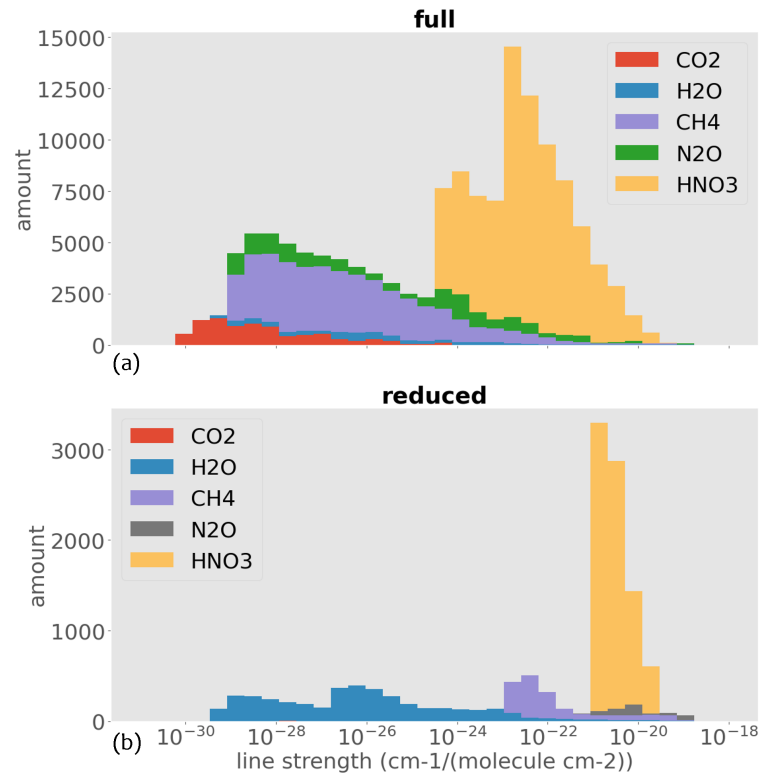


Figure 4.18. Amount of spectroscopic lines for CH_4 , N_2O , HNO_3 and CO_2 within the MUSICA IASI spectral window, as used (a) for the original MUSICA IASI retrieval and (b) for the MUSICA IASI retrieval simulator.

4.4.1.1. Simulation Approach

Following the discussions in Schneider et al. (2017), the recommended approach for harmonizing data from an isotope-enabled model with the remote sensing characteristics of the MUSICA IASI product is to simulate a theoretical IASI observation based on the modeled atmosphere. This is achieved by applying a forward operator that considers the IASI characteristics to the atmospheric state given by the model, analogous to the forward relation in Eqn. (4.1). Then, by applying Eqn. (4.6), this allows to calculate the averaging kernel matrix A_{mod} , which corresponds to the modeled atmospheric conditions (temperature, pressure, H_2O and δD profiles). Treating the modeled atmospheric state x_{mod} as the true atmospheric state, the relation in Eqn. (4.7) can then be used to derive the theoretical IASI observation \hat{x}_{mod} :

$$\hat{x}_{\text{mod}} = A_{\text{mod}}(x_{\text{mod}} - x_a) + x_a \quad (4.30)$$

As discussed in Sect. 3.2.2, Schneider et al. (2017) developed a method for simulating the IASI averaging kernels A_{mod} based on a simplified radiative transfer model. They already achieved a reliable prediction of the IASI kernels for a modeled atmosphere, despite strong limitations in the radiative transfer assumptions. However, aiming towards a most reasonable cross-comparison

between the MUSICA IASI {H₂O, δD} pair data and results from any isotope-enabled model, the thesis develops an improved approach of simulating A_{mod} . Specifically, this will be based on the full PRFFWD radiative transfer algorithm together with the post-processing for the optimal estimation {H₂O, δD} pair product. This ensures that the retrieval simulator is reliable as possible. However, such line-by-line calculations during the forward model are rather resource-intensive, which can be a handicap for comprehensive model-to-measurement cross-comparisons (Schneider et al., 2017). Therefore it is useful to think of reasonable simplifications for speeding up the radiative transfer calculations. With the intent not to restrict the actual PRFFWD algorithm, the thesis will give evidence that considering only the dominant spectroscopic lines of HITRAN2016 during the line-by-line calculations is a very efficient way for reducing the required computational resources without a great loss of the data quality.

Figure 4.5 and 4.17a show the line strengths and Fig. 4.18a the amount of spectroscopic lines for each MUSICA IASI target species that lie within the spectral window of 1190–1400 cm⁻¹ and that are considered during the MUSICA IASI retrieval. In fact, for every species also the spectroscopic features of their individual isotopologues are included. For instance, H₂O comprises in this context the values of H₂¹⁶O, H₂¹⁸O, H₂¹⁷O, HD¹⁶O, HD¹⁸O and HD¹⁷O (Gordon et al., 2017). As the main focus lies on the different isotopologues of H₂O, a first simplification is achieved by limiting the other species to the respective main isotopologue, i.e. ¹²CH₄ for CH₄, ¹⁴N₂¹⁶O for N₂O, H¹⁴N¹⁶O₃ for HNO₃ and ¹²C¹⁶O₂ for CO₂. This is justified by the fact that each of these isotopologues has a mean abundance of around 99 % relative to the respective weaker isotopologues.

Furthermore, the effect of each spectroscopic line on the radiative transfer calculations depends on the respective line strength. The stronger a specific line for a specific species, the more effective is the molecular absorption of the species at this wave number. As Fig. 4.18a shows, the lines in the spectral window of the MUSICA IASI retrieval comprise line strengths with more than ten orders of magnitude (from 10⁻³⁰ to 10⁻¹⁹ cm⁻¹/(molecule cm²)). For nearly every species the amount of considered lines increases significantly with decreasing line strength. However, the radiative transfer calculations are mainly determined by the respective strongest spectroscopic lines. Therefore, a second simplification is achieved by only considering those spectroscopic lines with line strengths above the 90th percentile for each target species (except for H₂O), respectively. The resulting distribution of available spectroscopic lines is shown in Fig. 4.17b and 4.18b. This leads to a drastic reduction of the line amounts for all species except H₂O. For instance, for CO₂, which has only weak lines in the chosen spectral windows, all lines disappear. For CH₄, N₂O and HNO₃ a large amount of strong lines remains, which still reflect the strongest and most dominant features.

Sensitivity tests revealed that by performing these line reductions the computation time of the

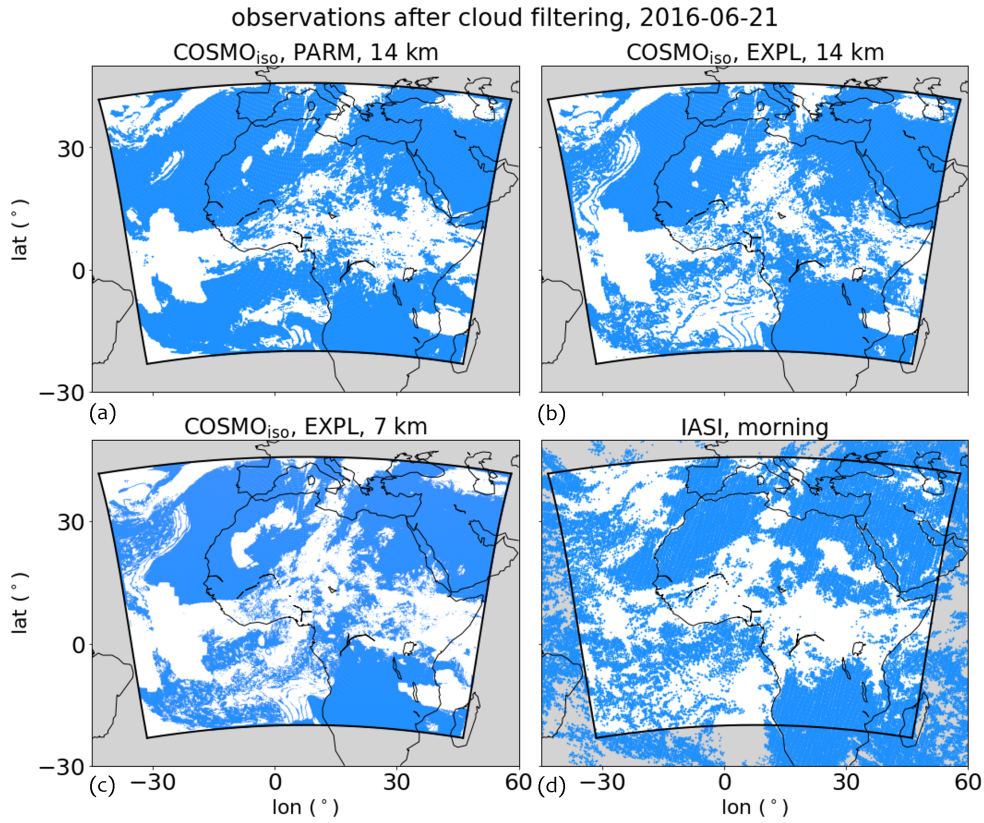


Figure 4.19. Available COSMO_{iso} data points for 21 June 2016, after applying the cloud filtering from Eqn. (4.32) and (4.33) to the output of the simulations (a) CISO1, (b) CISO2 and (c) CISO3 (see Table 3.2). Additionally, the horizontal distribution of cloud-free MUSICA IASI data are shown in (d).

PRFFWD forward model can be reduced by around 50 %. That these simplifications are not at the expense of the retrieval quality will be shown in the following sections.

4.4.1.2. Modifications for use with COSMO_{iso}

Within this study, the new retrieval simulator was initially developed and evaluated using data from the isotope-enabled model COSMO_{iso}. This section presents the adjustments that are necessary in order to run the PRFFWD algorithm exemplary for COSMO_{iso} simulations, which can be easily adapted to other models.

As first step, the COSMO_{iso} data are sampled according to the local overpass times of the Metop satellites (see Table 4.1). As the COSMO_{iso} uses the UTC time format t_{utc} as time coordinate, following expression is used to transform t_{utc} for each model data point by considering the respective longitudinal position x_{lon} to the local solar time t_{local} :

$$t_{\text{local}} = t_{\text{utc}} + 24 x_{\text{lon}}/360 \quad (4.31)$$

To match the morning (09.30 local time) and evening (21.30 local time) overpasses of Metop, only model data with t_{local} between 09–10 and 21–22 local time are considered, respectively. Analogous to the MUSICA IASI dataset, this yields two COSMO_{iso} maps per day, one for the morning and one for the evening data.

Further, the MUSICA IASI $\{H_2O, \delta D\}$ pair data are only available for cloud-free scenes (see Sect. 4.2.1). Therefore a corresponding cloud filtering needs to be performed also for the COSMO_{iso} data. This is a non-trivial task, as there is no binary cloud flag in the COSMO_{iso} output. Instead, the one-moment microphysics scheme provides information about the vertical profiles of frozen (q_i) and liquid (q_c) cloud particles. Following the cloud filter conditions for the IASI observations (no liquid clouds, small ice cloud contaminations possible, see Sect. 4.2.1), the following set of filter criteria is considered for selecting cloud-free COSMO_{iso} data points:

$$\sum_{k=z_{\text{srf}}}^{z_{\text{srf}}} q_{c,k} = 0 \text{ kg kg}^{-1} \quad (4.32)$$

$$\sum_{k=z_{\text{srf}}}^{z_{\text{srf}}} q_{i,k} \leq 10^{-10} \text{ kg kg}^{-1} \quad (4.33)$$

Even though the filter values are chosen somewhat arbitrarily, they lead to a reasonable cloud masking when comparing to IASI. Figure 4.19 shows the resulting distributions of available COSMO_{iso} data for three different model setups (model runs CISO1, CISO2 and CISO3, see Table 3.2). Overall, there is a good agreement between the cloud masks of the individual models and IASI. Differences become clear depending on the choice of convection treatment and horizontal resolution. While smaller grid cell sizes allow the representation of more detailed and sharp cloud occurrences, also the transition from parameterized to explicit convection leads to much finer structures. Compared with IASI, the explicit convection setup provides the most suitable cloud cover (in-detailed discussions of the different model setups are found in Sect. 6.2). Further, the atmospheric conditions provided by the model need to be interpolated to the vertical grid structure of the MUSICA IASI retrieval. While the top of the COSMO_{iso} configuration used here lies at 22 km, the PRFFWD settings used for the MUSICA IASI processing consider altitudes up to 55.6 km (see Fig. 3.6). Therefore, for height levels with missing model data the atmospheric variables temperature and pressure are determined according to monthly climatologies derived from the global and annual MUSICA IASI retrieval results. The H_2O and δD states above 22 km are set to the MUSICA IASI a priori states (see Sect. 4.2.2). As COSMO_{iso} does not provide information about the other trace gases (CH_4 , N_2O , HNO_3 and O_3), the respective MUSICA IASI a priori data are chosen for the full vertical profile.

Finally, the COSMO_{iso} output does not include any data about surface emissivities, therefore the land surface emissivities are taken from the global database of infrared land surface emissivities

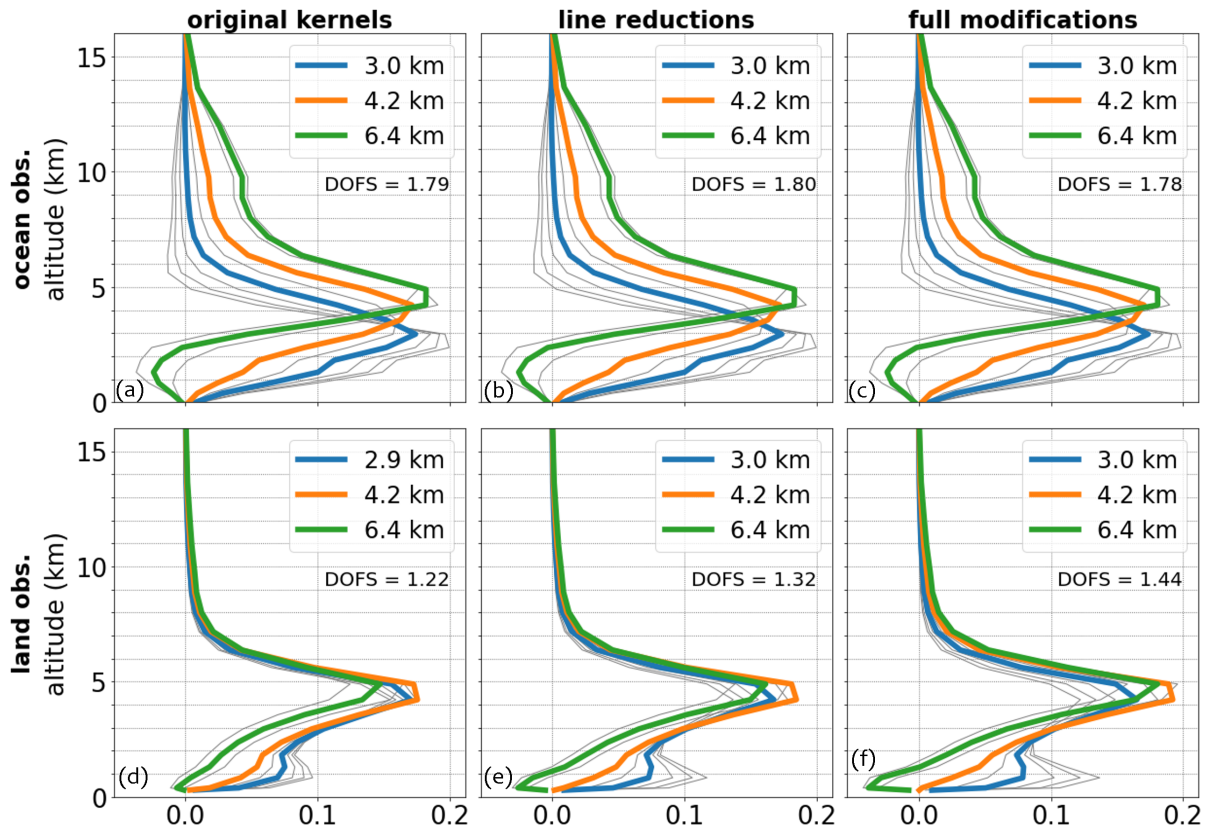


Figure 4.20. $\{H_2O, \delta D\}$ pair averaging kernels, for the original MUSICA IASI pair product (a,d), for a retrieval simulation with reduced spectroscopic lines (b,e) and for a full retrieval simulation including the full modifications required for the use with COSMO_{iso} data (c,f). The kernels are shown for a tropical ocean and land observation. The colored kernels correspond to the grid levels given in the legend, whereas the grey lines depict the kernels for the other grid levels.

(IREMIS, Seemann et al., 2008) and over ocean and ice observation the same external datasets as for the IASI retrieval are chosen (see Sect. 4.2.2).

4.4.2. Evaluation of Retrieval Simulations

This section discusses the quality of the retrieval simulations for $\{H_2O, \delta D\}$ pairs with respect to the original MUSICA IASI kernels and demonstrates exemplarily the effect of applying the retrieval simulator for COSMO_{iso} data when comparing the model with the actual IASI observations.

4.4.2.1. Theoretical Quality Assessment

Following the validation approach of Schneider et al. (2017), the quality of the new MUSICA IASI retrieval simulator will be assessed by comparing simulated averaging kernels of the $\{H_2O, \delta D\}$ pair product against the actual IASI kernels (see Fig. 4.6b). For this purpose, the retrieval

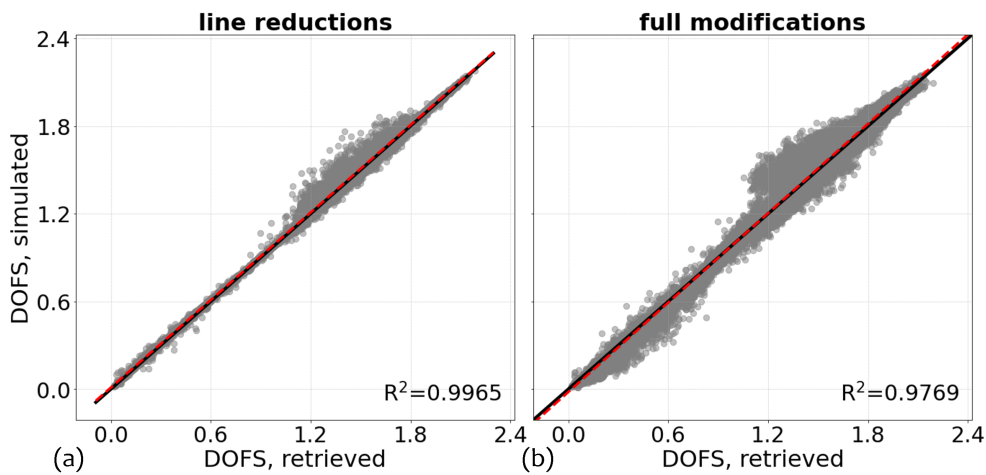


Figure 4.21. DOFS of the original (retrieved) MUSICA IASI $\{H_2O, \delta D\}$ pair kernels against the retrieval simulation with reduced spectroscopic lines (a) and with the full modifications (b). Data are shown for the whole Metop-A orbit with the ID 50189, during the 21 June 2016. Additionally, the squared Pearson's correlation coefficients (explained variance of distribution) are shown.

simulations are performed by using real IASI retrieval results from Sect. 4.2.2 as input, i.e. the retrieved states for temperature, H_2O , δD and the other interfering trace gases. In the case of a perfect simulation, the simulated IASI averaging kernels would match the real kernels, meaning that the retrieval simulator is capable of perfectly reproducing the original IASI kernels.

Figure 4.20 shows the comparison of the original IASI kernels (a,d) with the simulated kernels for a single ocean and land observation. For the kernel simulations in Fig. 4.20c,e only the line reductions from Sect. 4.4.1 are applied, while the full MUSICA IASI results of the retrieved states (temperature, pressure, H_2O , δD and interfering species) are used as input for the kernel simulation. In addition, Fig. 4.20d,f also includes those modifications necessary for the use with COSMO_{iso} data (use of IREMIS emissivities, interpolation of temperature and pressure above 22 km to IASI climatologies and use of a priori profiles for initializing the interfering species, see Sect. 4.4.1). For the ocean observation the retrieval simulator is capable of well reproducing the original IASI kernels for the mid-troposphere without any significant discrepancies. Also the land observation shows close agreement between the different kernels, even though small differences appear for the lower troposphere. The line reductions lead to a minor increase of the kernel values between 1–2 km and the respective DOFS value, what is even more pronounced for the fully modified simulation setup. A reason that partly explains these discrepancies is the different treatment of the emissivities over land surfaces.

Figure 4.21 provides a statistical perspective on the effect of the retrieval simulation setups from Fig. 4.20. It correlates the DOFS values of all $\{H_2O, \delta D\}$ pair averaging kernels along a full Metop orbit against the DOFS of respective kernel simulations. By only considering the line reductions,

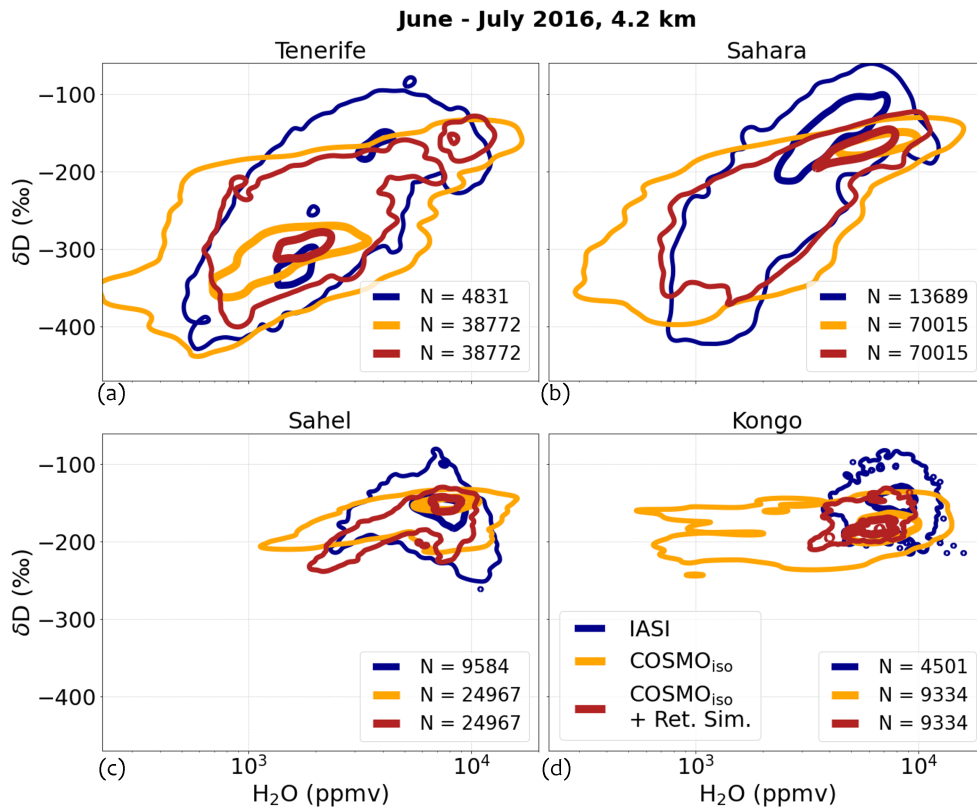


Figure 4.22. $\{H_2O, \delta D\}$ pair data for four different $4^\circ \times 4^\circ$ domains. The respective scatter (numbers given in the legends) are illustrated with the two-dimensional histogram contours for 50 and 99%. Contours are shown for the raw $COSMO_{iso}$ output (orange), the $COSMO_{iso}$ data processed with the MUSICA IASI retrieval simulator (red) and for the MUSICA IASI data (blue). The $COSMO_{iso}$ data are cloud-filtered according to Eqn. (4.32) and (4.33).

a squared correlation of more than 0.99 is achieved. Minor differences appear for DOFS values between 1.1 and 1.8, where the simulation tends to create higher values (in agreement with Fig. 4.20d,e). When considering the $COSMO_{iso}$ modifications, the correlation reduces to 0.9769, which is still a significant improvement compared to the simulated IASI retrievals from Schneider et al. (2017) (maximum correlation of 79 %).

4.4.2.2. Application to $COSMO_{iso}$ Data

This section shows illustrative case studies for demonstrating the effect of the MUSICA IASI retrieval simulator on $\{H_2O, \delta D\}$ data from $COSMO_{iso}$, when directly compared to the MUSICA IASI $\{H_2O, \delta D\}$ pair data. For this purpose, the model run CISO2 (14 km horizontal resolution, with explicit convection, June – July 2016) is used and processed with the full retrieval simulator, including the line reductions and the modifications required for handling the $COSMO_{iso}$ data structure (see Sect. 4.4.1).

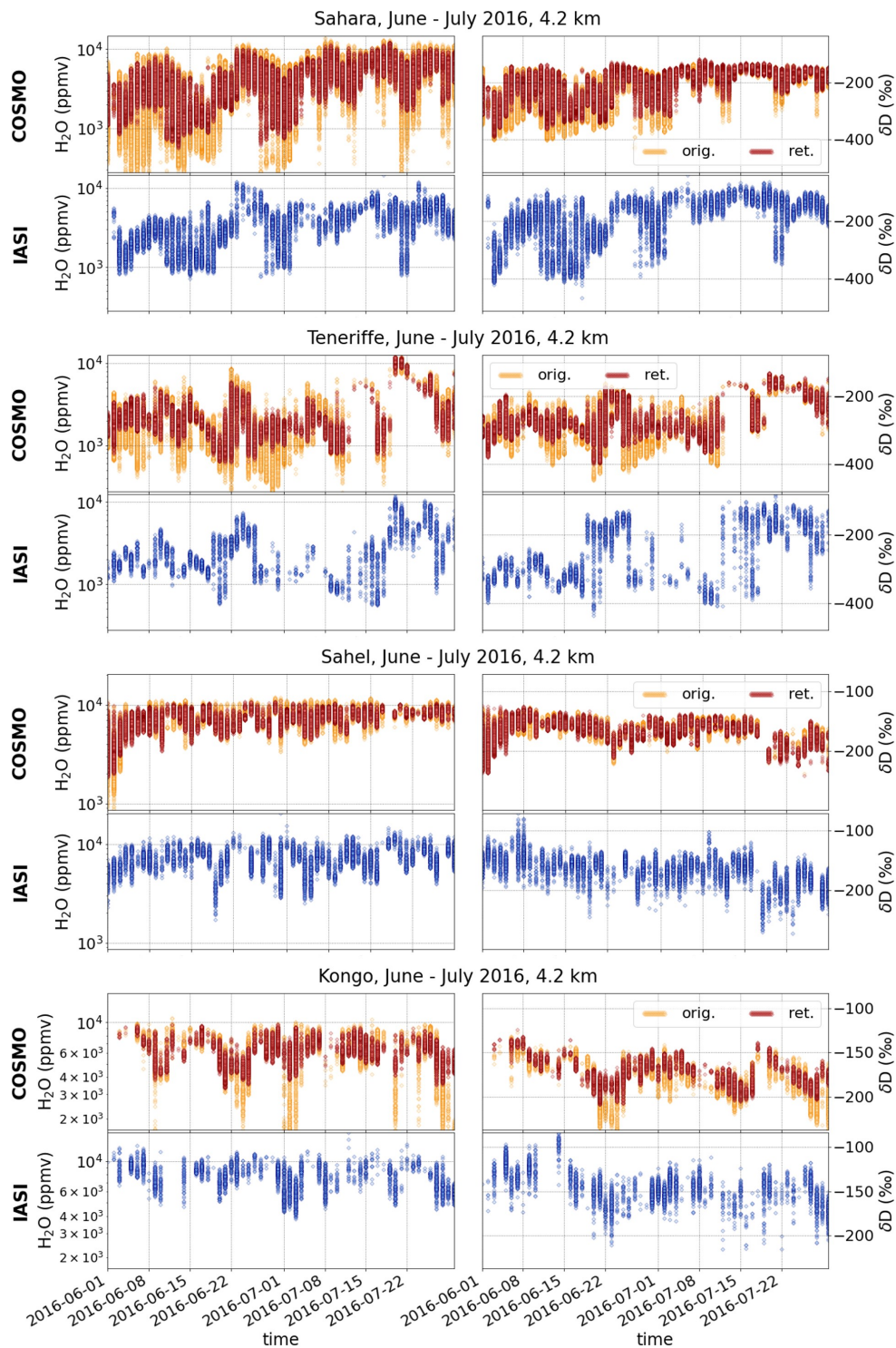


Figure 4.23. Time-series of the H_2O and δD data for the same domains as in Fig. 4.22. The orange scatter indicate the raw COSMO_{iso} output, the red scatter the COSMO_{iso} output processed with the retrieval simulator and the blue ones the IASI data.

Figure 4.22 illustrates the influence of the retrieval simulator in $\{H_2O, \delta D\}$ phase space. The two-dimensional histogram contours (50 and 99 %) are shown for the raw COSMO_{iso} output (sampled for the IASI overpasses), the COSMO_{iso} data processed with the retrieval simulator and the corresponding IASI data. Data are shown for four exemplary $4^\circ \times 4^\circ$ regions: Tenerife ($26^\circ - 30^\circ$ N, $14.5^\circ - 18.5^\circ$ W), Sahara ($25^\circ - 29^\circ$ N, $2^\circ - 6^\circ$ W), Sahel ($13^\circ - 17^\circ$ N, $0^\circ - 4^\circ$ W) and Kongo ($0^\circ - 4^\circ$ N, $12^\circ - 16^\circ$ E). The COSMO_{iso} data are cloud-filtered according to Eqn. (4.32) and (4.33). The square-shaped domain size is chosen such that the Tenerife box is centered around the actual island, as used in analyses by Dahinden et al. (2021) (note that the chosen Sahel domain is smaller than for Fig. 4.16).

Over Tenerife the application of the MUSICA IASI retrieval simulator leads to a strong improvement of the model data towards the IASI distribution. In particular, the unprocessed model data exhibit a larger variability for low H_2O and δD than IASI, which is reduced when including the simulated kernels. However, IASI reveals some moist and enriched data points (up to -100%) that are not covered in the processed COSMO_{iso} data, which is further discussed in Appendix A. Over the Sahara, the effect of the retrieval simulator appears to be similar to Tenerife. In contrast, the much moister tropical Sahel shows rather small effects, when applying the retrieval simulator. This suggests that during this period the Sahelian mid-troposphere exhibits homogeneous, high moisture profiles such that averaging around the adjacent altitude levels does not create significant changes. For the Kongo domain, the IASI $\{H_2O, \delta D\}$ pairs are constantly high, while COSMO_{iso} shows additional features with significantly lower H_2O . Again, the application of the retrieval simulator reduces this variability drastically, leading to a moisture range comparable to IASI.

The corresponding time-series of each domain further support the discussed findings (Fig. 4.23). For the subtropical domains Sahara (Fig. 4.23a) and Tenerife (Fig. 4.23b) there is a strong reduction of variability in the model, when the simulator is considered, making their time-series well comparable to the IASI time-series. As found in Fig. 4.22, the improvements are larger for drier model data. For moist observations over the Sahel and Kongo, the raw model data already represent well the IASI time-series and there is no considerable effect of the simulator. Only for occasional dry events within the Kongo domain, which are not observed by IASI, the retrieval simulator leads to significant improvements of the model data towards higher H_2O and δD values (compare with Fig. 4.22).

In summary, the MUSICA IASI retrieval simulator successfully merges the vertical $\{H_2O, \delta D\}$ pair characteristics of IASI into the modeled atmosphere of COSMO_{iso}, thereby allowing for a meaningful model-to-satellite cross-comparison.

4.5. Summary

This chapter presents an extension of the MUSICA IASI retrieval that aims at creating a high-quality water isotopologue pair product for the free troposphere. The retrieval processor from Schneider et al. (2021c) is an update of the version that was developed and validated against reference measurements during the MUSICA project (Schneider et al., 2016). The presented a posteriori processing step exploits their retrieval results and generates an optimal estimation $\{H_2O, \delta D\}$ pair product by harmonizing the averaging kernels of H_2O and δD , as proposed by Schneider et al. (2012). A further processing step by a posteriori reducing the strength of the underlying regularization is introduced. This increases the sensitivity of the $\{H_2O, \delta D\}$ pair retrieval product, especially for dry conditions, and enhances the vertical profile information between the boundary layer and the free troposphere. However, as trade-off the retrieval noise increases, but not beyond an unreasonable range ($\sim 12\%$ for H_2O and $\sim 30\%$ for δD). For a user-friendly data handling, supplementary filter flags are derived that perform a height-dependent data selection based on the quality of the $\{H_2O, \delta D\}$ pair results.

This a posteriori processing is herein applied to the MUSICA IASI full retrieval product, resulting in a novel space-borne dataset of tropospheric $\{H_2O, \delta D\}$ pair data. It consists of two global maps per day for all cloud-free IASI observations between October 2014 and June 2019. On a global average, the main vertical sensitivity lies between 2–7 km. It features best horizontal representativeness in terms of data quality and coverage for tropical and summertime sub-tropical regions. Despite a negative equator-to-pole gradient in the horizontal representativeness, there is still a satisfactory amount of reliable $\{H_2O, \delta D\}$ pair data in higher latitudes, ranging during summer up to the polar regions. The entire subject is referenced in a corresponding publication (see Diekmann et al., 2021b) and the full MUSICA IASI $\{H_2O, \delta D\}$ pair dataset is referenced via the DOI 10.35097/415 (Diekmann et al., 2021a).

In addition, this section includes the development of a new MUSICA IASI retrieval simulator, which considers the full line-by-line radiative transfer calculations as assumed during the MUSICA IASI retrieval. By drastically reducing the originally considered set of spectroscopic line parameters from HITRAN2016 for all target species except for H_2O and HDO, a computational speed-up of 50 % could be achieved. This is found to be not at the expense of the quality of the retrieval simulator. Further, the added value of applying the MUSICA IASI retrieval simulator on model data is shown in direct comparisons of the new MUSICA IASI $\{H_2O, \delta D\}$ pair data with respectively processed data from COSMO_{iso}. In particular for dry conditions, a strong push of the COSMO_{iso} $\{H_2O, \delta D\}$ distributions towards those from MUSICA IASI can be observed. In total, the MUSICA IASI retrieval simulator is very promising for providing meaningful, direct model-to-satellite comparisons.

5 | Lagrangian Process Attribution of $\{H_2O, \delta D\}$ Pairs over West Africa

Lagrangian approaches offer a variety of analysis methods for deeper understanding the sources and pathways of atmospheric moisture. For instance, Lagrangian backward trajectory models have proven to be powerful tools for investigating water isotope abundances in precipitation and water vapor with respect to air mass sources and the evolution of meteorological parameters along air pathways (e.g. Rindsberger et al., 1983, Pfahl and Wernli, 2008, Soderberg et al., 2013, Salamalikis et al., 2015). By relying on a moisture source attribution of Sodemann et al. (2008), Dütsch et al. (2018) performed a sophisticated process attribution by estimating the relative contributions of surface evaporation, cloud formation and air mass mixing to δD and $\delta^{18}O$ distributions over Europe. As the identification of these processes is performed based on tendencies in H_2O and on the geographical properties (trajectory altitude, type of underlying surface), this approach is not able to robustly account for effects of rain evaporation and equilibration.

Therefore, this chapter develops a new Lagrangian process attribution that considers the paired $\{H_2O, \delta D\}$ evolution along trajectories and thereby allows for identifying effects of the moist processes presented in Sect. 3.1. This process analysis is based on the theoretical framework of Noone (2012) with the idealized process curves that indicate effects of mixing and rain processes in the $\{H_2O, \delta D\}$ phase space (see Fig. 3.2). Even though some of these simple models (e.g. the Rayleigh distillation for cloud formation) implicitly involve the Lagrangian perspective, the process curves of Noone (2012) have never been explicitly applied in combination with air parcel trajectories and their water isotope signals. The aim herein is to use these curves together with hydrometeorological parameters in order to link characteristic $\{H_2O, \delta D\}$ signals along trajectory ensembles to underlying processes such as air mass mixing and microphysical interaction between rain and vapor. As a result, this allows for interpreting the isotopic composition at the trajectory ending points by means of the attributed processes along the trajectory pathways.

As a case study to demonstrate the power of this Lagrangian process attribution procedure, the Sahelian mid-troposphere is used as starting domain for the backward trajectories during the WAM season 2016. For this purpose, the tool LAGRANTO is used based on data from COSMO_{iso}. The model configuration is chosen such that the results of this Lagrangian process attribution procedure provide a general framework to meaningfully interpret observed $\{H_2O,$

δD) distributions over the Sahel, as shown for the MUSICA IASI data in Fig. 4.16.

Section 5.1 motivates the concept of the Lagrangian process attribution procedure for interpreting the general evolution of $\{H_2O, \delta D\}$ pairs along the WAM development (Sect. 5.2). In this context, the $\{H_2O, \delta D\}$ pair signals along individual trajectories (Sect. 5.3) and along trajectory ensembles (Sect. 5.4) will be investigated. Finally, Sect. 5.5 discusses the Sahelian $\{H_2O, \delta\}$ pairs with respect to the history of the individual trajectory ensembles.

5.1. Data and Methodology

As basis for the Lagrangian process attribution the backward trajectories from LAGRANTO are considered, whose technical details are described in Sect. 3.4 and Table 3.2. By considering the COSMO_{iso} simulation CISO2 (14 km horizontal resolution, explicit convection, June – July 2016), the 7-day backward trajectories are calculated for the Sahelian mid-troposphere (625 and 575 hPa, domain shown in Fig. 3.7), starting daily at 9 and 21 UTC between 08 June to 30 July 2016. The analysis of the trajectories will rely on the WAM phases as defined by Knippertz et al. (2017) and discussed in Sect. 2.4.

5.1.1. Trajectory Sorting

For the interpretation of the full ensemble of 12 720 trajectories, the trajectories will be sorted according to geographical and meteorological criteria along their atmospheric pathways. By considering the geographical position and altitude of the trajectory origins, the aim is to build clusters that represent the dominant transport patterns of the WAM (see Sect. 2.2). As transport is an important control factor for atmospheric moisture, such a dynamical clustering will give a useful first overview of the characteristic moisture evolution of the defined clusters (Nieto et al., 2006, Salih et al., 2015, Sy et al., 2018).

Further, a trajectory will be classified as precipitating, if its rain content q_r accumulated over the whole pathway exceeds the rather arbitrarily chosen threshold of 2 g kg^{-1} . Analogously, if the accumulated rain content is lower than 0.2 g kg^{-1} , the trajectory is assumed to be non-precipitating. An individual trajectory data point is classified as precipitating, if the specific content (q_r for rain, q_s for snow) is at least $10^{-5} \text{ g kg}^{-1}$. If both q_r and q_s fulfil this criterion, this precipitation is viewed as mixed-phase. Following the moisture source attribution of Sodemann et al. (2008), a moisture uptake along a trajectory is attributed to surface evaporation if the corresponding trajectory altitude z_{tra} is below the boundary layer height z_{bl} . As models tend to underestimate the boundary layer height, Sodemann et al. (2008) recommended for this purpose a scaling of z_{bl} with a factor of 1.5.

5.1.2. Trajectory-Based Process Attribution of $\{H_2O, \delta D\}$ Pairs

The aim of this work is to establish a framework for interpreting the isotopic composition in a region of interest with regard to moist processes occurring during the transport of air masses arriving in this region. For this purpose, this chapter develops a process attribution procedure by considering temporal changes not only in H_2O , but also in δD , thereby making use of the additional isotope information. The general concept behind the Lagrangian process attribution procedure is the following:

1. A priori selection of dominant processes that will be identified in the $\{H_2O, \delta D\}$ phase space
2. Categorization of trajectories or individual segments along trajectories that correspond to the processes of interest
3. Interpretation of the isotopic composition in the region of interest by means of the categorized trajectories

In this context, the Sahel and its $\{H_2O, \delta D\}$ pair variability during the monsoon 2016 will be used as region of interest (Sect. 5.2). The processes of interest are defined according to the idealized process curves of Fig. 3.2 describing effects of air mass mixing, Rayleigh condensation, and Super-Rayleigh signals (step 1). To illustrate the plausibility of those curves, first, they will be evaluated against the isotopic evolution along three characteristic example trajectories (Sect. 5.3). Then, following Sect. 5.1.1, the trajectories will be sorted into dominant transport clusters, whose averaged $\{H_2O, \delta D\}$ signals are identified in order to link segments of trajectories to the proposed process curves (step 2, Sect. 5.4). In turn, this provides statistical information of the relative process occurrence frequencies within the full trajectory ensemble, which eventually facilitates an improved interpretation of the isotopic composition in the target region and the atmospheric processes that determine this composition (step 3, Sect. 5.5).

5.2. Average $\{H_2O, \delta D\}$ Development Along the West African Monsoon

Figure 5.1 displays the average $\{H_2O, \delta D\}$ behaviour over the Sahel at 600 hPa along the monsoon season 2016, as given by the COSMO_{iso} grid point values in the considered domain. The model data are sorted according to the WAM phases from Knippertz et al. (2017), i.e. the pre-onset stage (Phase 1), the post-onset stage (Phase 2), the wet westerly regime (Phase 3) and the monsoon recovery (Phase 4) (for details see Sect. 2.4). The respective $\{H_2O, \delta D\}$ pair data are summarized

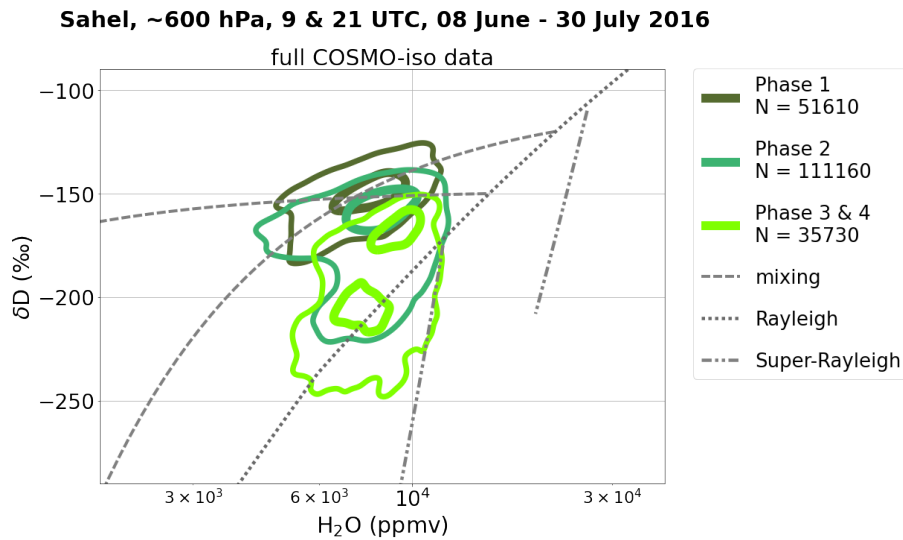


Figure 5.1. $\{H_2O, \delta D\}$ pair distributions at the grid points of COSMO_{iso} in the Sahelian mid-troposphere (600 hPa, see dashed box in Fig. 3.7) during 08 June–30 July 2016. The two-dimensional contours indicate the data distributions during the different monsoon phases as described by Knippertz et al. (2017). For each stage, the contours summarize 50 and 95 % of the respective data points and the total amount of data points is given in the legend. Additionally, this plot includes the idealized process curves that are marked with arrows in Fig. 3.2.

by normalized two-dimensional histogram contours (see Sect. 3.1.5).

While H_2O remains within a similar data range throughout the whole period, δD reveals a change with time. In particular, the development of the monsoon circulation in Phase 2 coincides with a shift of δD towards markedly lower values. The very wet Phases 3 and 4 show a slight shift to higher moisture content and lower δD with respect to Phase 2. This anti-correlation between H_2O and δD corresponds to the averaged intra-annual variability over the Sahel, as documented and discussed in Sect. 4.3.3.6 in the context of Fig. 4.16. Thus, these figures suggest that there may be moist processes whose effects can be observed more clearly in the paired $\{H_2O, \delta D\}$ phase space than in individual H_2O distributions. In this context, the backward trajectories serve to understand the mechanisms controlling the $\{H_2O, \delta D\}$ variability, as they shed light on the history of the air masses arriving at the target region.

5.3. Isotopic Process Attribution Along Single Trajectories

As a first step of the Lagrangian process analysis, the focus is on three individual trajectories: T1 as non-precipitating trajectory, T2 as trajectory with an accumulated q_r between 0.2 and 2 $g\ kg^{-1}$ and T3 as precipitating trajectory (see Sect. 5.1.1). This section analyzes their $\{H_2O, \delta D\}$ evolution to demonstrate how the idealized process curves from Fig. 3.2 can be interpreted for

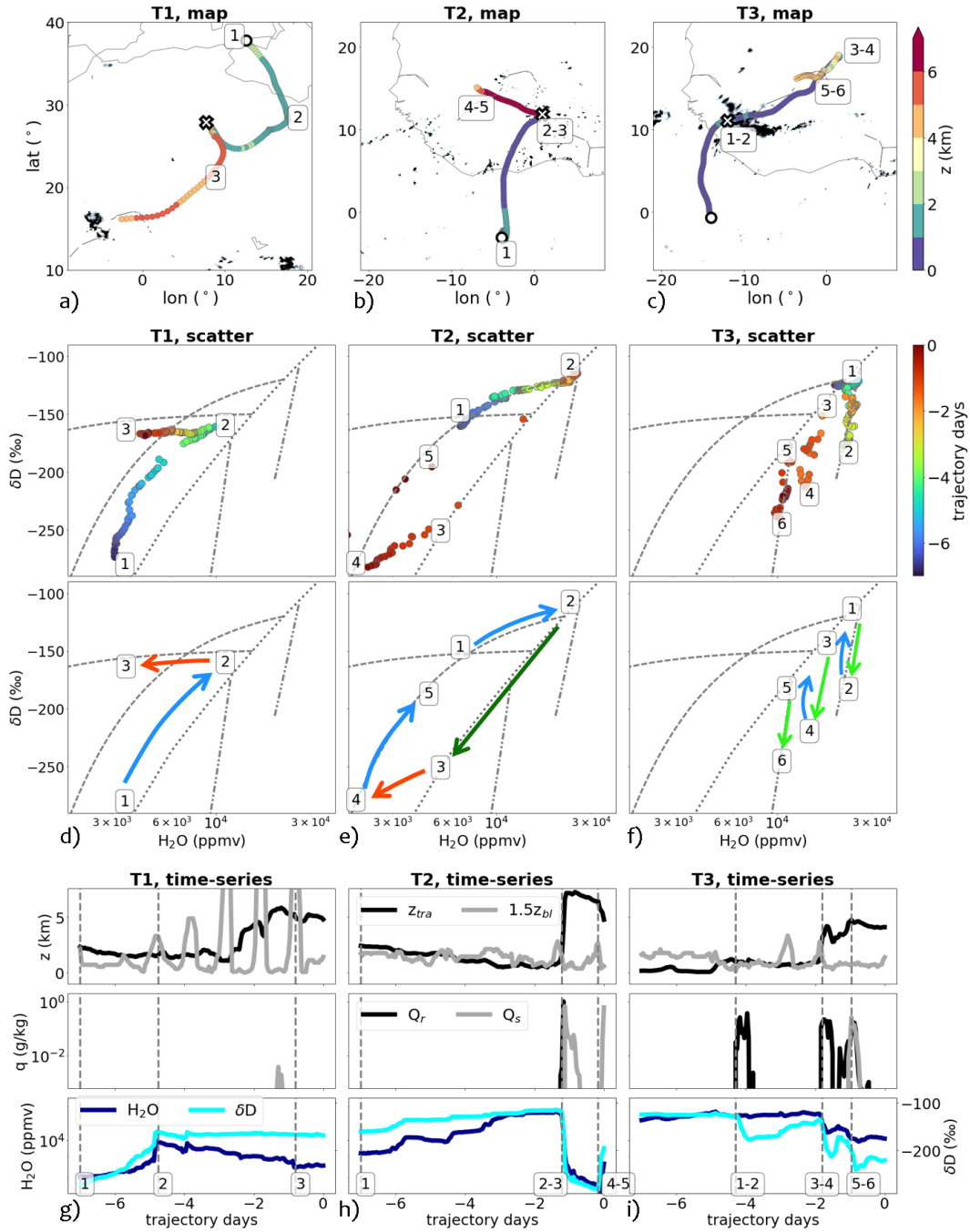


Figure 5.2. Overview of the trajectories T1, T2, and T3. (a), (b), and (c) show the geometrical pathways, color-coded with altitude. The black framed circle marks the trajectory origin. The black shades indicate areas with a rainfall $> 2 \text{ mm hr}^{-1}$ during the trajectory step marked with x. (d), (e) and (f) show the evolution in the $\{H_2O, \delta D\}$ phase space. The main variations between characteristic signals (marked with numbers) are illustrated with arrows (arrow colors according to Fig. 3.2). (g), (h) and (f) show the time-series of the altitude (z_{tra}), boundary layer height (z_{bl}), rain (q_r) and snow content (q_s), H_2O and δD along the respective trajectory.

individual air parcels. The target dates of T1, T2 and T3 are 09 UTC 05 July, 21 UTC 23 June and 09 UTC 19 July 2016. Figure 5.2 provides an overview of the properties of the chosen trajectories. The dominant mixing and microphysical processes along the trajectories are identified according to the temporal evolution of $\{H_2O, \delta D\}$ pairs and depending on the occurrence of hydrometeors. Arrows and markers illustrate the corresponding temporal evolutions in the $\{H_2O, \delta D\}$ phase space.

The first trajectory T1 does not show any significant precipitation along its pathway. It starts with a relatively dry signature at around 2.5 km above the Mediterranean Sea near Sicily (Point 1) and then crosses the Libyan coast (Point 2, Fig. 5.2a). During this time the trajectory moistens and enriches due to ocean evaporation. Therefore, in Fig. 5.2d the $\{H_2O, \delta D\}$ pairs follow the moist mixing line according to the blue arrow from marker 1 to 2. After this, dehydration sets in over the dry North African desert areas. At 30° N, T1 experiences a strong lifting to 4 km (see marker “x” in Fig. 5.2a), where it mixes with dryer mid-tropospheric air and moves southwestward. While H_2O decreases, δD remains mostly constant, leading to a mixing signature that follows the dry mixing curve (see orange arrow from 2 to 3).

T2 (middle column in Fig. 5.2) represents a trajectory with both strong mixing and precipitating effects. It originates in the lower troposphere ($z_{tra} \sim 3$ km) over the Gulf of Guinea and exhibits moistening and enrichment, while subsiding below 1 km and taking course towards the Guinea Coast. This moistening is associated with surface evaporation, while the trajectory penetrates into the boundary layer (Fig. 5.2h). This leads to an enrichment following the moist mixing line (see blue arrow from marker 1 to 2, in Fig. 5.2e) and results in higher moisture contents than for T1. Over the Sahel, a local convection event (see precipitation patterns in Fig. 5.2b) lifts the trajectory abruptly from ~ 1 to 6 km altitude (see marker “x” in Fig. 5.2b). As a consequence, precipitation forms and depletes the trajectory of its heavy isotopes in the vapor phase following a clear Rayleigh signature (dark green arrow from 2 to 3 in Fig. 5.2e). Thereafter, the air parcel appears to leave the convective cell and weak mixing with drier surrounding air occurs (orange arrow, from 3 to 4), leading to a less steep evolution than a pure Rayleigh process would imply. From 4 to 5, a slight moistening appears (blue arrow), as the trajectory subsides down to ~ 4.5 km.

For the third, precipitating trajectory T3 (right column in Fig. 5.2) the starting point is already associated with very moist and isotopically enriched conditions, as it is located near the surface of the tropical Atlantic ($z_{tra} < 200$ m) and is therefore strongly affected by surface evaporation over the relatively warm waters southwest of West Africa. After reaching the West African land mass, the trajectory crosses a westward propagating squall line (see precipitation patterns and marker “x” in Fig. 5.2c). Large rain drops fall through the air parcel, while the air parcel remains constantly at an altitude in the range of 1–2 km, hence without being lifted by convection. This suggests that this rain formed at higher altitudes and fell from above into the trajectory pathway.

While H_2O remains high during this event, δD shows a sharp drop by more than 50 ‰ (Fig. 5.2f, light green arrow from marker 1 to 2). This depletion is stronger than would be predicted using the Rayleigh model and thus penetrates into the Super-Rayleigh regime. Along its northeastward path over the Sahel, the trajectory first enriches, likely due to surface evapotranspiration (blue arrow from 2 to 3), until it finally interacts with a second squall line and exhibits once again a pull in the $\{\text{H}_2\text{O}, \delta\text{D}\}$ phase space towards the Super-Rayleigh regime (light green arrow from 3 to 4). However, at this time the air parcel is lifted to 4 km and changes its flow direction by 180° , consistent with the propagation direction of the squall line. A subsequent enrichment (blue arrow from 4 to 5) defines the isotopic composition for the injection into the next convective updraft, where the occurrence of snow particles (Fig. 5.2i) is accompanied by Super-Rayleigh signals (light green arrow from 5 to 6).

In summary, the analysis of the selected trajectories reveals that, by using the theoretical process curves from Fig. 3.2, the temporal evolution of $\{\text{H}_2\text{O}, \delta\text{D}\}$ pairs along air parcels can be divided into moist and dry mixing, drying and depletion due to Rayleigh condensation, and processes that deplete the vapor beyond the prediction by the Rayleigh model. Only by considering the whole isotopic history of an air parcel, it is possible to fully explain its target position in the $\{\text{H}_2\text{O}, \delta\text{D}\}$ phase space.

5.4. Isotopic Process Attribution Along Trajectory Ensembles

In the next step, the aim is to test the usefulness of the idealized process curves for interpreting larger trajectory ensembles during the monsoon period 2016. For this purpose, the trajectory ensemble is sorted into transport clusters (see Sect. 5.1.1) and investigate the importance of different processes along the individual clusters. This section addresses the question to which extent and for which meteorological conditions the mixing, Rayleigh, and Super-Rayleigh process curves from Fig. 3.2 are useful to explain the isotopic signals along the transport clusters. As these clusters are most representative during the active monsoon (see Fig. 5.5), the focus for the Lagrangian process attribution procedure in the following will be on trajectories during the post-onset stage (Phase 2).

5.4.1. Identification of Transport Clusters

As discussed in Sect. 5.1.2, the full ensemble of 12 720 trajectories is disentangled into meteorologically meaningful clusters of trajectories that experience a similar moisture history (analogous to the moisture transport pathways described in Sect. 2.2). Taking into account the characteristic regions of the trajectory origins as well as the relative position of origin altitude

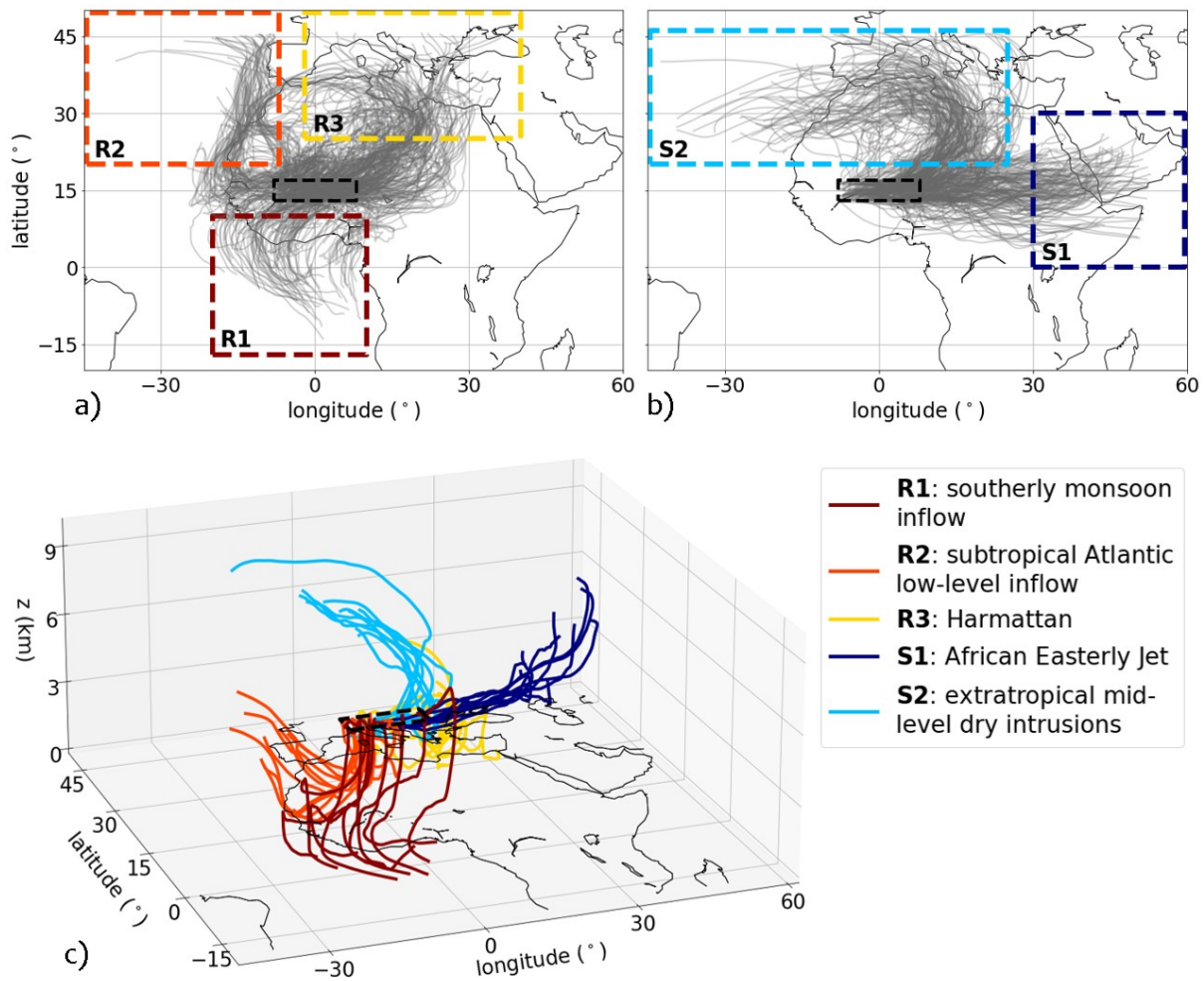


Figure 5.3. Trajectory clusters for (a) the monsoon inflow (R1, dark red box), the low-level subtropical Atlantic inflow (R2, orange), the Harmattan (R3, yellow), and (b) the African Easterly Jet (S1, dark blue) and the extratropical dry intrusions (S2, cyan). (c) shows a 3D view of selected trajectories for each transport cluster. The black box marks the target region over the Sahel.

against target altitude, a rough distinction is made for rising (R1 to R3) and subsiding (S1 and S2) transport clusters (see Sect. 5.1.1). Their main averaged properties (see Fig. 5.3 and 5.4) are briefly characterized:

R1 This cluster represents the *southerly monsoon inflow* (dark red trajectories in Fig. 5.3), originating from the lower troposphere over the Gulf of Guinea with high contents of H_2O and δD (see Fig. 5.4c and d). It advances on an anticyclonic path towards the Sahel, where it is lifted into the middle troposphere due to moist convection (see Fig. 5.4a; Marsham et al., 2013). This ascent into colder and dryer regions is associated with intense precipitation (Fig. 5.4e and f), leading to a strong depletion in δD (Fig. 5.4d).

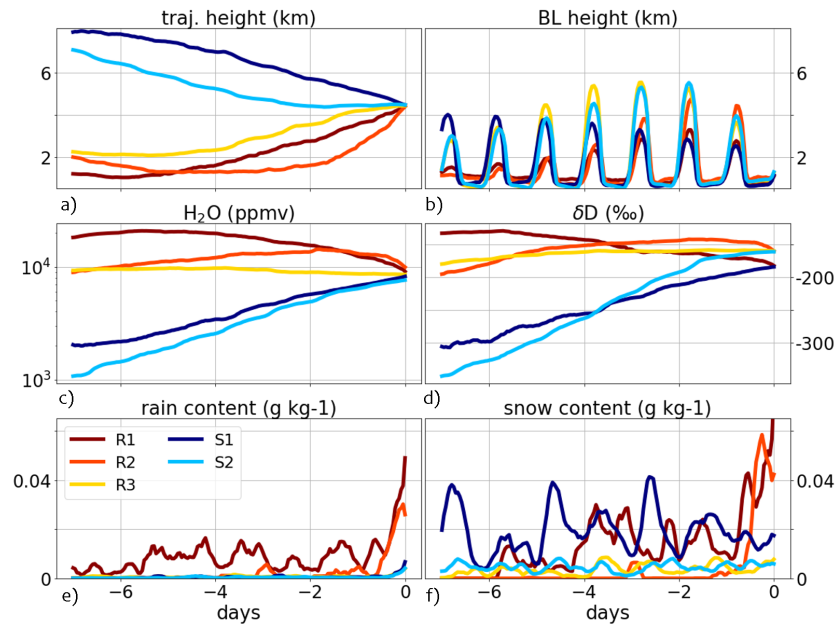


Figure 5.4. Averaged time series for the transport clusters R1 (dark red), R2 (orange), R3 (yellow), S1 (dark blue) and S2 (cyan). Shown are (a) the trajectory altitude, (b) boundary layer height ($1.5 z_{bl}$), (c) H_2O , (d) δD_v , (e) rain content q_r and (f) snow content q_s . For (e) and (f), a running mean over 6 time steps is used. To underline the daily cycles, in particular for the boundary layer height, only the trajectories with the same starting time of 09 UTC at the Sahelian mid-troposphere are herein considered.

- R2** The orange trajectories in Fig. 5.3 indicate the *subtropical Atlantic low-level inflow* with the trade winds that get deflected eastward towards the Sahel as a response to the Saharan heat low (Nieto et al., 2006, Lavaysse et al., 2009). Its initial moisture is lower than for R1, but increases during the transport over the Atlantic (Fig. 5.4c and d). Similar to R1, it experiences a convective lifting over the Sahel (Fig. 5.4a), but ends up with more enriched δD .
- R3** The trajectories in yellow (Fig. 5.3) originate in the lower troposphere over the Mediterranean Sea and follow the Etesian winds towards the African continent (Tyrlis and Lelieveld, 2013). Over North Africa, this cluster moves along the eastern side of the Atlas mountains and then feeds the relatively dry *Harmattan* (Hall and Peyrill e, 2006). As the surface evaporation over North Africa is small, there is hardly any change in moisture (Fig. 5.4c) as well as no significant contribution to the Sahelian precipitation (Fig. 5.4e,f). At the target location, it shows δD values similar to R2 (Fig. 5.4d).
- S1** The *African Easterly Jet* (AEJ) inflow is represented by the dark blue trajectories (Fig. 5.3). It is characterized by a low-latitude easterly flow that transports dry air masses from the upper troposphere (Fig. 5.4a) from East Africa down to the Sahelian mid-troposphere (Cook,

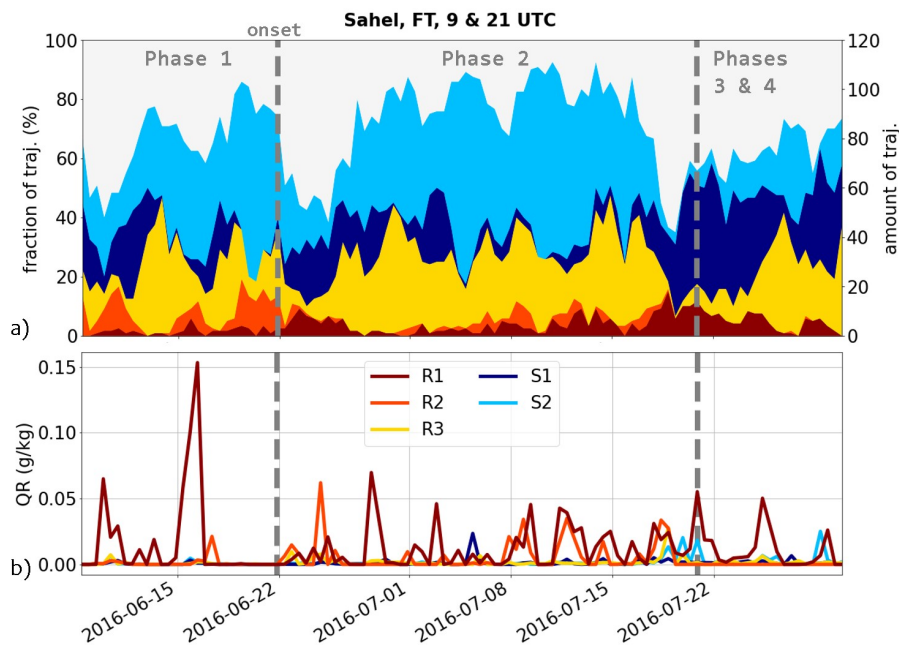


Figure 5.5. Time series of (a) the relative contribution of each transport cluster to the total amount of trajectories arriving every 12 hours in the target region in the Sahelian free troposphere. The monsoon phases described in Knippertz et al. (2017) are separated by dashed lines. (b) Rain water content accumulated along the last 12 hours of each trajectory before arriving in the target region and averaged over each cluster.

1999, Sy et al., 2018). Through deep tropical convection, frozen precipitation falls into the AEJ (Fig. 5.4f). During its subsiding path into moister tropospheric regions, H_2O and δD increase and converge towards the values of R1 (Fig. 5.4c and d).

S2 The cyan trajectory ensemble in Fig. 5.3 describes *extratropical mid-level dry intrusions*, which feed into the anticyclonic circulation above the Saharan surface heat low (Cook, 1999, Lavaysse et al., 2009). As this flow originates from the mid-latitude upper troposphere, it reveals very low moisture contents (Roca et al., 2005), even lower than for S1 (Fig. 5.4c and d). During its subsiding transition towards the Sahel (Fig. 5.4a), a strong moistening and enrichment takes place. The reason for this evolution is the elevated boundary layer height over the Sahara (Hall and Peyrill e, 2006), which mixes water vapor into S2 (Fig. 5.4b). At its target position, its δD resembles the values of the rising extratropical clusters (R2 and R3).

Figure 5.5a shows the relative contributions of each transport cluster to the target region as a function of time. The clusters represent together up to 90 % of the air transported into the Sahelian mid-troposphere. The unclassified trajectories mainly originate above the West African continent with no characteristic large-scale transport. Even though the relative contribution of the monsoon inflow (R1) is comparably low in terms of number of trajectories ($< 10\%$), it is

nonetheless the major driver of precipitation for the Sahel during the post-onset stage (Phase 2, e.g. compared to R2 in Fig. 5.5b). The pre-onset Phase 1 shows marked fluctuations associated with synoptic-scale disturbances described in Knippertz et al. (2017), leading to single rainfall events during June (e.g. Maranan et al., 2019). As the monsoon has not fully developed yet, the fraction of trajectories from the subtropical Atlantic (R2) is higher than in the other phases (Fig. 5.5a). The actual monsoon onset is characterized by a breakdown and then re-establishment of the AEJ as indicated by the dark blue trajectories S1. The fraction of monsoon trajectories in Phase 2 clearly increases compared to Phase 1 and precipitation events are now more frequent (Fig. 5.5b). Finally, the unusual flow situation during Phase 3 (and to a lesser extent Phase 4) is reflected in a clear shift of the fractions of transport clusters. Extratropical intrusions almost disappear entirely with a surge in AEJ inflow. The monsoon inflow, which causes marked precipitation events, increases at the expense of the Harmattan inflow.

In summary, the trajectory clustering according to their source regions reflects well the major transport contributions for the Sahelian troposphere during the monsoon season 2016. The clusters separate the trajectories into rising and subsiding transport patterns that bring moist and dry air masses to West Africa from different regions.

5.4.2. Mixing Processes

As discussed in Sect. 5.3, air mass mixing plays a crucial role in evaluating the isotopic evolution along a trajectory, in particular if no rain processes occur. Therefore, to extract pure mixing effects in the $\{H_2O, \delta D\}$ phase space, all non-precipitating trajectories are selected (see Sect. 5.1.1).

Figure 5.6 shows the $\{H_2O, \delta D\}$ pair data along the non-precipitating trajectories for each transport cluster. Even though the rising clusters R1 and R2 show on average strong occurrences of precipitation along their pathways (see Fig. 5.4e), still non-precipitating trajectories appear for both (7 and 56 %). The non-precipitating trajectories of R1, R2, and R3 show clear isotopic signals towards the moist mixing line. Moisture uptake from ocean evaporation and continental evapotranspiration represents a very moist mixing member and is opposed to the relatively dry conditions in the free troposphere. For instance, the non-precipitating trajectories in the monsoon cluster R1 start with very moist and enriched values above the Gulf of Guinea and subsequently mix with the drier and more depleted mid-tropospheric air masses while they advance over West Africa. Similar mixing structures are apparent for the Atlantic inflow (R2) and the Harmattan (R3), with substantially larger numbers of non-precipitating trajectories. As their initial moisture is much more variable than for R1, both moistening and drying occur along the non-precipitating trajectories of R2 and R3, closely following the moist mixing curve.

For the subsiding clusters S1 (AEJ) and S2 (extratropical intrusions), the non-precipitating trajectories are predominant ($\sim 90\%$). As they typically originate in the upper troposphere,

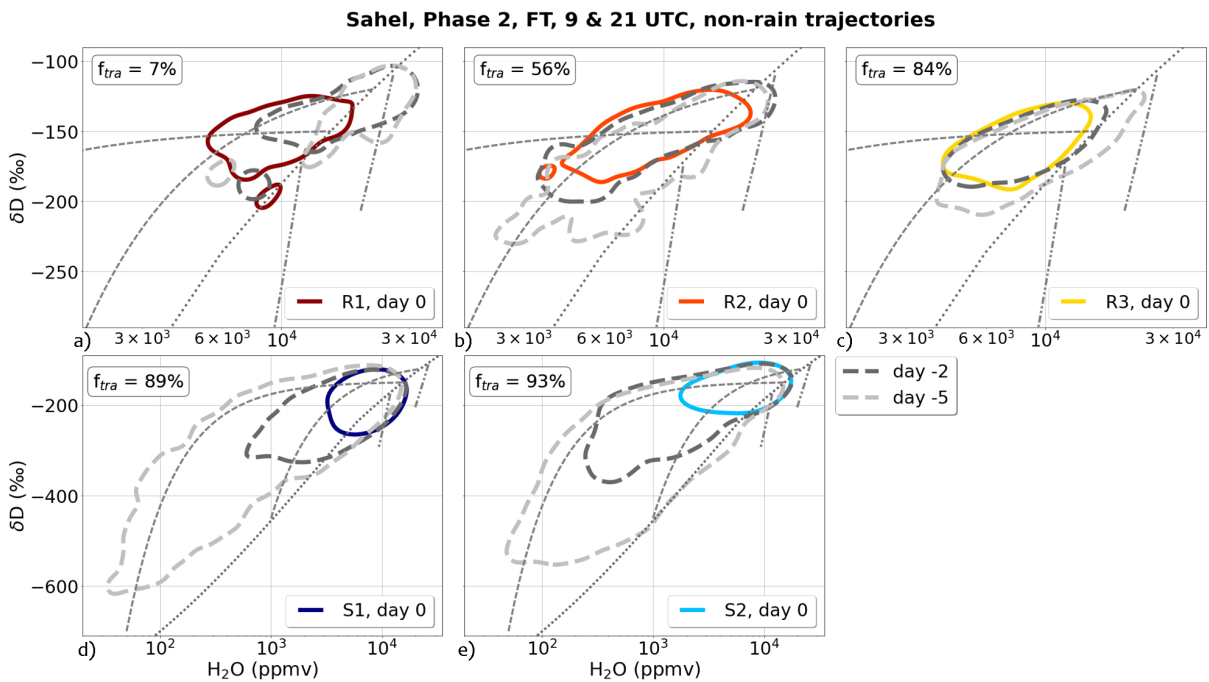


Figure 5.6. $\{H_2O, \delta D\}$ pair distributions for the non-precipitating trajectories of each transport cluster. The relative fractions f_{tra} of corresponding trajectories in each cluster are given in the respective plots. The solid, colored contours comprise 95 % of the data for the last 24 hours of the trajectory before reaching the target region (day 0), the dashed, dark gray contours the data of 2 days before arrival and the light gray contours the data of 5 days before arrival. The underlaid gray process curves are the same as in Fig. 3.2. Note the much larger axis ranges shown in the bottom two panels.

their starting points constitute very dry and depleted end members, while in this case the mid-tropospheric air masses act as moister end members. Thus, during the subsidence of S1 and S2 strong signals along the dry mixing curve develop, until the moisture approaches values similar to the target moisture of the rising trajectories.

In summary, even though the non-precipitating trajectories of the rising and subsiding transport clusters start with significantly different isotopic signals, mixing homogenizes to first order their $\{H_2O, \delta D\}$ pairs when arriving over the Sahel. Dehydration and moistening along the respective trajectories is well described by the theoretical moist and dry mixing curves.

5.4.3. Rayleigh Processes

To identify Rayleigh processes along the transport clusters, the focus is now on the precipitating trajectories (see Sect. 5.1.1). Thereby only the transport clusters R1 and R2 are considered, because only these two clusters include trajectories that exhibit a significant rain amount and therefore fulfill the rain criterion (see Fig. 5.5b).

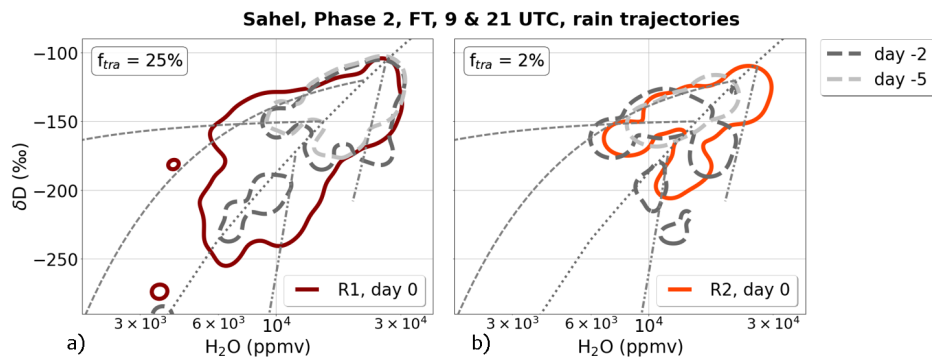


Figure 5.7. $\{H_2O, \delta D\}$ pair distributions analogous to Fig. 5.6, but for the precipitating trajectories of transport clusters R1 and R2 only.

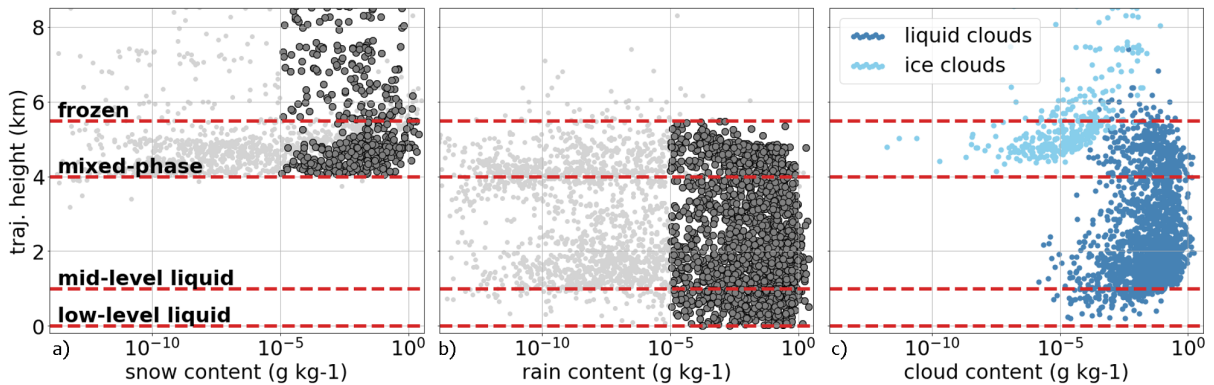


Figure 5.8. Classification of precipitating segments along the precipitating trajectories of R1. (a) shows snow content plotted against the trajectory altitude and (b) the respective rain content. The dark gray scatter indicate the data points, where the respective precipitation content exceeds $10^{-5} \text{ g kg}^{-1}$. In (c), liquid and frozen cloud contents are plotted against trajectory altitude.

In addition to signatures along the moist mixing curve, a clear Rayleigh signal is evident for both clusters (see Fig. 5.7). In particular during the last 24 hours before arrival, when the convection peaks, the $\{H_2O, \delta D\}$ pairs are distributed along the theoretical Rayleigh curve and indicate a depletion that cannot be explained with the mixing curves alone. Additionally, also values appear below the Rayleigh curve. This is either due to further Rayleigh processes with curves that are shifted towards lower δD values (which is rather unlikely, since the plotted Rayleigh curve is already chosen for relatively high surface temperature and relative humidity, see Fig. 3.2), or there are processes that lead to an enhanced depletion and create signals in the Super-Rayleigh area, as documented for trajectory T3 in Sect. 5.3.

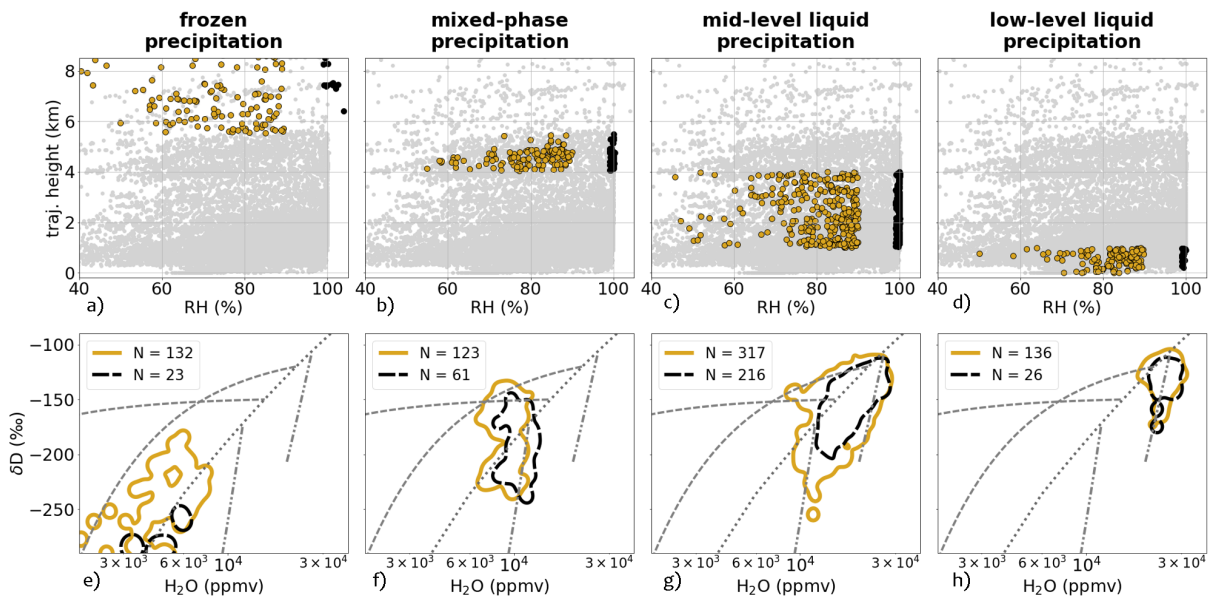


Figure 5.9. $\{H_2O, \delta D\}$ pair distributions of water vapor for the different classes of precipitating segments within the transport cluster R1, further separated according to relative humidity. (a)–(d) relative humidity against trajectory altitude with black (golden) dots indicating the saturated (unsaturated) conditions. (e)–(h) contours that summarize 95% of the $\{H_2O, \delta D\}$ pair data of the ambient water vapor for the saturated (black) and unsaturated (golden) points in the different precipitation classes.

5.4.4. Super-Rayleigh Processes

This section sheds light upon the Super-Rayleigh signatures in the $\{H_2O, \delta D\}$ pair distributions that develop during the precipitating ascent of the monsoon flow (R1) into the Sahelian mid-troposphere.

For this purpose, this section considers the precipitating trajectories of R1 and further classifies the precipitating points of the corresponding trajectories based on the altitude and phase of the precipitation (see Fig. 3.1). Following the precipitation thresholds of Sect. 5.1.1, a distinction is made for frozen (q_s), mixed-phase (q_s and q_r), and mid- and low-level liquid (q_r) precipitation. The respective altitude ranges are shown in Fig. 5.8. While the low-level liquid class represents the near-surface and mostly sub-cloud rain, the other precipitation classes can go along with liquid or frozen clouds (Fig. 5.8c). Further, we distinguish between saturated ($RH \geq 99\%$) and unsaturated ($RH \leq 90\%$) conditions of the ambient vapor (Fig. 5.9a–d). In line with the saturation adjustment of COSMO_{iso}, the unsaturated data points are cloud-free (near-cloud points may exhibit minor cloud contents due to the 3D interpolation when tracing q_c and q_i along the trajectories).

Figures 5.9a, e depict the trajectory points with frozen precipitation q_s , with only few corresponding data points appearing within the chosen axis range of the $\{H_2O, \delta D\}$ plot. These snow

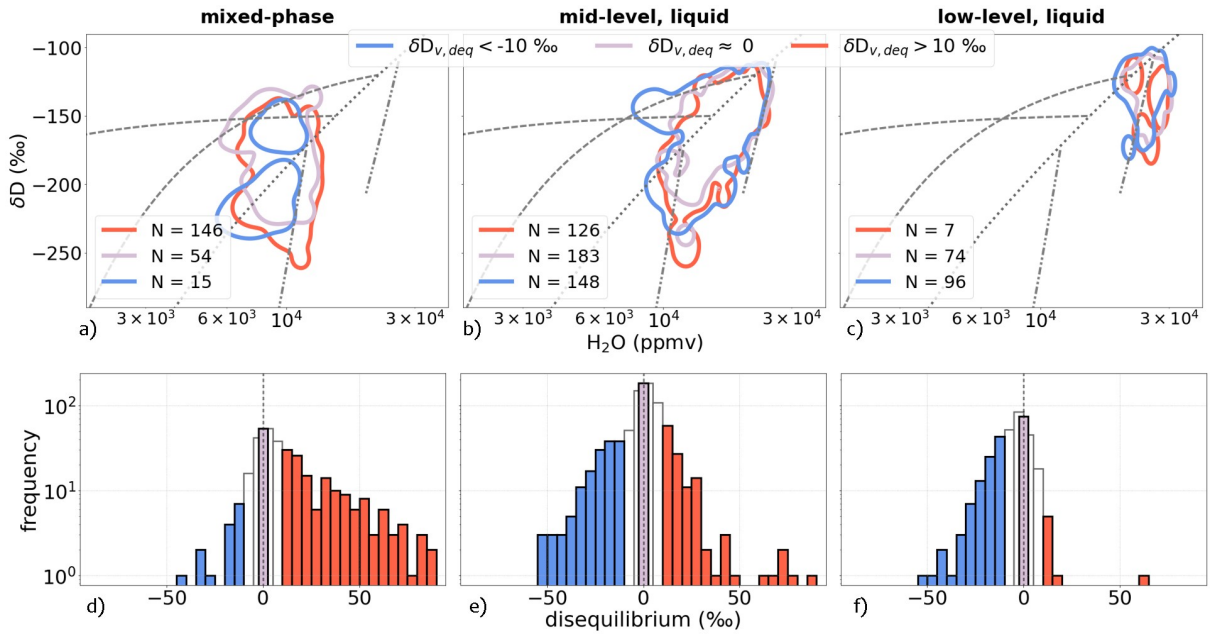


Figure 5.10. Isotopic disequilibrium $\delta D_{v,deq}$ between water vapor and rain drops for the precipitation classes from Fig. 5.8 where liquid precipitation occur. In (a)–(c), the blue contours summarize 95 % of the $\{H_2O, \delta D\}$ pairs with $\delta D_{v,deq} < -10 \text{ ‰}$ (vapor is significantly more depleted in heavy isotopes than the equilibrium vapor from precipitation), the lilac contours pairs with $|\delta D_{v,deq}| < 2.5 \text{ ‰}$ (no significant disequilibrium), and the red contours pairs with $\delta D_{v,deq} > 10 \text{ ‰}$ (vapor is significantly more enriched in heavy isotopes compared to the equilibrium vapor from precipitation). (d)–(f) show the histograms for $\delta D_{v,deq}$ for the corresponding classes of precipitation and disequilibrium.

particles are assumed to have formed at high altitudes, for instance within the deep convective parts of mesoscale convective systems, and fall through the trajectories on their way down. As sublimation is assumed not to fractionate, the isotopic composition of the ambient vapor gathers around the Rayleigh condensation curve with no significant Super-Rayleigh signals.

If these snow particles fall into the melting zone, where the air temperature is around $\sim 0^\circ \text{C}$, an area of mixed-phase precipitation develops (Fig. 5.9b, f). The melting process itself is non-fractionating, but it initiates fractionating interactions between the newly formed liquid drops and the ambient vapor. The contours in the $\{H_2O, \delta D\}$ phase space reveal that for both saturated and unsaturated conditions Super-Rayleigh signatures appear. Even in saturated conditions an isotopic flux can occur and equilibrate the rain drops with the ambient vapor. In case of sub-saturation, rain evaporation can take place. Both effects have the potential to further deplete the water vapor (see Sect. 3.1) and may thus explain the depleted Super-Rayleigh signatures inside the melting zone.

During the sedimentation of the liquid drops through a convective system, the Super-Rayleigh signatures are less pronounced for the saturated trajectory points, but are still notably pronounced

for unsaturated and cloud-free conditions (Fig. 5.9c, g). This depletion results, for instance, from rain evaporation in the unsaturated area below the stratiform cloud shield of a squall line (Fig. 3.1).

Figures 5.9d, h show the trajectory points, where rain drops occur near the surface and below the convective cloud base. Here, the air parcels are mostly unsaturated and indicate sharp tendencies towards the Super-Rayleigh area. In agreement with Risi et al. (2008a), this hints towards effects of sub-cloud rain evaporation in unsaturated downdrafts (Fig. 3.1).

To further improve the distinction between contributions of rain evaporation and isotopic equilibration, the degree of isotopic disequilibrium $\delta D_{v,deq}$ between the vapor and the liquid condensate is analyzed (see Eqn. 3.7). Figure 5.10a–c show the distribution of $\delta D_{v,deq}$ in the $\{H_2O, \delta D\}$ phase space for the precipitation classes from Fig. 5.8 with liquid precipitation. Even though the $\{H_2O, \delta D\}$ pair distributions for negative, low, and positive disequilibrium are highly similar (Fig. 5.10a–c), the corresponding histograms of $\delta D_{v,deq}$ show fundamental differences (Fig. 5.10d–f). In the melting zone, a strong imbalance towards positive values of $\delta D_{v,deq}$ appears, i.e., the vapor in equilibrium with the rain drops would be more depleted than the actual vapor (Fig. 5.10d). These rain drops have formed from melting snow and therefore reflect the isotopic composition at the condensation altitudes of the snow particles. Within deep convective systems, condensation can occur up to the tropopause level, imprinting highly depleted signatures on the precipitation (Celle-Jeanton et al., 2004, Risi et al., 2008a). If saturated conditions prevail at lower altitudes, the fall of this precipitation with low δD_r through a region with relatively higher δD induces an equilibrating isotopic flux from the vapor to rain. This decreases δD , while H_2O remains constant, contributing to the development of Super-Rayleigh signals within the melting zone (Fig. 5.10a). With decreasing height, the histogram of $\delta D_{v,deq}$ shifts to lower values (Fig. 5.10e), while equilibration and rain evaporation proceed and reduce the grade of disequilibrium (see Sect. 3.1). Eventually, in the sub-cloud zone the imbalance in $\delta D_{v,deq}$ changes sign (Fig. 5.10f), featuring equilibrium vapor from precipitation with a higher δD than the sub-cloud vapor. As here unsaturated conditions prevail (Fig. 5.9h), rain evaporation is strongly enhanced, leading to an enrichment of heavy isotopes in the rain drops and as a consequence to negative $\delta D_{v,deq}$.

Figures 5.9 and 5.10 reveal another interesting feature with more enriched values along the mixing curves (at around -150 ‰). In particular for the mid-level liquid precipitation, a clear mixing signal stands out that correlates with sub-saturation (Fig. 5.9g) and negative disequilibrium (Fig. 5.10b). The suggestion is that this feature is a result of synoptic-scale intrusions that transport dry and depleted air masses as rear-to-front flow into a convective system (Fig. 3.1; Kurita, 2013).

In summary, to account for the Super-Rayleigh signals in water vapor in the presence of precipitation, it is not sufficient to think of an isolated process but rather to consider the full interaction of microphysical processes that occur within and around a convective cell. The

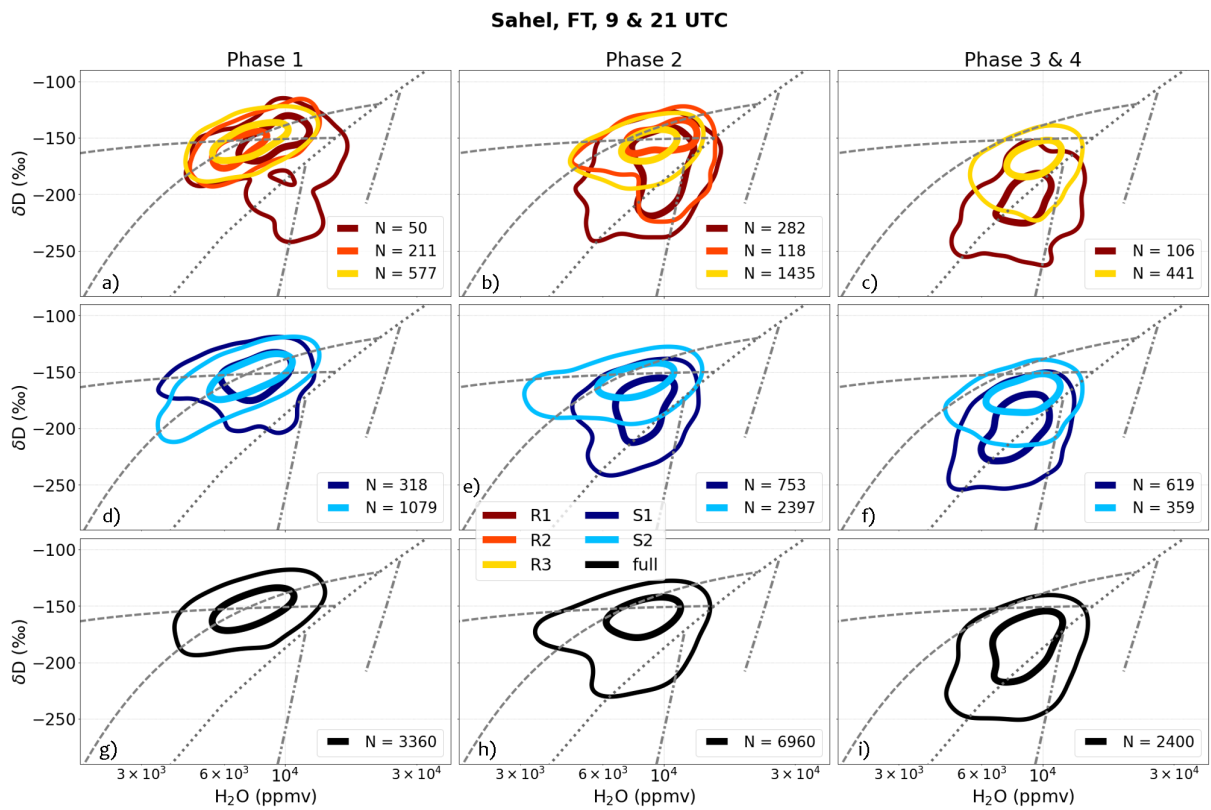


Figure 5.11. $\{H_2O, \delta D\}$ pair distributions in the target region (Sahelian free troposphere) for each transport cluster (colors, see Fig. 5.3) and monsoon phase in 2016 (columns, see Sect. 2.4). The shown contours mark 50 and 95% of all data points for (a)–(c) clusters R1, R2, and R3, (d)–(f) clusters S1 and S2, and (g)–(i) all trajectories.

depletion due to Rayleigh condensation during the convective updraft is superposed by additional depleting contributions of evaporation and equilibration of the falling rain drops. However, the two Super-Rayleigh lines marked in the $\{H_2O, \delta D\}$ phase space constitute rough bounds for the Super-Rayleigh area, as they frame the altitude range, where interactions between vapor and liquid precipitation can occur (from the melting zone to the surface).

5.5. Understanding $\{H_2O, \delta D\}$ Pairs over the Sahel

The discussion of Fig. 5.1 already showed that COSMO_{iso} simulates markedly different isotopic distributions in the Sahelian free troposphere during the three monsoon phases documented in Knippertz et al. (2017) (see Sect. 5.2). By combining the attributed signals in the $\{H_2O, \delta D\}$ phase space along the transport clusters, one can examine to which extent mixing and microphysical processes along the trajectory pathways explain this temporal evolution. To this end, Fig. 5.11 shows the same two-dimensional histogram contours for $\{H_2O, \delta D\}$ pairs as in Fig. 5.1 but now

separated by transport cluster. As a reference, results for the five transport clusters combined are added (Fig. 5.11g–i).

A direct comparison to Fig. 5.1 reveals that the $\{H_2O, \delta D\}$ pairs of the full trajectory ensemble (Fig. 5.11) are in line with the full COSMO_{iso} grid point values (Fig. 5.11). During the pre-onset stage (Phase 1) the $\{H_2O, \delta D\}$ distribution is governed by (mostly dry) mixing processes between moister and drier air masses that converge along the ITD. After the monsoon onset (Phase 2), convective processes (i.e., condensation, evaporation, and diffusive equilibration of rain drops) prevail and lead to a strong shift of δD towards the Rayleigh and Super-Rayleigh areas. This shift can temporally be enhanced during particularly wet monsoon periods (Phase 3 and 4). As the contours of the full trajectory ensemble result from contributions of each transport cluster (Fig. 5.5), their individual inspection allows for further disentangling the isotopic variability during the monsoon.

During the pre-onset stage (Phase 1; left column in Fig. 5.11) all clusters reveal robust and similar isotopic signals along the mixing curves. Despite the significantly different mixing history of the rising (R1, R2 and R3) and the subsiding (S1 and S2) clusters, two-way mixing harmonizes their $\{H_2O, \delta D\}$ pairs in the Sahelian troposphere. Only within the monsoon inflow (R1, red contours) occasional convective events create departures from the mixing curves towards the Rayleigh and Super-Rayleigh lines. The dominating extratropical intrusion cluster S2 (cyan contours, 1079 trajectories) agrees well with R2 and R3 but with a tendency towards drier and more depleted air. In contrast, the AEJ cluster S1 (dark blue contours) shows mild indications of both mixing along the dry mixing line and towards the Rayleigh and Super-Rayleigh area. Consequently, the contours for all trajectories combined are dominated by mixing along the moist mixing curve.

With the transition to the post-onset stage (Phase 2; middle column in Fig. 5.11), the frequent convection over the Sahel causes a general shift from relatively enriched air towards lower δD , while H_2O remains high. The condensation processes associated with the monsoon convection pushes the $\{H_2O, \delta D\}$ pair distributions of R1 and R2 towards the Rayleigh line. Additionally, the increased convection enhances effects such as diffusive equilibration and partial rain evaporation. As herein data are considered in the free troposphere, i.e. in the melting zone of falling snow particles, strong isotopic signals develop for R1 and R2 towards the lower Super-Rayleigh line. Because of the strong relation of monsoon precipitation with the air masses transported by the AEJ (Sy et al., 2018, Niang et al., 2020), the isotopic composition of cluster S1 merges with the signals of R1. By contrast, the northwesterly subtropical clusters R3 (Harmattan) and S2 (extratropical mid-level dry intrusions) remain around the mixing curves with only slight tendencies towards the Rayleigh curve. This emphasizes the existence of a subtropical mixing barrier that hinders the isotopic exchange between subtropical and tropical transport clusters as discussed in Yang and Pierrehumbert (1994) and Niang et al. (2020). The resulting contrast

between the effects of mixing and microphysical processes are well represented in the contours of the full ensemble.

Finally, in the unusually wet Phases 3 and 4 (right column in Fig. 5.11) the tendencies towards moister and less enriched air amplify. The monsoon inflow (R1) and the AEJ inflow (S1) further drop to lower δD , as convective processes increase and foster Rayleigh and Super-Rayleigh signatures. During this period convection is so widespread that also the sub-tropical clusters R3 and S2 show indications of reduced mixing and increased Rayleigh signals. As already shown in Fig. 5.5, the low-level Atlantic inflow cluster R2 is not present during these phases (Knippertz et al., 2017). The isotopic composition of all trajectories clearly reflects the shift from the mixing to the Rayleigh line with a marked extension towards the Super-Rayleigh area.

To summarize, the comparison of the $\{H_2O, \delta D\}$ pairs of the transport clusters from Fig. 5.11 against the COSMO_{iso} grid point values from Fig. 5.1 reveals that the identified process curves along different transport pathways provide a useful framework for better understanding the actual evolution of the isotopic composition in the Sahelian mid-troposphere during the WAM.

5.6. Summary

The aim of the proposed Lagrangian process attribution procedure is to provide a framework for interpreting the isotopic composition of tropospheric moisture in a chosen target region by means of individual moisture pathways. In this procedure, the evolution of paired distributions of H_2O and δD is tracked along backward trajectories. Analyzing the two-dimensional $\{H_2O, \delta D\}$ phase space, a separation of effects due to mixing and precipitation processes (condensation, evaporation, and equilibration) is possible by following the theoretical process curves of Noone (2012). They usually refer to single processes occurring along idealized air parcel trajectories. However, an application of these curves that explicitly identifies processes occurring along actual trajectories has never been done so far.

As a showcase for the Lagrangian process attribution, this chapter demonstrates how the interpretation of mid-tropospheric $\{H_2O, \delta D\}$ pair data over the Sahel during the West African Monsoon season 2016 can be improved by considering the past transport pathways and moist processes of inflowing air masses. For this purpose, data from a high-resolution, convection-permitting COSMO_{iso} simulation are considered and Lagrangian backward trajectories are computed starting in the Sahelian mid-troposphere. By analyzing the $\{H_2O, \delta D\}$ evolution along individual trajectories as well as clusters of trajectories, this allows for identifying characteristic effects of: (1) mixing of moist air masses that were enriched due to surface evaporation and moist advection against subsiding air masses that transport dry and depleted signals from the upper troposphere; (2) condensation associated with convection that follows the Rayleigh model; (3) partial rain evaporation and isotope equilibration of rain drops formed from melting snow

that both lead to a depletion of water vapor beyond the Rayleigh prediction, thereby accounting for the so-called Super-Rayleigh area. This complements earlier work from Worden et al. (2007) and Risi et al. (2010a), who attributed the enhanced depletion in tropical mid-level water vapor to rain evaporation and dry mixing.

In summary, the combination of the aforementioned processes, which are closely connected with the dominant transport pathways over West Africa, ultimately determine the prevailing signals in $\{H_2O, \delta D\}$ pairs in the Sahelian mid-troposphere at a given time. A publication presenting the new Lagrangian process attribution procedure is realized in Diekmann et al. (2021c).

6 | Analysis of $\{H_2O, \delta D\}$ Pair Data During the West African Monsoon

The prime intent of this section is to perform an in-depth analysis of the isotopic composition of water vapor in the West African troposphere along the WAM development. This is achieved through a synthesis of the previous results of this study, namely the generation of the new MUSICA IASI $\{H_2O, \delta D\}$ pair dataset together with the MUSICA IASI retrieval simulator from Chap. 4 and the trajectory-based analysis of process curves in the $\{H_2O, \delta D\}$ phase space from Chap. 5. The supplementary use of further simulations from ICON-ART_{iso} and COSMO_{iso} and of precipitation data from GPM IMERG (datasets technically discussed in Sect. 3.4) allows for studying the $\{H_2O, \delta D\}$ variability of the WAM and its control mechanisms on three different scales:

- Quasi-climatological analysis in order to put the WAM development in an intra-annual context
- Detailed investigation of seasonal monsoon development during the season 2016
- Convective-scale evaluation of isotopic impact from monsoonal precipitation events during WAM season 2016

For this purpose, the domains over the Guinea Coast and the Sahel are considered analogous to Knippertz et al. (2017), i.e. latitudinal bounds of 0° to 7.5° N and 7.5° to 15° N, respectively, with common longitudinal bounds of 8° W to 8° E. These domains are found to be useful for documenting the shift of maximum precipitation from the coastal to the Sahelian regions and thereby for defining the WAM onset (see Chap. 2). In this framework, the joint analysis of model and satellite data addresses the question, which model configurations are useful for reproducing the $\{H_2O, \delta D\}$ variability as observed from remote sensing and, in more general, the overall development of the WAM. Model settings to be investigated are the horizontal resolution and treatment of convection. Additionally, the benefits of applying the retrieval simulator will be emphasized. Together with the $\{H_2O, \delta D\}$ process curves from the Lagrangian process attribution, the overall research question of the multi-scale analysis is to examine, to which extent the paired $\{H_2O, \delta D\}$ distributions can provide reliable information about the underlying dynamical and

microphysical processes within the monsoon system.

Following the concept above, Sect. 6.1 analyzes the intra-annual variability of $\{H_2O, \delta D\}$ pair distributions over West Africa, Sect. 6.2 the seasonal development during the WAM period 2016 and Sect. 6.3 the variability on the convective scale.

6.1. Annual Cycle over West Africa

The aim of this section is to put the WAM development and its effect on the isotopic composition of mid-tropospheric water vapor into a quasi-climatological, intra-annual context. For this purpose, characteristic $\{H_2O, \delta D\}$ pair distributions from MUSICA IASI retrievals and free-running ICON-ART_{iso} simulations (IISO1–4, see Table 3.2) are investigated for the different seasons of the year. Precipitation distributions from GPM IMERG (RGPM2, see Table 3.2) and from ICON-ART_{iso} itself will provide supplementary diagnostic information about the meteorological conditions over West Africa.

6.1.1. Remote Sensing-Based Analysis

Figure 6.1 contrasts the averaged annual cycle of MUSICA IASI $\{H_2O, \delta D\}$ pair data with the monthly averaged GPM IMERG precipitation (RGPM2, see Table 3.2) for the Guinea Coast and the Sahel from 2015 to 2018. Specifically, the months February, May, August and November are chosen as representative periods for boreal winter, spring, summer and autumn. Additionally, the $\{H_2O, \delta D\}$ pair plots include the idealized process curves from Fig. 3.2. The horizontal distribution of monthly averaged precipitation (RGPM2) for the individual periods as chosen in Figure 6.1 are shown in Fig. 6.2.

Throughout the whole year, the Guinea Coast shows high H_2O and δD values (Fig. 6.1a). During the winter dry season, the ITD is located far south around the coastal region, leading to moderate precipitation (Fig. 6.2a) and to considerable contributions of the dry northeasterly Harmattan trade winds over the Guinea Coast. This can be seen in the $\{H_2O, \delta D\}$ pair data during February, where mainly clear mixing signals down to 2000 ppmv develop (blue contours in Fig. 6.1a). A drop in the averaged δD with a simultaneous increase in H_2O becomes apparent during the transition from the dry to the first rainy season (cyan contours in Fig. 6.1a), when the ITD moves northwards and allows a marked intensification of organized convection (Fig. 6.1c and 6.2b). As response, large-scale and intense precipitation patterns develop over the eastern Tropical Atlantic close to the coastal regions (Fig. 6.2c). This reduces the dry $\{H_2O, \delta D\}$ signals along the mixing curves and induces a shift towards the Rayleigh curve with even some Super-Rayleigh signals. These signals with high H_2O ($\sim 10\,000$ ppmv) and relatively depleted δD (below -200 ‰) remain as a weaker but persistent feature during the summer and autumn months, even though

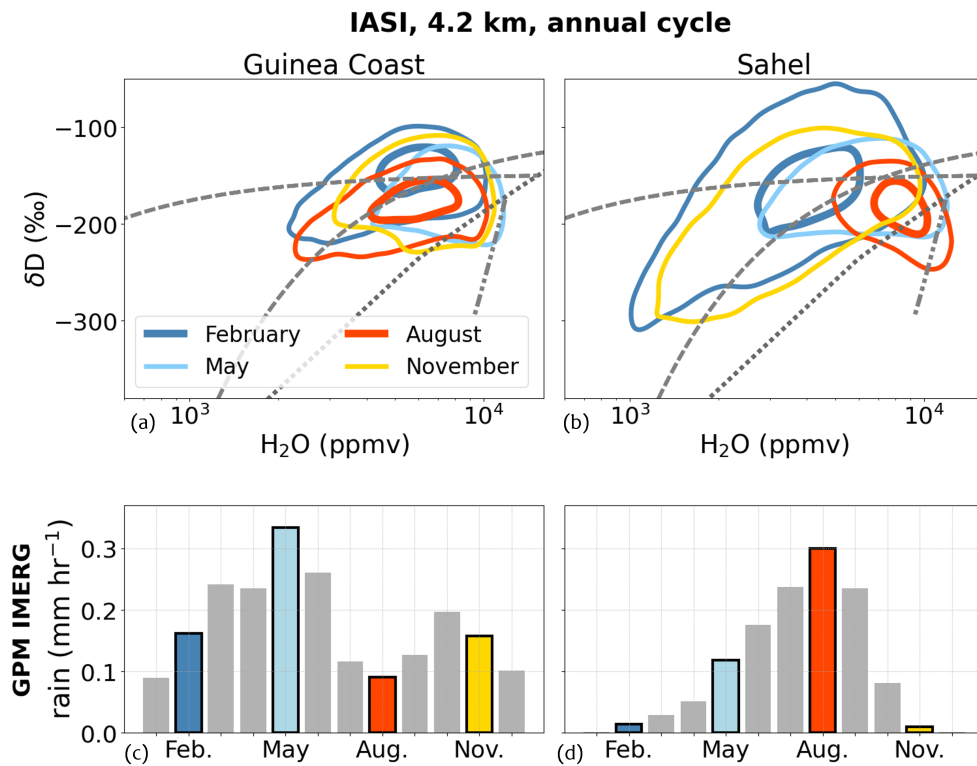


Figure 6.1. Upper row: MUSICA IASI $\{H_2O, \delta D\}$ pair distributions over (a) the Guinea Coast and (b) the Sahel, averaged for February, May, August and November for the period from 2015 to 2018. The 90 % and 50 % contours are drawn for each month. Lower row: Monthly averaged precipitation for the period 2015 to 2018 over (c) the Guinea Coast and (d) the Sahel (dataset RGPM2 from Table 3.2). The color-coded bars correspond to the time-periods of the respective contours in (a) and (b).

the precipitation reduces drastically during August (Fig. 6.1c and Fig. 6.2c). This little dry season during August goes along with a stretch of the $\{H_2O, \delta D\}$ pair contour towards lower H_2O and δD along the moist mixing curve (down to 2000 ppmv and -230 ‰), similar to the winter signals.

The reason for the coastal precipitation minimum during August is the shift of the ITD and maximum precipitation band towards the Sahel associated with the WAM onset, which leads to drastic contrasts between the Sahelian winter (dry) and summer (wet) season. Due to its more northern position, the Sahelian troposphere witnesses a stronger dehydration than the Guinea Coast during winter times. Here, the ITD is located at the coastal regions and there is almost no precipitation occurring over the Sahel (see Fig. 6.1d and 6.2a). The descending branch of the Hadley cell is found to feed dry and depleted air masses into the North and West African troposphere (Frankenberg et al., 2009), which are then transported by the far-south reaching Harmattan trade winds towards the Sahel. With only weak evaporation taking place, this creates a strong mixing signal in the $\{H_2O, \delta D\}$ distribution of the Sahelian troposphere with reaching

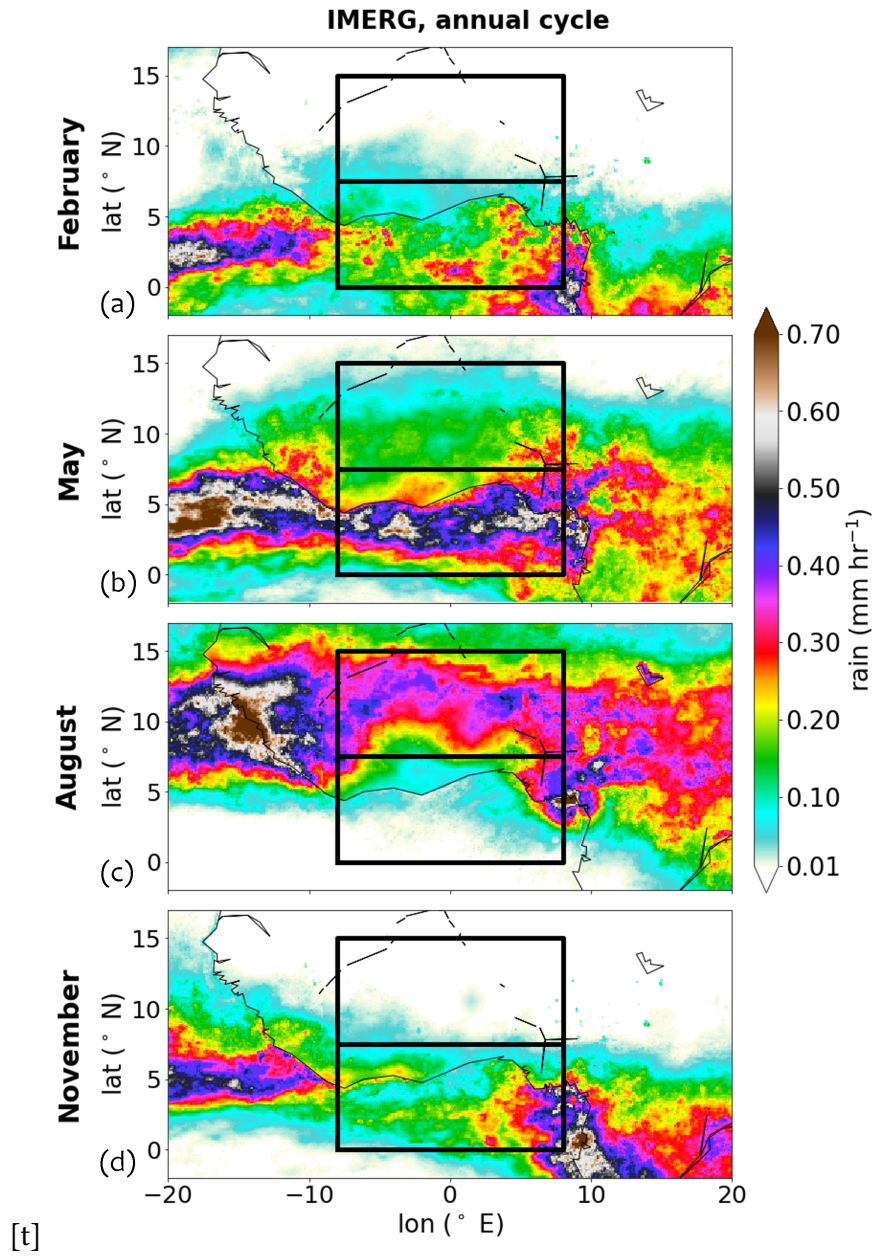


Figure 6.2. Horizontal distributions of monthly averaged precipitation from GPM IMERG (RGPM2 from Table 3.2) over West Africa for the colored months from Fig. 6.1. The black boxes denote the chosen target regions for the Guinea Coast (0° – 7.5° N) and the Sahel (7.5° – 15° N).

minimum values of 1000 ppmv for H_2O and -300‰ for δD (see blue contour in Fig. 6.1b). This is in agreement with Sect. 5.4.2, where this moist mixing curve was found to well represent the near-surface moistening along the Harmattan and the Atlantic westerly inflow.

Even though the Sahelian winter is mostly dry and rain-free, rare but heavy precipitation events may form as response to strong tropical-extratropical interactions between the AEJ, the monsoon southerlies and the Harmattan trade winds (e.g. Knippertz and Martin, 2005, Knippertz and Fink, 2009, Davis et al., 2013). Occasional northward shifts of the ITD allow the moist southerly winds from the Gulf of Guinea to reach (for boreal winter) unusually high latitudes and to thereby penetrate moisture towards the dry Sahel. This explains the occurrences of high moisture contents even in the dry winter season (up to 10000 ppmv), ultimately resulting in a large variability in the corresponding $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair distribution (see blue contour for February and yellow contour for November in Fig. 6.1b).

The winter distribution of Sahelian $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair data reveals another interesting feature in the range of 2000–8000 ppmv. The corresponding contour exhibits highly enriched δD values up to -60‰ that are far above the shown process curves. González et al. (2016) identified similar structures over Tenerife and found a correlation with high dust abundances that originate from the African continent and are transported within the Saharan air layer to Tenerife. Such a dust outbreak is associated with strong vertical mixing, injecting near-surface humidity efficiently to higher altitudes. This is in very good agreement with highly enriched MUSICA IASI $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair signals that are observed over Tenerife and are closely linked to high dust abundances (see Appendix A). As an outbreak of Saharan dust can even spread to the coastal West Africa through the strong northeasterly trade winds (Knippertz and Fink, 2006), the suggestion is that such dust events might explain also the highly enriched values over the Sahel during February. This would be in agreement with Noone (2012), who stated that the mixing curves shift to higher δD values, if the contributions of land sources (evapotranspiration) increase. Due to non-fractionating plant transpiration, evapotranspired water constitutes a more enriched mixing member than water evaporated from ocean surfaces (equilibrium fractionation).

As discussed above, the WAM is characterized by the shift of maximum precipitation from the Guinea Coast to the Sahel during summer (see Fig. 6.1c,d and 6.2b,c). The precipitation peak over the Sahel during August goes along with a strong increase in H_2O (data range between 5000–10 000 ppmv), whereas δD shows decreasing tendencies (see discussions of Fig. 4.16). This anti-correlation between precipitation amount and abundances of heavy water isotopes is often referred to as *amount effect* (e.g. Dansgaard, 1964, Lee and Fung, 2008, Risi et al., 2008a, Tharamal et al., 2017). Various mechanisms have been proposed for explaining this anti-correlation, such as dry mixing due to convective (Risi et al., 2008a) and meso-scale (Kurita, 2013) downdrafts as well as the occurrence of partial rain evaporation (e.g. Worden et al., 2007, Risi et al., 2010b,

Noone, 2012), which all have been identified to be pronounced within deep convection (Lacour et al., 2018). However, Fig. 6.1 shows that the enhanced depletion during summer is strongly associated with signals below the Rayleigh prediction and creates $\{H_2O, \delta D\}$ structures following the marked Super-Rayleigh line. In Sect. 5.4.4, this Super-Rayleigh regime was attributed to microphysical rain processes within the melting zone of a convective system (see Fig. 3.1), namely the evaporation of falling rain drops in unsaturated areas as well as the equilibration between water vapor and relatively depleted rain drops formed from melting snow for saturated conditions. Both effects are linked to convective activity and are therefore capable of explaining the correlation of the increased precipitation during summer with the generation of depleted Super-Rayleigh signals in the MUSICA IASI $\{H_2O, \delta D\}$ pair distributions.

In summary, the isotopic composition of mid-tropospheric water vapor reflects well the fundamentally different meteorological conditions over the Guinea Coast and the Sahel throughout the year. The coastal domain shows constantly high values in H_2O and δD , with only minor impacts due to convective activity towards slightly higher H_2O and lower δD (during the rain seasons in spring and autumn) and due to dry intrusions mainly towards lower H_2O (during the dry seasons in winter and summer). In contrast, the Sahel exhibits a tremendous variability in the $\{H_2O, \delta D\}$ phase space: During the very dry and mostly rain-free winter months the isotopic composition largely reflects air mass mixing with large contrasts in H_2O and δD , whereas the wet and rainy monsoonal summer period is characterized by strong signals of microphysical rain interactions.

6.1.2. Model-Based Analysis

This section provides a model-based analysis of the intra-annual variability of $\{H_2O, \delta D\}$ pairs and precipitation for the Guinea Coast and the Sahel. For this purpose, the multi-year simulations of ICON-ART_{iso} IISO1–4 (see Table 3.2) are evaluated for the period 2017–2019 analogous to the MUSICA IASI $\{H_2O, \delta D\}$ pair data and the GPM IMERG precipitation data, as discussed in the previous section. The only difference is the change of month for representing the boreal summer, from August to July, as this better represents the Sahelian precipitation maximum in ICON-ART_{iso} (see discussions and figures below). However, this section does not intend to quantitatively evaluate a direct one-to-one comparison of the individual seasons as observed from IASI and modeled by ICON-ART_{iso}, as for the latter a free-running setup was chosen (nudging of atmospheric variables not yet implemented). Furthermore, as mentioned in Sect. 4.4, the MUSICA IASI retrieval simulator is not yet capable of processing data from ICON-ART_{iso}. Thus, the aim of the following section is to analyze the averaged annual cycle as directly modeled by ICON-ART_{iso} and its capability of reproducing the general $\{H_2O, \delta D\}$ pair behaviour from

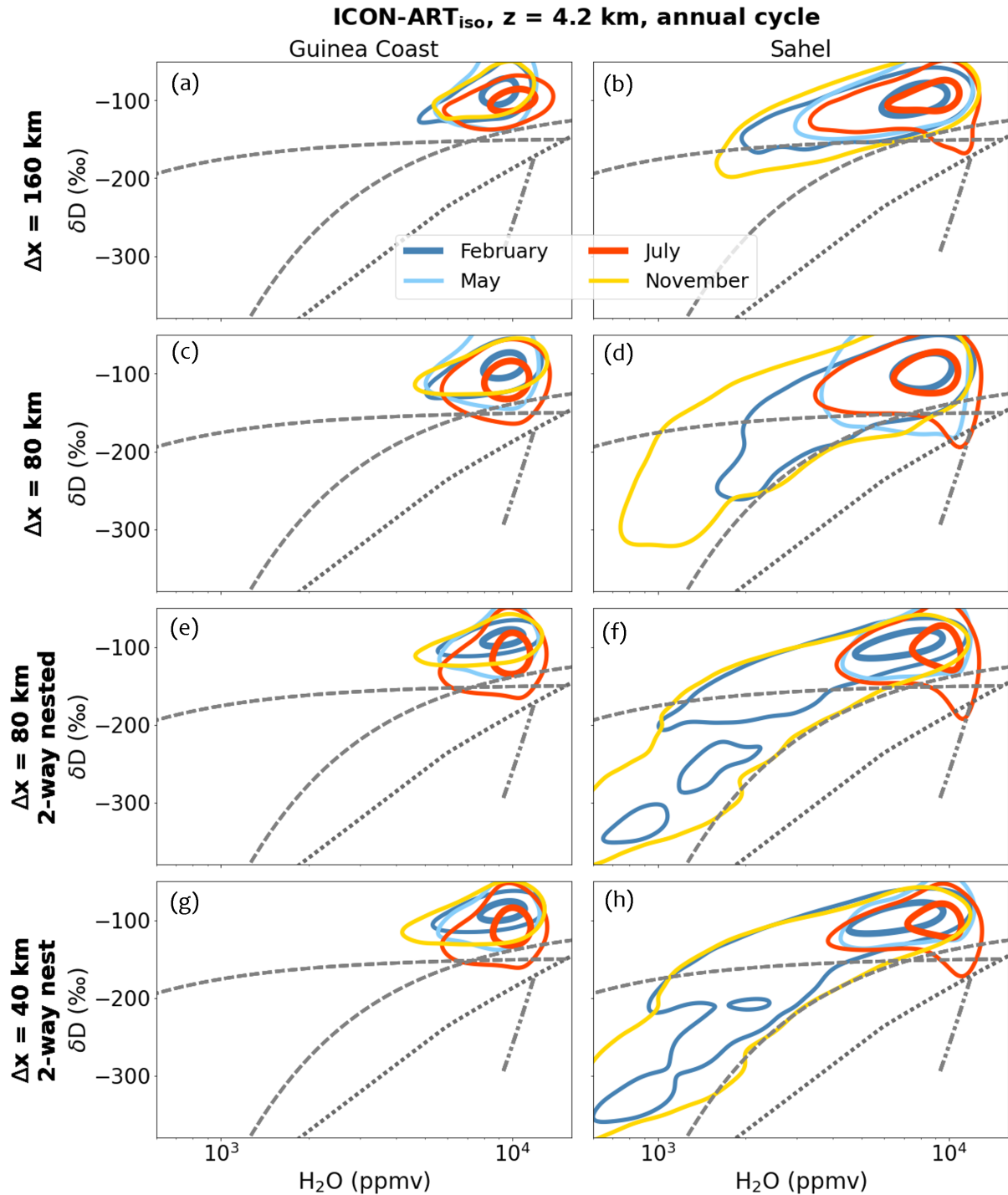


Figure 6.3. $\{\text{H}_2\text{O}, \delta D\}$ pair data from the ICON-ART_{iso} simulations (a) IISO1, (b) IISO2, (c) IISO3 and (d) IISO4 (see Table 3.2), monthly averaged for the years 2017–2019. The contours and seasons are analogous to Fig. 6.1, except for the summer contour (red), which here shows the data for July, as for the ICON-ART_{iso} simulations the maximum precipitation sets in earlier compared to GPM IMERG (compare Fig. 6.4 with Fig. 6.1c,d).

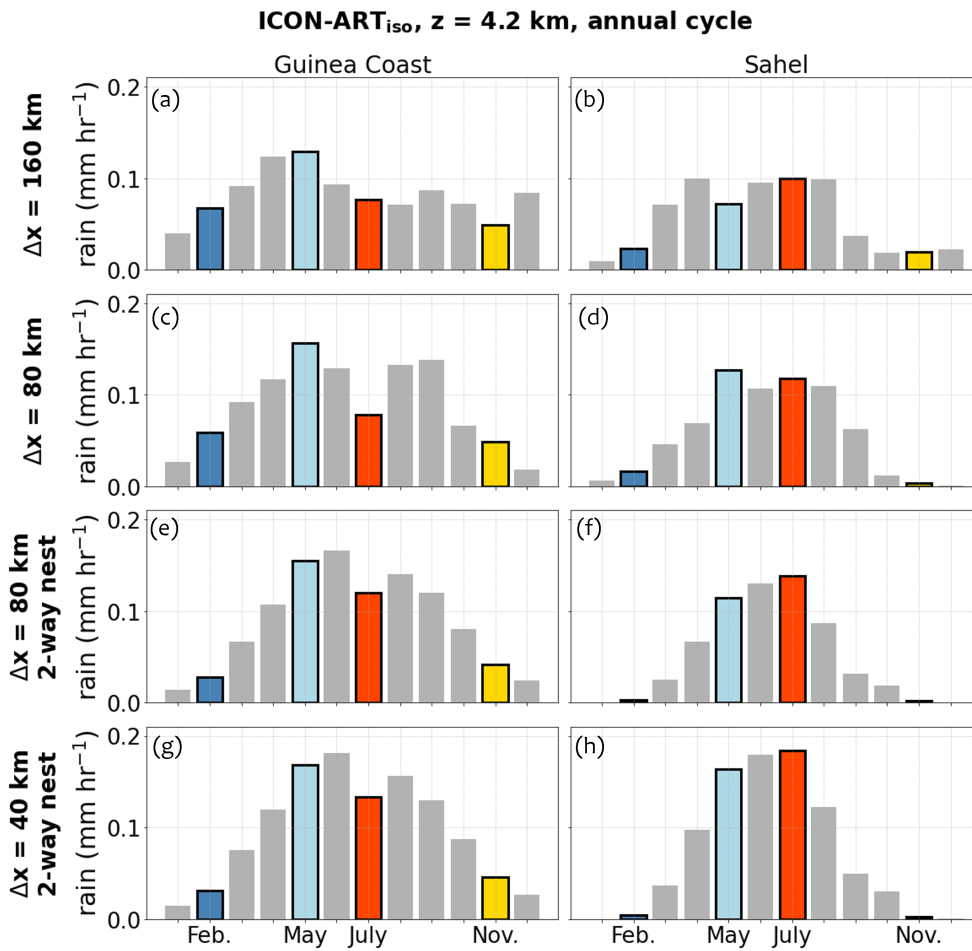


Figure 6.4. Monthly averaged precipitation for 2017–2019 from the ICON-ART_{iso} simulations (a) IISO1, (b) IISO2, (c) IISO3 and (d) IISO4 (see Table 3.2). The colored bars correspond to the months chosen in Fig. 6.3.

MUSICA IASI, with both aspects being addressed from a rather qualitative perspective.

The distinctive feature of the different ICON-ART_{iso} simulations IISO1–4 (see Table 3.2) is mainly the choice of horizontal resolution. IISO1, IISO2 and IISO3 are run globally with 160 and 80 km grid cell size, whereas IISO4 represents a nested, refined domain (40 km grid cell size) over Africa with a two-way-interaction to the global parent domain of IISO3 (see Sect. 3.4). Analogous to the IASI data in Fig. 6.1, Fig. 6.3 contrasts the annual cycle of mid-tropospheric $\{H_2O, \delta D\}$ pairs over the Sahel with the respective data over the Guinea Coast for the different ICON-ART_{iso} simulations. Figure 6.4 shows the monthly averaged precipitation and Fig. 6.5 the corresponding horizontal distributions of precipitation for each ICON-ART_{iso} simulation, respectively.

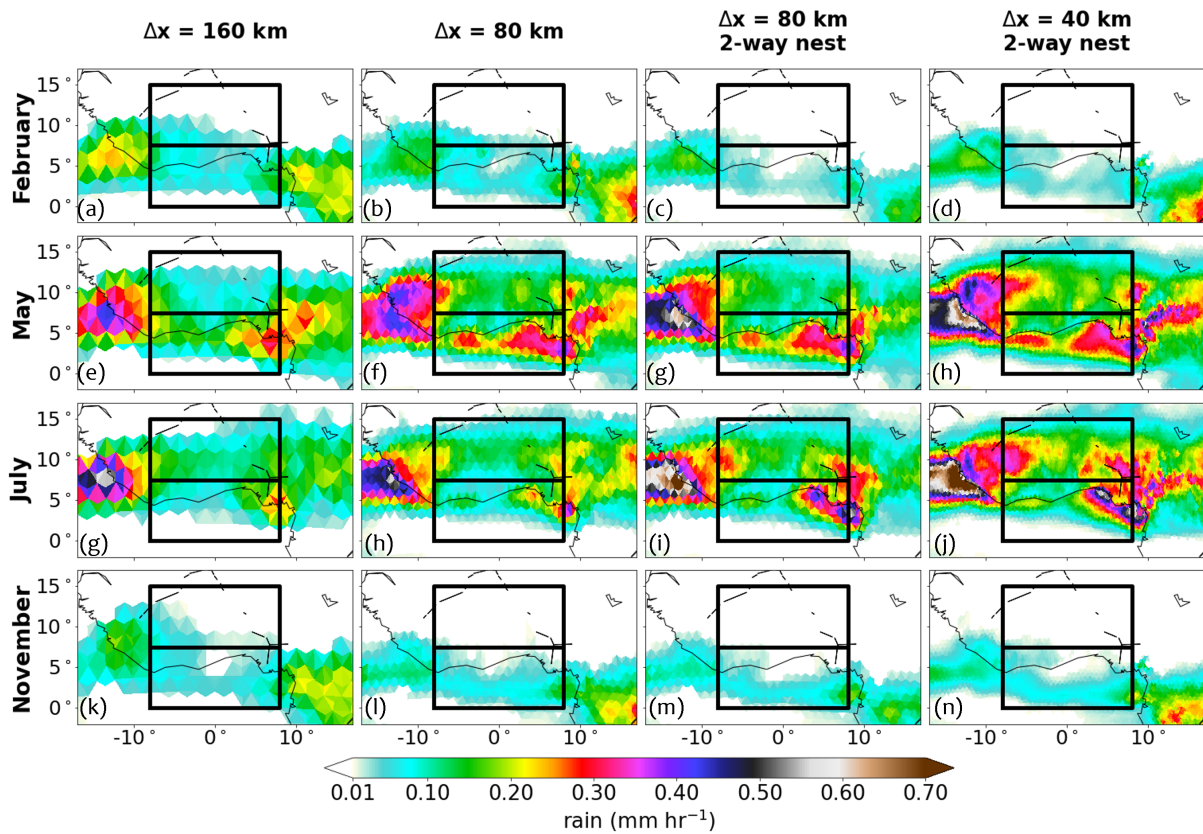


Figure 6.5. Horizontal distributions of the monthly averaged precipitation for the periods chosen in Fig 6.3. Data are shown for the ICON-ART_{iso} simulations IISO1 (first row), IISO2 (second row), IISO3 (third row) and IISO4 (fourth row), see Table 3.2. Analogous to Fig. 6.2, the black boxes indicate the chosen target regions over the Guinea Coast and the Sahel.

Overall, all four simulations are more or less able to reproduce the general contrast of mid-tropospheric $\{H_2O, \delta D\}$ pair behaviour over the Guinea Coast and the Sahel as observed by IASI. In terms of absolute values, ICON-ART_{iso} appears to develop a bias towards higher H_2O and δD values. A potential reason for this bias might be a too extensive contribution from surface evaporation, as evaporation generates more enriched signals and thereby shifts mixing curves to higher δD (Noone, 2012). At the current stage, ICON-ART_{iso} assumes monthly climatologies for the isotopic composition of evaporated water, which might be improved in the future by considering an interactive computation of fractionation during evapotranspiration.

The Guinea Coast exhibits persistently moist and enriched signals throughout the whole year, even though the moist bias shifts the contours to higher H_2O and δD when compared with the IASI data in Fig. 6.1 (note that H_2O is shown on a logarithmic scale, such that the contour size automatically shrinks for a shift to higher values). The increase of the horizontal resolution leads to a better representation of the relatively weak depletion in δD during the transition from

boreal winter to summer (compare blue and red contours in the left row of 6.3). For the global 80 km resolution, the depleting effect is even slightly larger when a two-way-interaction with the refined nest is assumed (Fig. 6.3c and e).

This goes along with corresponding improvements in the annual precipitation cycle. For instance, IISO1 (horizontal resolution of 160 km) exhibits only a less pronounced intra-annual variability in precipitation (Fig. 6.4a and first column in 6.5), while IISO4 (40 km) shows much stronger contrasts between the relatively dry (February, Fig. 6.5d) and rainy (May, Fig. 6.5h) seasons. Even though the two-way-nesting-functionality leads to a considerable increase of precipitation during summer also for the parent domain (compare Fig. 6.5j and k with Fig. 6.2c), the heavy coastal precipitation events during May are still underestimated in IISO3 and IISO4 (compare Fig. 6.5g and h with Fig. 6.2b). This is probably caused by a delayed representation of the precipitation maximum during spring towards the summer, while the precipitation maximum during autumn is shifted to earlier times of the year (compare Fig. 6.4e and g with Fig. 6.1c). Therefore, this leads to relatively less depleted $\{H_2O, \delta D\}$ pair signals over the Guinea Coast during May and November (Fig. 6.3e and g) compared to the respective contours from the IASI data (Fig. 6.1a).

Over the Sahel, all four simulations are able to reflect the general intra-annual $\{H_2O, \delta D\}$ pair variability, despite the bias towards higher δD values. During the dry months February and November, the refining of the horizontal grid size leads to a significant increase of variability towards lower H_2O and δD (see right column in Fig. 6.3) together with a slight reduction of precipitation (see right column in Fig. 6.4). This suggests that the model runs with higher resolution simulate stronger northeasterly Harmattan trade winds and thereby lead to a more efficient transport of dry air masses to West Africa. As discussed for the IASI data in the previous section, this induces marked mixing signals with dry values in the $\{H_2O, \delta D\}$ phase space and pushes the precipitation southwards (see first and last row in Fig. 6.5).

During summer, the consideration of finer resolutions improves the simulation of WAM precipitation significantly. Only IISO4 (nest with 40 km grid cell size) is capable of reproducing a reasonable precipitation peak during summer (Fig. 6.4h), also positively affecting the rain of the parent domain (Fig. 6.4f). The respective horizontal distributions show analogous results (third row in Fig. 6.5), though revealing that even for IISO4 the precipitation is still largely underestimated compared to GPM IMERG (Fig. 6.2c). Nevertheless, the simulated precipitation is already sufficient for generating robust $\{H_2O, \delta D\}$ signals towards the Rayleigh and Super-Rayleigh regimes for each IISO1–4 during July, if the δD bias is neglected (right column in Fig. 6.3). In particular for IISO4, a clear distinction of the median contour of July (50 % of corresponding data) against the respective February contour is achieved (Fig. 6.3h). Through the two-way-interaction this signal propagates to the parent domain (IISO3, Fig. 6.3f).

The direct comparison of the Sahel $\{H_2O, \delta D\}$ pair contours from ICON-ART_{iso} against the

MUSICA IASI contours reveals that ICON-ART_{iso} generates largely drier signals along the mixing curves throughout the whole year. Based on the discussions from Sect. 4.4, it is likely that a height offset between the MUSICA IASI product and the ICON-ART_{iso} data might exist, especially for dry conditions, as identified during the evaluation of the retrieval simulator with COSMO_{iso} data (see Sect. 4.4.2). This emphasizes the benefits of considering the MUSICA IASI retrieval simulator for the ICON-ART_{iso} data, because this would merge the vertical characteristics of the MUSICA IASI {H₂O, δ D} pair product to the model data and thereby probably reduce the enhanced variability of ICON-ART_{iso} for dry conditions. This, however, needs to be tested and evaluated in future work.

To summarize, despite the free-running model setups with relatively coarse horizontal resolutions, ICON-ART_{iso} is capable of reproducing the general evolution of MUSICA IASI {H₂O, δ D} pair distributions over the Guinea Coast and the Sahel, if the diagnosed moist bias is neglected. By increasing the horizontal resolution from 160 to 80 km and by additionally incorporating a nested, refined domain with 40 km grid cell size, the paired {H₂O, δ D} distributions develop larger variability and contrasts between dry and wet seasons. The isotopic composition reflects improvements of the monsoon precipitation as well as of dynamical features such as the dry Harmattan trade winds. For instance, the simulations with finer horizontal grids are able to reproduce the annual cycle of averaged precipitation much more realistic. This leads to a more pronounced representation of the observed anti-correlation of δ D and H₂O (or precipitation) associated with the monsoon activity, even though in terms of absolute values the monsoon precipitation is still underestimated.

6.2. Seasonal Monsoon Development

As a next step towards documenting and understanding the isotopic development within the monsoon system, the focus will be on the WAM season 2016. More specifically, the aim is to investigate the evolution of {H₂O, δ D} pair distributions along the monsoon phases, as defined in Knippertz et al. (2017), i.e. the pre-onset stage (Phase 1), the post-onset stage (Phase 2), the wet westerly regime (Phase 3) and the monsoon recovery (Phase 4) (see Sect. 2.4). The tropospheric {H₂O, δ D} pair information from the MUSICA IASI {H₂O, δ D} pair dataset will be compared to the COSMO_{iso} simulations CISO1 (14 km grid cell size, parameterized convection), CISO2 (14 km, explicit convection) and CISO3 (7 km, explicit convection). For all three simulations, the application of the MUSICA IASI retrieval simulator will adjust the modeled results according to the vertical characteristics of the MUSICA IASI pair product. The inter-comparison of the COSMO_{iso} precipitation against the half-hourly GPM IMERG precipitation data (RGPM1) will

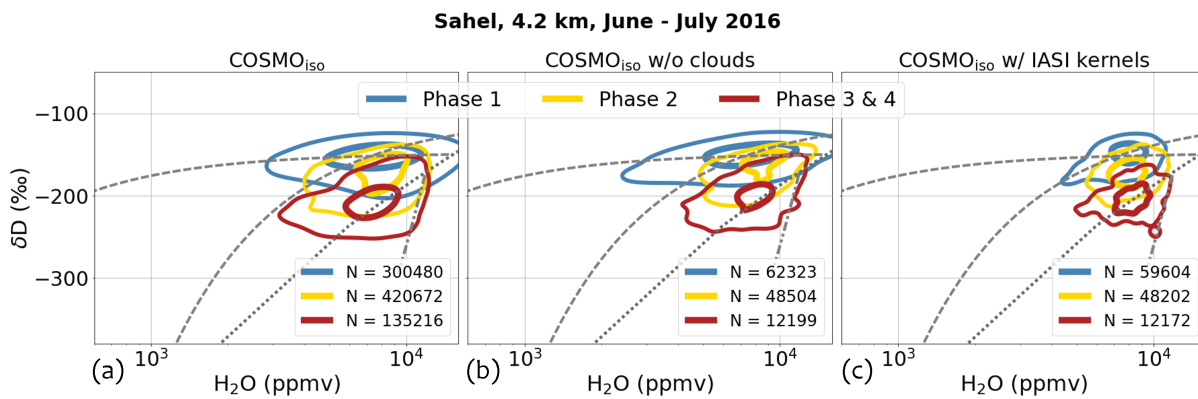


Figure 6.6. $\{H_2O, \delta D\}$ pair distributions at 4.2 km over the Sahel during the WAM season 2016 (monsoon phases defined in Sect. 2.4), as simulated by COSMO_{iso} (dataset CISO2, see Table 3.2). Panel (a) shows the raw model output, sampled according to the Metop overpass times (09.30 and 21.30 local time). For (b), the data from (a) are cloud-filtered according to Eqn. (4.32) and (4.33). Panel (c) shows the results, after simulating the MUSICA IASI averaging kernels for the cloud-filtered model data (b) and filtering out data points with averaging kernels of low quality (quality filtering according to Sect. 4.3.2).

further support the interpretation of the $\{H_2O, \delta D\}$ pair variability. For details about the individual datasets, please refer to Table 3.2.

6.2.1. Preparing the Model Data

In order to foster quantitative model-to-satellite comparisons, the MUSICA IASI retrieval simulator is considered for harmonizing the COSMO_{iso} $\{H_2O, \delta D\}$ data with the vertical sensitivity of the remotely sensed MUSICA IASI $\{H_2O, \delta D\}$ pair product. This is achieved, first, by sorting the model data according to the general availability of MUSICA IASI observations (09.30 and 21.30 local time, cloud-free scenes) and, second, by simulating and applying the MUSICA IASI $\{H_2O, \delta D\}$ pair averaging kernels for the atmospheric conditions provided by COSMO_{iso} (details about processing can be found in Sect. 4.4).

Figure 6.6 displays how the Sahelian $\{H_2O, \delta D\}$ distributions from COSMO_{iso} are affected during the different monsoon phases in 2016 when the MUSICA IASI retrieval simulator is applied. The removal of cloudy data points according to the cloud criteria from Eqn. (4.32) and (4.33) leads to a general reduction of variability mainly in H_2O (compare Fig. 6.6a and b). The data amount reduces by 80 % during Phase 1 and by 88–90 % during the other phases, as the latter are marked by strongly increased convective activity and thus exhibit more cloud occurrences. For Phase 1 and 2, the minor signals below the Rayleigh curve disappear almost completely as response to the cloud filtering. By additionally applying the retrieval simulator and precluding data points with low-quality averaging kernels (analogous to the quality filtering described in Sect. 4.3.2),

the overall $\{H_2O, \delta D\}$ variability is even more reduced. In particular, the relatively strong mixing signals before the monsoon onset (Phase 1) are damped. Phase 3 & 4 depict the only $\{H_2O, \delta D\}$ pair contours with noticeable Super-Rayleigh signals (Fig. 6.6c).

6.2.2. Analysis of Monsoon Season 2016

After adjusting the $COSMO_{iso}$ data with respect to the vertical MUSICA IASI pair data characteristics, foundations are laid for a meaningful direct model-to-satellite comparison for the WAM season 2016. Figure 6.7 shows the $\{H_2O, \delta D\}$ pair distributions over the Guinea Coast and the Sahel, as observed by IASI and modeled in the $COSMO_{iso}$ simulations CISO1–3. A zoomed view on the respective $\{H_2O, \delta D\}$ data during the post-onset stage (Phase 2) is given in Fig. 6.8. The daily averaged precipitation from each dataset is shown as mean over the chosen domains in Fig. 6.9 and as horizontal distribution in Fig. 6.10.

Over the Guinea Coast, the overall evolution of the MUSICA IASI $\{H_2O, \delta D\}$ pairs is well reproduced by $COSMO_{iso}$. Marked and persistent mixing signals are apparent for all datasets throughout the different monsoon phases (left column in Fig. 6.7), as also observed on the annual scale (Fig. 6.1a). $COSMO_{iso}$ captures the overall decrease of daily precipitation during the transition from the pre- to the post-onset stage, even though it misrepresents individual coastal precipitation events (left column in Fig. 6.9). For instance, CISO1 partly overestimates the pre-onset precipitation, in particular over the eastern Tropical Atlantic, while the precipitation in CISO2 and CISO3 appears to be rather fragmented and (in terms of total values) underestimated (first row in Fig. 6.10). During Phases 2 and 3, the coastal precipitation is overall lower in $COSMO_{iso}$ compared to GPM IMERG. Especially in Phase 3, the remarkable increase of precipitation over the eastern Tropical Atlantic, as seen in the RGPM1 data, is not covered in any of the $COSMO_{iso}$ simulations (third row in Fig. 6.10). From the isotopic perspective, this precipitation feature results in a clear $\{H_2O, \delta D\}$ signal of the MUSICA IASI data below the Rayleigh curve, which is missed out by $COSMO_{iso}$ (compare red contours in the left column of Fig. 6.7).

The Sahelian $\{H_2O, \delta D\}$ pair distributions basically reflect the observed transition from the mixing to the Rayleigh and Super-Rayleigh regimes as response to the increase of convective activity (compare Fig. 6.1b with right column of Fig. 6.7). Phase 1 is governed by air mass mixing with occasional convective events generating minor Rayleigh features (Fig. 6.7b). The mixing signals are slightly overestimated in CISO1, but are improved when convection is considered explicitly, whereas the increase of the horizontal resolution does not lead to a significant impact in the $\{H_2O, \delta D\}$ data of Phase 1.

The increase of precipitation from Phase 1 to Phase 2 is more pronounced in CISO1 than in

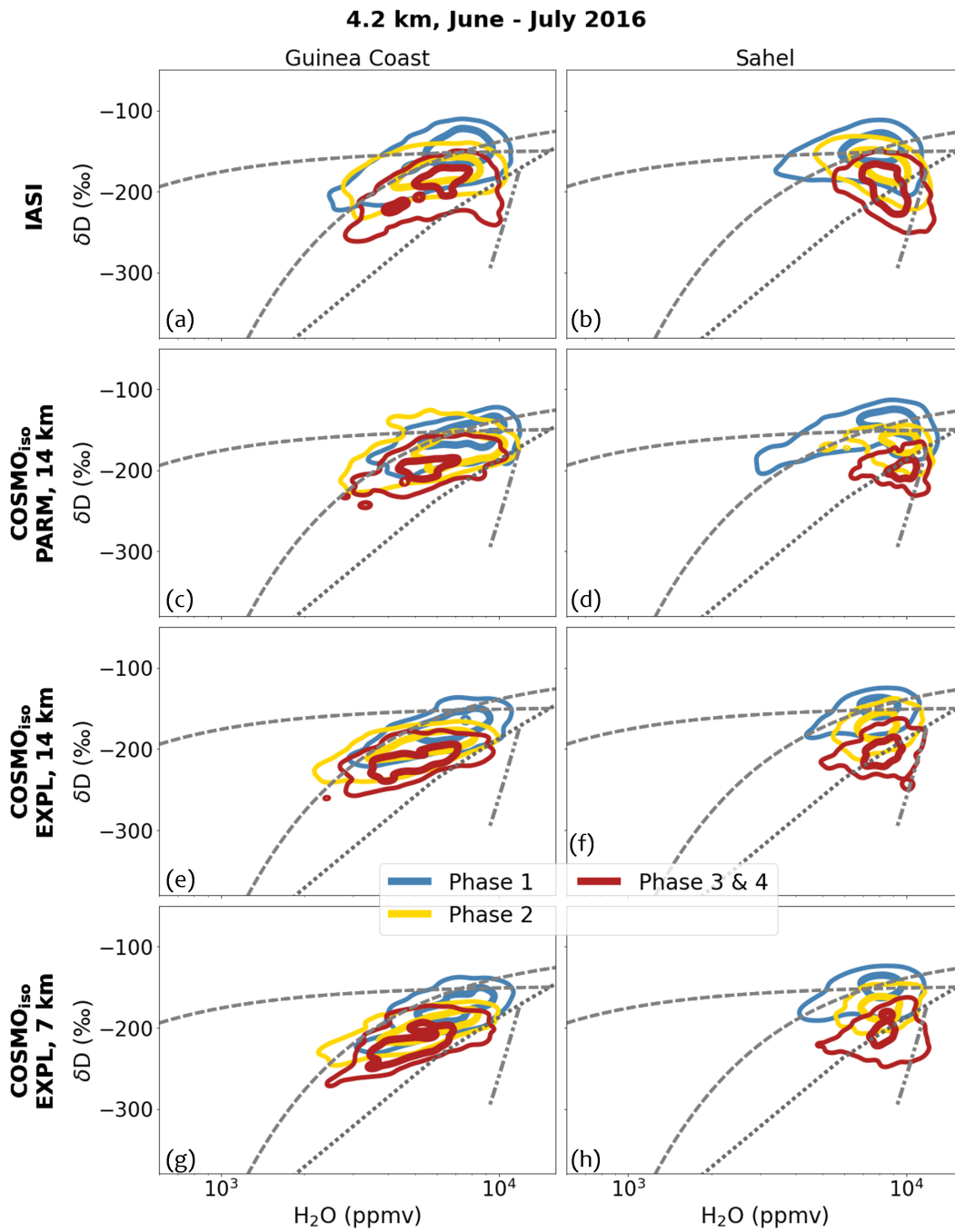


Figure 6.7. $\{H_2O, \delta D\}$ pair distributions at 4.2 km over the Guinea Coast (left column) and the Sahel (right column) during the WAM season 2016 (see Sect. 2.4). Data are shown from (a,b) the MUSICA IASI $\{H_2O, \delta D\}$ pair dataset and the COSMO_{iso} simulations (c,d) CISO1, (e,f) CISO2 and (g,h) CISO3 (see Table 3.2). All COSMO_{iso} data are post-processed with the MUSICA IASI retrieval simulator.

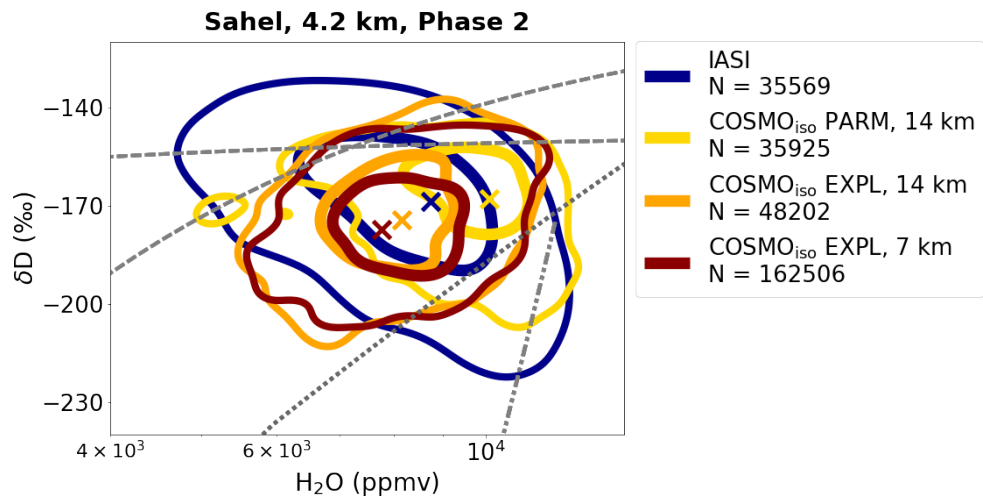


Figure 6.8. Zoomed view on the $\{H_2O, \delta D\}$ pair distributions over the Sahel during the post-onset stage (Phase 2) of the WAM season 2016. The MUSICA IASI data are shown by the blue contours, the CISO1 results in yellow, CISO2 in orange and CISO3 in dark red. The weighted means of the underlying two-dimensional histogram contours are indicated by the colored x-markers.

CISO2 and CISO3, as the latter two appear to underestimate the averaged monsoon precipitation (see right column in Fig. 6.9 and second row in 6.10). Additionally, despite showing a moist bias towards higher H_2O , the $\{H_2O, \delta D\}$ pair distributions of CISO1 seem to better reproduce the Super-Rayleigh features from the MUSICA IASI data, while CISO2 and CISO3 shift the $\{H_2O, \delta D\}$ contours towards lower H_2O and δD values along the Rayleigh curve (see x-markers in Fig. 6.8). By relying on these rather statistical views, one could assume that CISO1 with the parameterized convection may be better for simulating the general monsoon precipitation as well as its effects on the isotopic composition of the mid-tropospheric water vapor. As this would be contrary to previous studies (see discussions in Sect. 3.4), a more detailed view on these features might be useful.

A Hovmoeller diagram of the hourly and meridionally averaged precipitation over the Sahel domain from RGPM1 and CISO1–3 is shown in Fig. 6.11. The remotely sensed precipitation shows distinct and strongly organized mesoscale convective systems (MCS) with a fast westward propagation (from 8° to -8° E within slightly more than one day, see Fig. 6.11). The comparison to the model data shows that these convective events are quite poorly represented in terms of sharpness and intensity, when considering the parameterized convection (CISO1, Fig. 6.11b). In addition, CISO1 appears to develop a longitudinal band of relatively moderate precipitation that stretches over the whole domain and is daily recurring for noon and afternoon times. This is contrary to the observations within RGPM1, where similar precipitation cycles might appear only on a very occasionally and regionally limited basis, e.g. around 0° E (Fig. 6.11a). This means

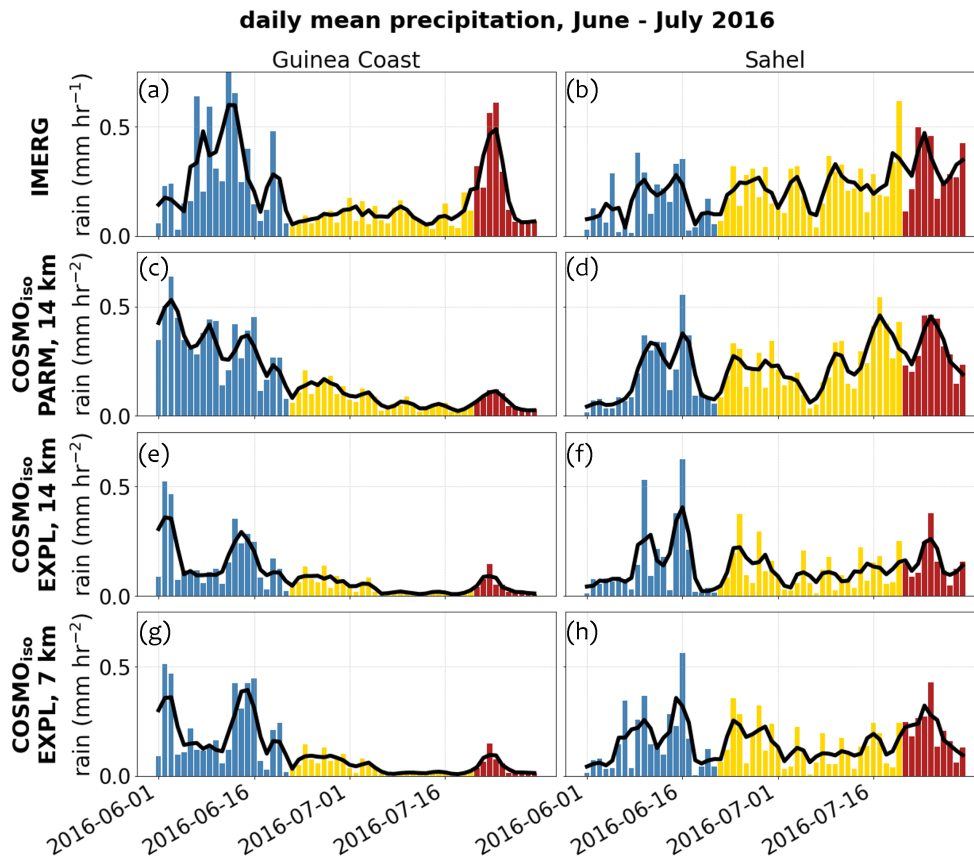


Figure 6.9. Daily averaged total precipitation over the Guinea Coast (left column) and the Sahel (right column) during the WAM season 2016, derived from the half-hourly GPM IMERG precipitation data (RGPM1, a,b) and from the hourly precipitation provided by the full COSMO_{iso} simulations (c,d) CISO1, (e,f) CISO2 and (g,h) CISO3. The colors correspond to the monsoon phases, as shown in Fig. 6.7.

that, by averaging, CISO1 logically shows increased precipitating values, but for the wrong reasons.

However, when switching off the parameterized convection schemes, the overall simulated precipitation patterns improve significantly (Fig. 6.11c and d). On the one hand, the MSCs are much more distinct and intense, perhaps even too much, as herein they reach higher precipitation values than in RGPM1. A similar over-estimation of convective precipitation as response to the explicit convection configuration is documented in Marsham et al. (2013), Kniffka et al. (2019) and Pante and Knippertz (2019). On the other hand, the erroneous longitudinal precipitation bands vanish completely, leading to lower daily mean values of the total precipitation. Further, the increase of the horizontal resolution achieves the generation of more local precipitation events between the MSCs passages, with ultimately showing the best performance compared to RGPM1, even though some features might still be underestimated.

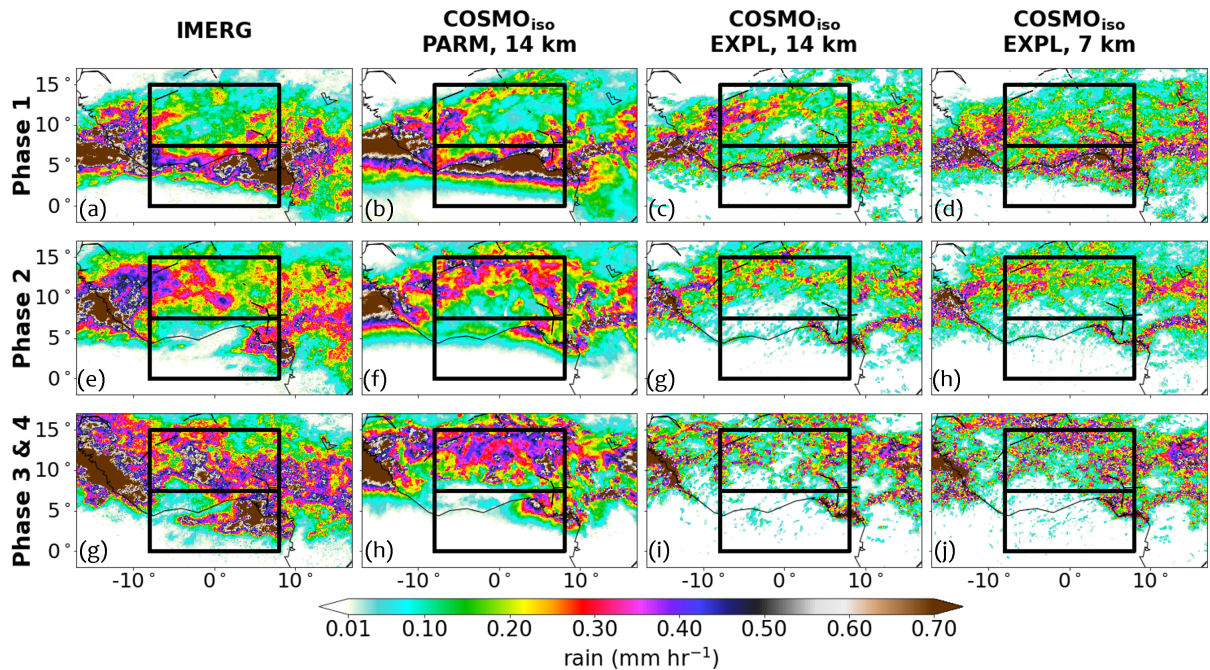


Figure 6.10. Horizontal distributions of daily averaged precipitation over West Africa for the different WAM phases, derived from the half-hourly GPM IMERG precipitation data (RGPM1, first column) and from the hourly precipitation provided by the COSMO_{iso} simulations (c,d) CISO1, (e,f) CISO2 and (g,h) CISO3. Analogous to Fig. 6.2, the black boxes frame the target regions over the Guinea Coast and the Sahel.

With keeping in mind these findings about the strongly varying precipitation characteristics for parameterized and explicit convection, this allows for a more sophisticated interpretation of the $\{H_2O, \delta D\}$ pair distributions during the monsoon activity (Fig. 6.8). CISO1 tends to generate largely more moderate precipitation, what explains the moist bias as well as the more pronounced Super-Rayleigh features. As Dansgaard (1964) stated, weaker precipitation events lead to a more efficient rain evaporation that reflects in stronger fractionation between the ambient vapor and the rain drop, if the evaporated fraction of the rain drop is small enough (see further discussions in Sect. 6.3; Risi et al., 2010a, Noone, 2012). That is, the less intense but more large-scale precipitation patterns within CISO1 foster the evaporation of falling rain, what increases H_2O (causing the moist bias) and depletes δD (creating the Super-Rayleigh signals) in the vapor. Vice versa, the runs with explicit convection (CISO2 and CISO3) generate more focused and intense convection with heavier precipitation and thereby exhibit less Super-Rayleigh depletion. Instead they generate more Rayleigh condensation due to the stronger vertical updrafts, leading to the shift of $\{H_2O, \delta D\}$ contours downwards the Rayleigh curve during the post-onset stage (Fig. 6.8). During Phase 3 & 4, the observed precipitation increase in RGPM1 is captured by all COSMO_{iso} simulations (right column in Fig. 6.9), with showing similar model performance like Phase 2

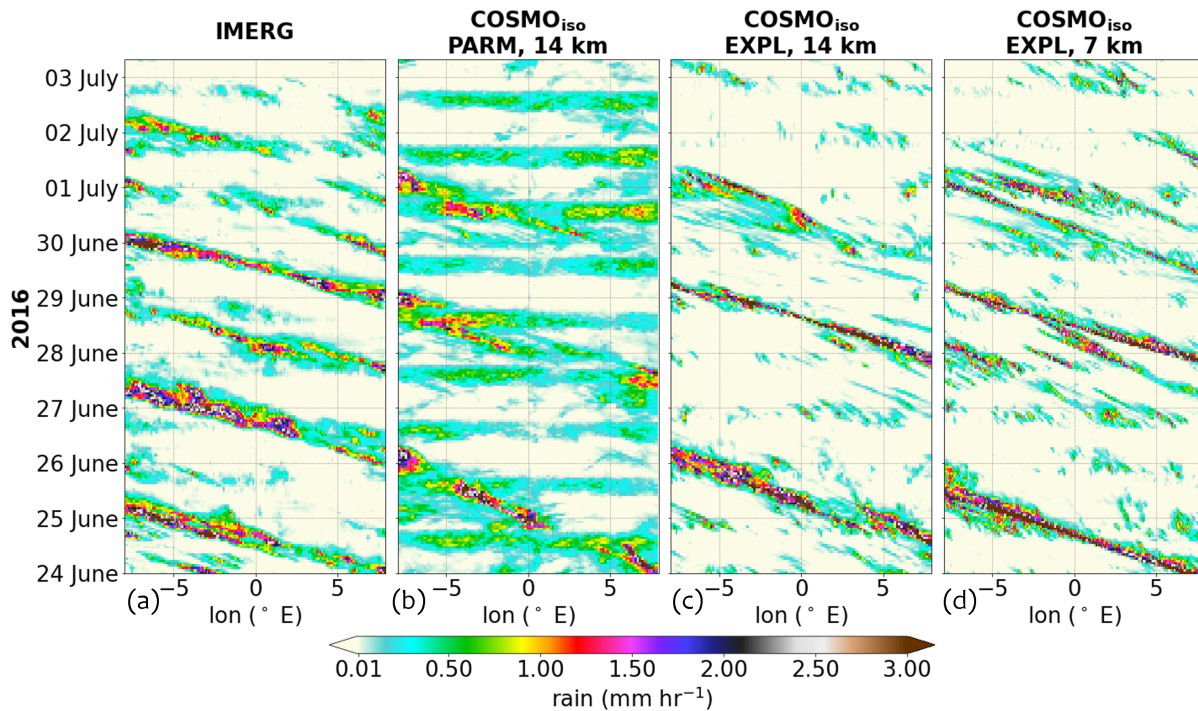


Figure 6.11. Hovmoeller diagram for the hourly and meridionally averaged precipitation over the Sahelian domain ($7.5^\circ - 15^\circ$ N), showing the results (a) for GPM IMERG (RGPM1), and for the COSMO_{iso} simulations (b) CISO1, (c) CISO2 and (d) CISO3 (see Table 3.2).

(overestimated precipitation averages in CISO1, more fragmented and underestimated in CISO2 and CISO3). However, the ongoing convective activity over the Sahel further enhances the microphysical cloud and rain effects in the $\{H_2O, \delta D\}$ phase space. Now all runs show marked Rayleigh and Super-Rayleigh features, with the former being again more pronounced for the explicit convection setups (see aforementioned discussion).

In summary, this section highlights the evolution of mid-tropospheric $\{H_2O, \delta D\}$ pairs during the WAM period 2016, as observed in the MUSICA IASI data and in different COSMO_{iso} simulations, which are post-processed with the MUSICA IASI retrieval simulator. While the Guinea Coast shows consistently and constantly marked mixing signals throughout the whole monsoon period, strong Rayleigh and Super-Rayleigh signals develop over the Sahel after the monsoon onset as response to the increased precipitation. The change from parameterized to explicit convection in COSMO_{iso} improves the representation of the monsoon precipitation in terms of daily cycles and amount. However, the isotopic perspective reveals that the explicit convection also tends to overestimate the convective precipitation associated with the development of MCSs. In contrast, the effect of refining the horizontal resolution from 14 to 7 km is in this analysis rather negligible.

6.3. Convective-Scale Variability

After investigating the isotopic composition of mid-tropospheric water vapor associated with the WAM activity in a quasi-climatological, intra-annual context (Sect. 6.1) as well as in a more detailed seasonal perspective (Sect. 6.2), this section aims for an analysis of $\{H_2O, \delta D\}$ pairs on the convective scale. It develops and evaluates a method for interpreting the remotely sensed MUSICA IASI $\{H_2O, \delta D\}$ pair data with respect to individual convective systems. The idea is to identify those specific observations that are associated with monsoon precipitation and then to contrast the corresponding $\{H_2O, \delta D\}$ signals to those from observations without any precipitation occurrences.

However, as the MUSICA IASI retrieval is currently only applicable for cloud-free scenes, this product per se cannot provide information about occurring precipitation events and their impacts on the isotopic composition of water vapor. Instead, the herein applied approach is to use the GPM IMERG precipitation dataset and to identify air masses with and without significant precipitation impact. Based on this, the MUSICA IASI observations can then be sorted into corresponding clusters and analyzed with respect to the idealized process curves from Fig. 3.2.

6.3.1. Clustering Approach

The basis for the rain clustering is the half-hourly GPM IMERG precipitation dataset (RGPM1, see Table 3.2) that is used to compute daily averaged precipitation distributions on the native $0.1^\circ \times 0.1^\circ$ grid. As the focus is on air masses that have been affected by precipitation events rather than on the individual precipitation events themselves, the daily RGPM1 data are averaged on a horizontal $1^\circ \times 1^\circ$ grid in order to perform daily statistics for the rain occurrences. A grid box will be treated as non-rain event, if the rain inside the box accumulated during the previous three days remains below 0.2 mm hr^{-1} . However, this is a simplified assumption, as it does not take into account the advection of air masses that might have experienced precipitation within adjacent grid boxes (e.g. compare with MCS propagation as proxy for wind speed in Fig. 6.11). In the future, this issue may be addressed, for instance, by also considering the Lagrangian perspective for ensuring rain-free conditions.

Further, to identify air masses that were actually affected by precipitation, a grid box is classified as post-rain event, if the precipitation on the previous day exceeds the threshold of 3 mm hr^{-1} . This definition is chosen such that it allows for identifying events, where IASI observed cloud-free air masses that have been affected by precipitation during their past. Again, this definition might be distorted due to advection. However, the threshold for the post-rain events is chosen rather high (compare with GPM IMERG precipitation distribution in Fig. 6.11a) such that mainly the well-organized and intense convective systems with strongest impact on the isotopic composition

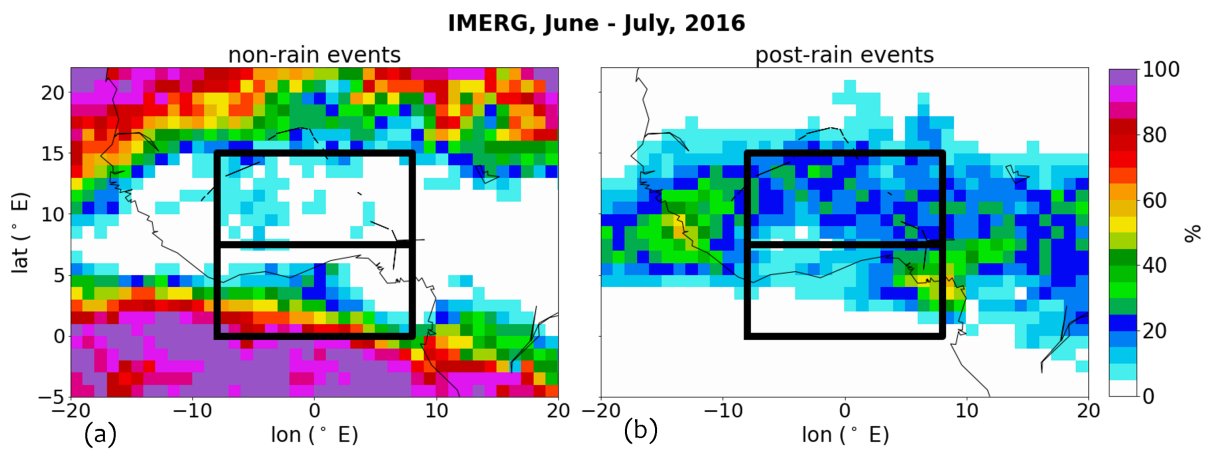


Figure 6.12. Fraction of days with (a) non-rain and (b) post-rain events for the WAM months June – July, derived by considering the daily averaged precipitation from the half-hourly GPM IMERG precipitation dataset (RGPM1) and evaluated on a $1^\circ \times 1^\circ$ grid. The black-framed boxes depict the chosen target domains over the Guinea Coast and the Sahel.

of mid-tropospheric water vapor are captured. In the final step, each MUSICA IASI $\{H_2O, \delta D\}$ pair observation will be checked depending on its geographical location, whether one of these two criteria are fulfilled, with ultimately resulting in two distinct MUSICA IASI data clusters, i.e. for non-rain and post-rain events.

6.3.2. Clustering of Monsoon Season 2016

The aforementioned rain clustering into non-rain and post-rain events is applied based on the half-hourly GPM IMERG precipitation data (RGPM1) during the period June – July 2016. Figure 6.12 provides a horizontal overview about the relative fraction of days that are identified as non-rain or post-rain events, for each considered $1^\circ \times 1^\circ$ grid box over West Africa. Following the main precipitation distribution during summer, as shown in Fig. 6.2c, this clustering successfully captures the areas associated with strong precipitation. It mainly reflects the WAM activity, i.e. the predominance of precipitation over the Sahel, while the Guinea Coast is mostly rain-free with regional exceptions for its easternmost areas. Here, around the coastal areas of Nigeria and Cameroon, a localized precipitation peak develops (e.g. discussed in Nlend et al., 2020) and appears to be even more frequent and persistent than the convective precipitation events over the Sahel.

Figure 6.13 shows the MUSICA IASI $\{H_2O, \delta D\}$ distributions that result when sorting the data points into non-rain and post-rain clusters, respectively. For the Guinea Coast, mostly non-rain events appear. As can be seen in the backward trajectories from Sect. 5.4.1 in Fig. 5.3, the moist southerly monsoon flow travels over the Gulf of Guinea towards the coastal West Africa, but

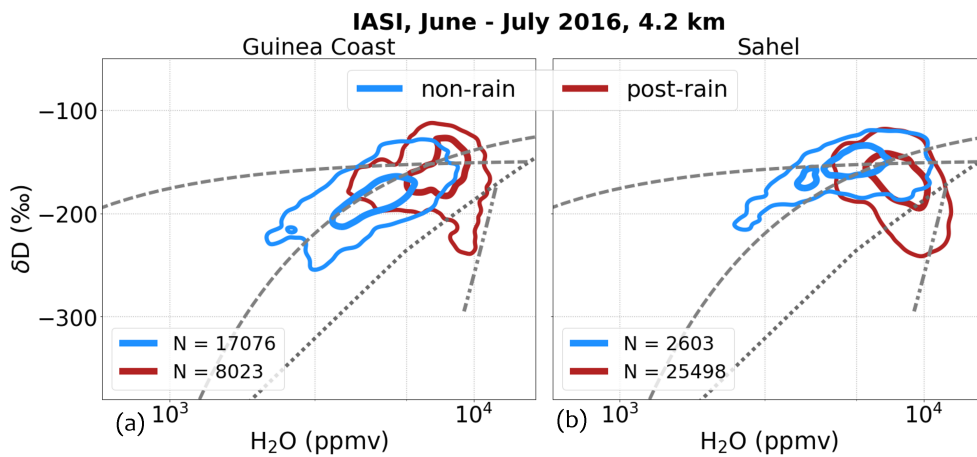


Figure 6.13. MUSICA IASI $\{H_2O, \delta D\}$ pair data over the Guinea Coast and the Sahel during the WAM months June – July 2016, clustered into non-rain and post-rain events based on daily averaged precipitation from GPM IMERG (RGPM1).

without generating intense convection (see Fig. 6.12). This reflects in marked mixing signals for the non-rain cluster. However, also a considerable amount of post-rain events develops within the Guinea Coast domain. Most likely these are linked to the above discussed precipitation hotspot over the eastern Gulf of Guinea. Interestingly, despite the strong precipitation contribution, these data show mainly an increased moistening along the mixing line with only minor Rayleigh and Super-Rayleigh features. This might be explained by taking into account the observation of Dansgaard (1964) that rain evaporation is reduced for more intense precipitation events (analogous to discussion in Sect. 6.2.2). As large rain drops fall faster, there is less time for evaporation, leading to less depletion of the ambient vapor (Stewart, 1975). Additionally, this precipitation hotspot appears to be sufficiently large-scale and persistent such that it leads to a significant moistening, which in turn fosters saturation and further reduces the rain evaporation efficiency (Risi et al., 2008a, Lee and Fung, 2008, Yoshimura et al., 2010). This is in agreement with the post-rain data being generally moister than the non-rain events with only minor Rayleigh and Super-Rayleigh features (Fig. 6.13a). As discussed in Sect. 3.1.4 and 5.4.4, in saturated conditions the fractionating effect of isotopic equilibration between the rain drop and the ambient vapor achieves a maximum, which could account for the minor Super-Rayleigh features seen in the $\{H_2O, \delta D\}$ data of the post-rain cluster over the Guinea Coast.

Over the Sahel domain very few non-rain events appear, whereas the post-rain cluster is predominant (Fig. 6.13b). The few non-rain data show marked mixing features that are slightly moister than the corresponding data over the Guinea Coast, as the evapotranspiration from wet surfaces in West Africa is found to further moisten the Sahelian troposphere (Nieto et al., 2006). In contrast to the Guinea Coast, the Sahelian post-rain events exhibit strong Rayleigh and

Super-Rayleigh features. As discussed in Sect. 2.3, the precipitation associated with squall line events consists of strong convective precipitation in the leading edge and moderate precipitation below the trailing stratiform cloud shield. The latter is characterized by smaller rain drops and the contribution of the dry rear-to-front inflow (e.g. the AEJ), with both features leading to a more efficient rain evaporation. Together these effects generate strong fractionation in the water vapor and thereby account for the observed Super-Rayleigh features in the post-rain cluster over the Sahel.

To conclude, despite the relatively simple clustering approach, this is an efficient method for identifying air masses that were affected by convective precipitation events against those without any significant precipitation contribution. By investigating the corresponding MUSICA IASI $\{H_2O, \delta D\}$ pair data, the non-rain events are clearly associated with air mass mixing. In contrast, the post-rain events are in general moister than the non-rain events and show distinct Super-Rayleigh signals. The latter is particularly pronounced over the Sahel, where squall line events together with the dry rear-inflow jet lead to intense rain evaporation.

6.4. Summary

In this section, the isotopic composition of mid-tropospheric water vapor over the Sahel is documented and analyzed with regards to the underlying control mechanisms. The focus is on dynamical and microphysical processes associated with the WAM, which, for the first time to date, is being studied from the $\{H_2O, \delta D\}$ perspective. This includes the cross-comparison between data from remote sensing as well as from isotope-enabled models.

Clear differences between the Guinea Coast and the Sahel emerge on the intra-annual scale. The MUSICA IASI $\{H_2O, \delta D\}$ pairs together with the GPM IMERG precipitation data reflect that the coastal areas are continuously moist with distinct mixing signals, whereas the Sahel shows a markedly strong variability from mostly dry and rain-free winter months to a wet rain season during boreal summer. While during winter marked mixing signals due to the strong influence of dry northeasterly trade winds develop, the monsoon months exhibit a strong anti-correlation between the increasing H_2O and decreasing δD , as already observed in Sect. 4.3.3.6 in Fig. 4.16 and in Sect. 5.2 in Fig. 5.1. As this behaviour is closely linked to the Super-Rayleigh regime, it can be attributed to microphysical rain processes (evaporation and equilibration), which are strongly increased during the monsoonal convective activity. ICON-ART_{iso} is able to reproduce the general $\{H_2O, \delta D\}$ variability over the Guinea Coast and the Sahel, despite a bias towards higher H_2O and δD values. The general performance of ICON-ART_{iso} in capturing the general WAM features can be improved by considering finer horizontal grids (evaluated from 160 to 40 km). The incorporation of a refined, nested domain over West Africa with a two-way-interaction achieves more realistic results within the global, parent domain. However, a more meaningful

one-to-one comparison between ICON-ART_{iso} and MUSICA IASI requires further developments for making the MUSICA IASI retrieval simulator applicable also to ICON-ART_{iso} data.

The in-detailed analysis of the $\{H_2O, \delta D\}$ variability during the WAM season 2016 is based on isotopic data from MUSICA IASI and various COSMO_{iso} simulations, which are post-processed with the MUSICA IASI retrieval simulator. The enhanced depletion over the Sahel during summer, as discussed above, is clearly linked to the onset of monsoon convection. As the horizontal resolutions considered for COSMO_{iso} are already rather high, the increase of 14 to 7 km leads only to minor improvements in precipitation and $\{H_2O, \delta D\}$ distributions, whereas the switch from parameterized to explicit convection has an enormous impact on the general WAM simulation. While the parameterized setup can hardly represent the daily cycles and maximum sensitivity of the monsoon precipitation reasonably, the explicit treatment of convection is able to simulate distinct and organized convective systems comparable to the GPM IMERG precipitation data. However, the $\{H_2O, \delta D\}$ distributions reveal that the explicit convection setup might generate too strong vertical convective updrafts leading to stronger Rayleigh condensation as the MUSICA IASI data would suggest.

Finally, the convective-scale clustering of MUSICA IASI $\{H_2O, \delta D\}$ pair data according to convective precipitation events, as provided by GPM IMERG, allows for a distinction between non-rain and post-rain events. In the absence of rain, the former shows marked mixing signals. The post-rain cluster is associated with moistening along the mixing line for large-scale and persistent precipitation over the eastern Gulf of Guinea, whereas the more separated precipitation events over the Sahel (e.g. squall lines) together with the contribution of the dry AEJ go along with strong Super-Rayleigh signals, indicating a more efficient evaporation of falling rain. That is, the paired $\{H_2O, \delta D\}$ phase space is capable of reflecting the differing impact of rain on tropospheric water vapor depending on the mechanisms and conditions under which the precipitation event developed and occurred.

Since the miss-representation of meteorological WAM features appears to be a common issue for most atmospheric models (Bielli et al., 2010, Meynadier et al., 2010), the paired analysis of stable water isotopes constitutes a powerful tool for evaluating effects of dynamical and microphysical processes and the model performances in reproducing them.

Part IV.

Conclusions

7 | Conclusions

The prime goal of this thesis was to advance the use of isotopic information of tropospheric water vapor for investigating the atmospheric water cycle. A special focus was on paired distributions of H_2O and δD over the Sahel zone during the West African Monsoon (WAM). This region is characterized by complex interactions between large-scale dynamics and microphysical cloud and rain processes, ultimately leading to a highly variable and still poorly understood tropospheric moisture budget.

7.1. Research Highlights

Figure 7.1 provides a summary of the main research results achieved by the three-step concept, which was the overall guideline of this thesis. As a first step, the lack of observational data over West Africa was addressed by developing and providing a novel dataset of tropospheric $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair data based on radiance measurements of the satellite-based sensor IASI. This dataset is an a posteriori processed extension of the MUSICA IASI retrieval product as presented in Schneider et al. (2021c). From the independently retrieved H_2O and δD states for cloud-free scenes, an improved optimal estimation product for $\{\text{H}_2\text{O}, \delta\text{D}\}$ pairs is derived with best sensitivity in the mid-troposphere. In total, the complete dataset comprises two global maps of $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair data per day for the whole period from October 2014 to June 2019. Due to its unique combination of coverage and resolution in space and time, this dataset is highly promising for studying atmospheric moisture pathways. In addition, a newly developed MUSICA IASI retrieval simulator, which considers the same radiative transfer algorithm as the actual MUSICA IASI retrieval, proves as valuable tool for achieving meaningful model-to-satellite comparisons with the MUSICA IASI $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair product. In its current state, this retrieval simulator is only applicable to data from the isotope-enabled model $\text{COSMO}_{\text{iso}}$, but should be extended in the future.

The second step develops a model-based framework for allowing improved interpretations of paired $\{\text{H}_2\text{O}, \delta\text{D}\}$ distributions. This consists of a novel Lagrangian process attribution procedure, where the isotopic evolution of trajectories is investigated and attributed to underlying moisture

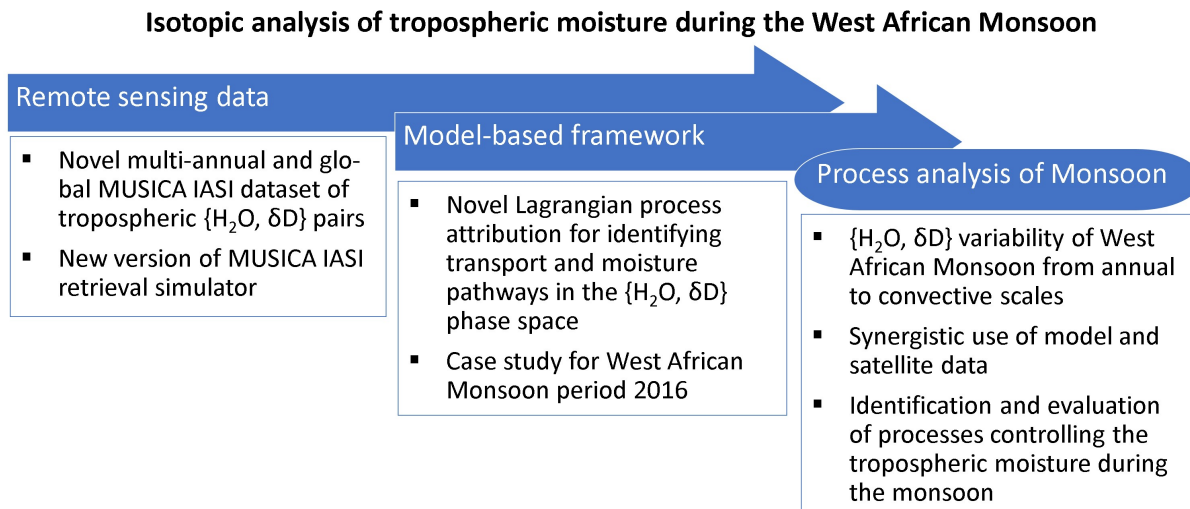


Figure 7.1. Illustrative scheme of the three-step concept followed throughout the thesis, as shown in Fig. 1.1, but extended with the main research highlights for each analysis step.

processes based on idealized process curves from Noone (2012). Following the main focus of this thesis, this procedure was applied to backward trajectories for the Sahelian mid-troposphere during the WAM season 2016, calculated with the trajectory tool LAGRANTO and based on data from COSMO_{iso}. As a result, this procedure identified occurrences of air mass mixing and microphysical rain processes along the dominant transport pathways of the WAM, with each leading to characteristic $\{H_2O, \delta D\}$ signals. The combination of these processes is thereby capable of determining the final isotopic composition in the Sahelian troposphere during the monsoon. In particular, the enhanced depletion in δD for high H_2O contents, an often discussed issue in literature (e.g. Worden et al., 2007, Risi et al., 2010a, Tremoy et al., 2012, Lacour et al., 2018), could be attributed to mid-tropospheric rain processes such as evaporation and equilibration.

Finally, the third step aimed at an extensive and multi-scale analysis of the isotopic composition of the WAM and its control factors. It achieved a synthesis of the different data products, including $\{H_2O, \delta D\}$ data from MUSICA IASI as well as from the isotope-enabled models ICON-ART_{iso} and COSMO_{iso}, with the latter being post-processed with the new MUSICA IASI retrieval simulator. Additionally, this analysis considered half-hourly and monthly precipitation data from GPM IMERG, a unified satellite precipitation product provided by NASA. This variety of datasets allowed to investigate the isotopic composition of the WAM system from the annual, seasonal and convective perspectives. A common finding is the clear distinction of $\{H_2O, \delta D\}$ variability between the Guinea Coast and the Sahel and the attribution to mixing and rain processes. In particular, the Sahel is characterized by a tremendous variability from mostly dry and rain-free winter months, where air mass mixing prevails the $\{H_2O, \delta D\}$ signals in the troposphere, to the wet rainy season during the monsoonal summer period, which exhibits the aforementioned

enhanced depletion in δD for high H_2O . A clustering of the MUSICA IASI $\{H_2O, \delta D\}$ pairs based on the GPM IMERG precipitation data further confirmed the attribution of this anti-correlation between H_2O and δD to effects of rain processes associated with the increased convective activity (as discussed in the previous paragraph). Furthermore, the identified links between characteristic water isotope signals and the dynamical and microphysical features of the WAM system represent a useful tool for performing model evaluations. For instance, this study showed that the simulation of the general WAM development (e.g. spatio-temporal evolution of dynamical features and monsoon precipitation on annual scales) can be improved by increasing the horizontal resolution (here from 160 to 40 km). This effect is reduced for convection-resolving grid sizes (7–14 km), but at such scales, turning off the convection parameterization and thus explicitly resolving convective processes appears to strongly improve the representation of individual convective systems. While the convection parameterization of COSMO_{iso} excessively generates moderate precipitation over the Sahel, the explicit setup leads to much more realistic daily cycles of the monsoon precipitation, albeit tending to overestimate the maximum rain amount of individual convective systems.

As an overall conclusion, this thesis demonstrates the potential of paired distributions of H_2O and δD for investigating the tropospheric moisture in general and, for the first time so far, with focus on the complex monsoon system over West Africa, which could in principle be extended to any other region. By providing a novel comprehensive dataset of observational water isotope information and new model-based analysis frameworks, this thesis has already made essential contributions to a variety of related studies, but also lays the foundation for developing promising follow-up projects (as outlined in Sect. 7.2). The core belief is that the synergistic analysis of isotopic data from observations and atmospheric models, together with sophisticated process attribution approaches, is of great value for improving the general understanding of the hydrological cycle and its representation in weather and climate models.

7.2. Perspectives

To further illustrate the potential of the results obtained within this thesis, some examples of ongoing and planned activities that incorporate them are presented here.

For instance, Dahinden et al. (2021) present an analysis of moisture transport pathways for the North Atlantic with special focus on the island of Tenerife. This is a multi-platform analysis, including the new MUSICA IASI $\{H_2O, \delta D\}$ pair data, which are compared with data from COSMO_{iso} simulations, ground-based remote sensing platforms as well as with airborne in-situ measurements. For the comparison with the MUSICA IASI data, the new MUSICA IASI retrieval simulator is applied to the simulations of COSMO_{iso}.

An extensive and multi-platform analysis of the isotopic composition associated with boundary layer processes and their relation to macro-scale cloud properties is provided by the EUREC4A-iso campaign (<http://eurec4a.eu/overview/eurec4a-iso>, Elucidating the Role of Clouds-Circulation Coupling in Climate) with support of the World Climate Research Programme (WCRP). Among a variety of observational and modeling datasets with a focus on the atmospheric conditions around the Barbados Cloud Observatory, the same MUSICA IASI $\{H_2O, \delta D\}$ pair product will be included.

Further, by using the LAGRANTO trajectory ensemble from this work, Ertl et al. (2021) presents a first, data-driven approach for interpreting the MUSICA IASI $\{H_2O, \delta D\}$ pair data. They develop a semi-supervised approach that automatically clusters trajectories with similar geographical and isotopic properties. The aim is to quantitatively identify characteristic $\{H_2O, \delta D\}$ structures that correspond to underlying moisture processes.

An approach of incorporating the observational MUSICA IASI $\{H_2O, \delta D\}$ pair data into the data assimilation scheme of isotope-enabled models is presented in Toride et al. (2021) and in Khosrawi et al. (2021). In these studies, the authors were able to provide evidence that the assimilation of synthesized MUSICA IASI observations is indeed capable of improving the quality of meteorological analysis fields within numerical weather prediction. Current research activities examine the impact of considering the actual MUSICA IASI $\{H_2O, \delta D\}$ pair data during this data assimilation procedure.

Within the project S5P+Innovation H₂O-ISO (<https://s5pinnovationh2o-iso.le.ac.uk/>) under the coordination of the European Space Agency (ESA), a retrieval processor is being developed that generates total columns of H₂O and δD based on data from the sensor TROPOMI onboard the Sentinel-5P satellite. This product is currently being evaluated by using the MUSICA IASI $\{H_2O, \delta D\}$ pair data together with the ICON-ART_{iso} simulations discussed here. As TROPOMI offers high sensitivity near the ground, where the MUSICA IASI retrieval mostly lacks sensitivity, Schneider et al. (2021b) present a method for a posteriori combining results from both sensors. The combined product exhibits strong vertical profile information, i.e. the variability in the boundary layer can be clearly distinguished from the free troposphere. This approach is currently performed and validated only for CH₄, however, it is planned to extend this to the retrieval results of H₂O and δD . This is of particular interest, as both sensors will be jointly operated on the upcoming Metop Second Generation satellites. Thereby each satellite will provide more than 1 million perfectly collocated observations of IASI and TROPOMI per day on a global scale for the next 20 years. This enormous availability of concurrent observations from both sensors will offer highly promising opportunities for investigating the hydrological cycle, as such a combined product would be capable of tracing water vapor close to its evaporative sources up to the mid-troposphere, where it is affected by large-scale transport and microphysical cloud and rain processes.

Part V.

Appendix

A | Dust Analysis of {H₂O, δD} Pairs over Tenerife

In this appendix, light will be shed on strongly enriched MUSICA IASI {H₂O, δD} signals in the mid-troposphere (δD above -100 ‰) for H₂O values below 10 000 ppmv, as identified over Tenerife and the Sahara (see Fig. 4.22a,b). Several studies found a strong correlation of similar structures over Tenerife to high abundances of dust that was uplifted over the Sahara and transported to Tenerife within the Saharan air layer (SAL) (Dyroff et al., 2015, Schneider et al., 2015, 2016, González et al., 2016). The SAL is a warm, dry and dust-laden layer between 1–5 km with a dust maximum between 2–3 km and is capable of spreading dust from Saharan outbreaks over long distances of several thousand kilometres (e.g. Parker et al., 2005, Knippertz and Fink, 2006, Rodríguez et al., 2011, Wang and Liu, 2014). As strong vertical mixing processes may occur within the SAL, near-surface humidity can be efficiently uplifted from the boundary layer to higher altitudes, leading to highly enriched mid-tropospheric water vapor (González et al., 2016). Therefore, this appendix addresses the question, whether the aforementioned enriched {H₂O, δD} signals over Tenerife may be attributed to the enhanced vertical moisture transport during SAL events. To this aim, sun-sky radiometry measurements of aerosol optical depth (AOD, measured with a wavelength of 500 nm), provided by the Aerosol Robotic Network (AERONET; Giles et al., 2019), is used. Figure A.1 shows the AOD time-series as measured in Izana (Tenerife) from 01 June to 31 July 2016. Measurements with a AOD higher than 0.05 are here treated as dust-laden, while the upper threshold of 0.02 is assumed to determine dust-free measurements. As next step, the MUSICA IASI {H₂O, δD} pair data over Tenerife are then classified according to the chosen dust criteria into dust-free and dust-laden data points.

The resulting distributions of classified MUSICA IASI observations are shown in Fig. A.2a. Additionally, the two-dimensional histogram contours for the full IASI observations as well as for the raw COSMO_{iso} output (dataset ID CISO2, see Table 3.2) and for the COSMO_{iso} output processed with the MUSICA IASI retrieval simulator is shown (Fig. A.2b, analogous to Fig. 4.22, but here shown for 4.9 km, because at this altitude the enriched {H₂O, δD} signals appear to be more pronounced). It becomes apparent that the data with high values in H₂O and δD strongly correlate with dust abundances. This thereby supports the findings of Schneider et al. (2015, 2016) and González et al. (2016), that dust-laden SAL events are associated with a significant

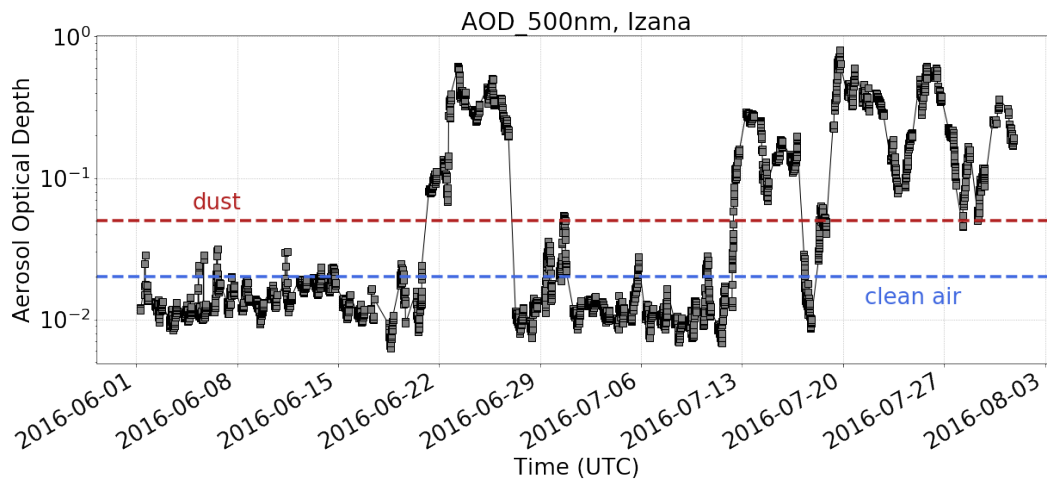


Figure A.1. Time-series of the AOD over Izana (Tenerife) based on data from AERONET (Giles et al., 2019). The red dashed line indicates the lower threshold for identifying dust-laden measurements, whereas the blue dashed line indicates the upper threshold for identifying dust-free measurements.

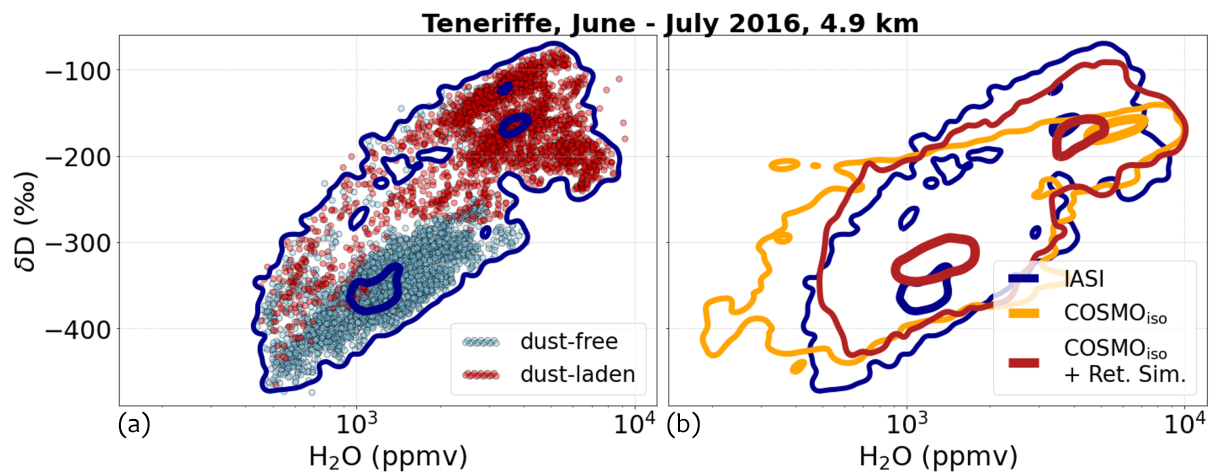


Figure A.2. (a) MUSICA IASI $\{H_2O, \delta D\}$ pair data at 4.9 km, classified according to the AOD from Fig. A.1. The blue contours summarize 50 % and 99 % of all (unclassified) MUSICA IASI observations over Tenerife for the given period. (b) Same as Fig. 4.22, but here for 4.9 km. Data from the COSMO₁₅₀ simulation CISO2 (see Table 3.2) are shown by the orange contours, and after applying the MUSICA IASI retrieval simulator on CISO2, the resulting data are shown by the contours.

moistening and enrichment of the mid-troposphere over Tenerife. Further, only by applying the retrieval simulator the COSMO_{iso} model is able to reflect these enriched values up to -100 ‰, as observed by the IASI sensor. This again emphasizes the value of considering the MUSICA IASI retrieval simulator for direct model-to-IASI comparisons.

B | Yearly $\{\text{H}_2\text{O}, \delta\text{D}\}$ Pair Data and Precipitation Distributions

This section provides yearly $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair distributions at 4.2 km over the Guinea Coast and the Sahel together with the respective precipitation histograms. Figures B.1 – B.4 show the MUSICA IASI $\{\text{H}_2\text{O}, \delta\text{D}\}$ pair data for the months February, May, August and November and the monthly averaged GPM IMERG precipitation data (RGPM2, see Table 3.2) individually for the years 2015–2018. These yearly data have been used to compute the multi-annual averages shown in Fig. 6.1.

Figures B.5 – B.16 show the $\{\text{H}_2\text{O}, \delta\text{D}\}$ distributions for the months February, May, July and November as well as the monthly averaged precipitation, both from the ICON-ART_{iso} simulations IISO1–IISO4 (see Table 3.2) individually for the years 2017–2019. These yearly data have been used to compute the multi-annual averages shown in Fig. 6.3 and 6.4.

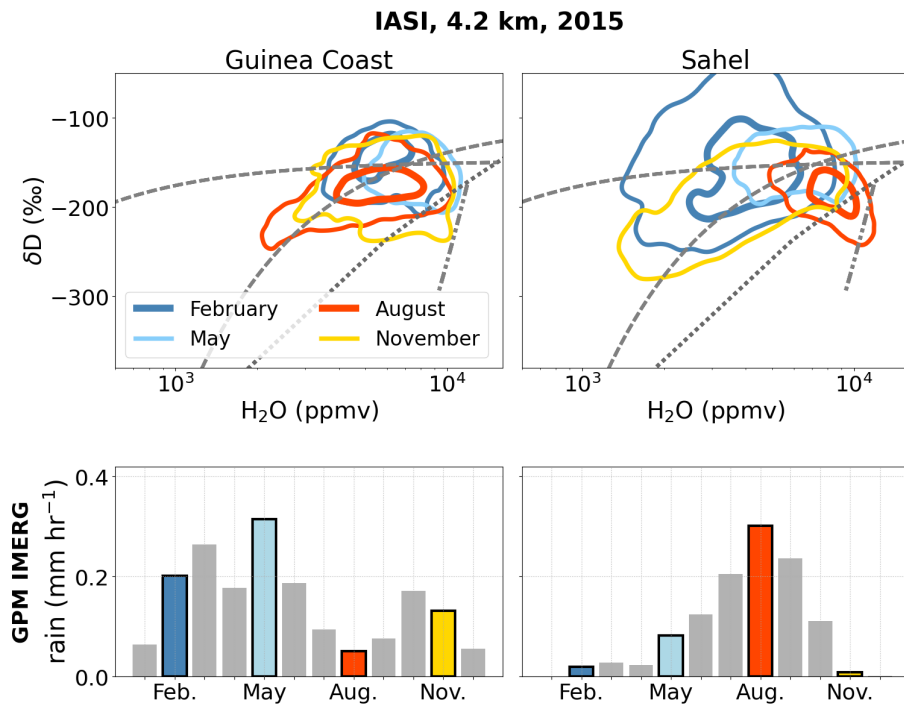


Figure B.1. $\{H_2O, \delta D\}$ pair distributions from MUSICA IASI and monthly averaged precipitation data from GPM IMERG (RGPM2), for the year 2015.

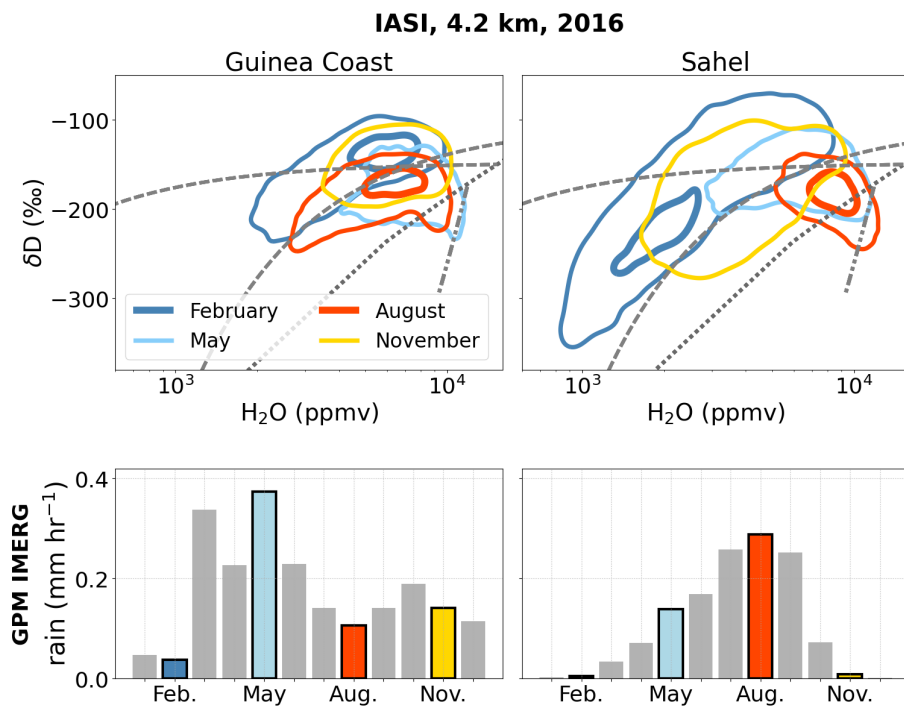


Figure B.2. $\{H_2O, \delta D\}$ pair distributions from MUSICA IASI and monthly averaged precipitation data from GPM IMERG (RGPM2), for the year 2016.

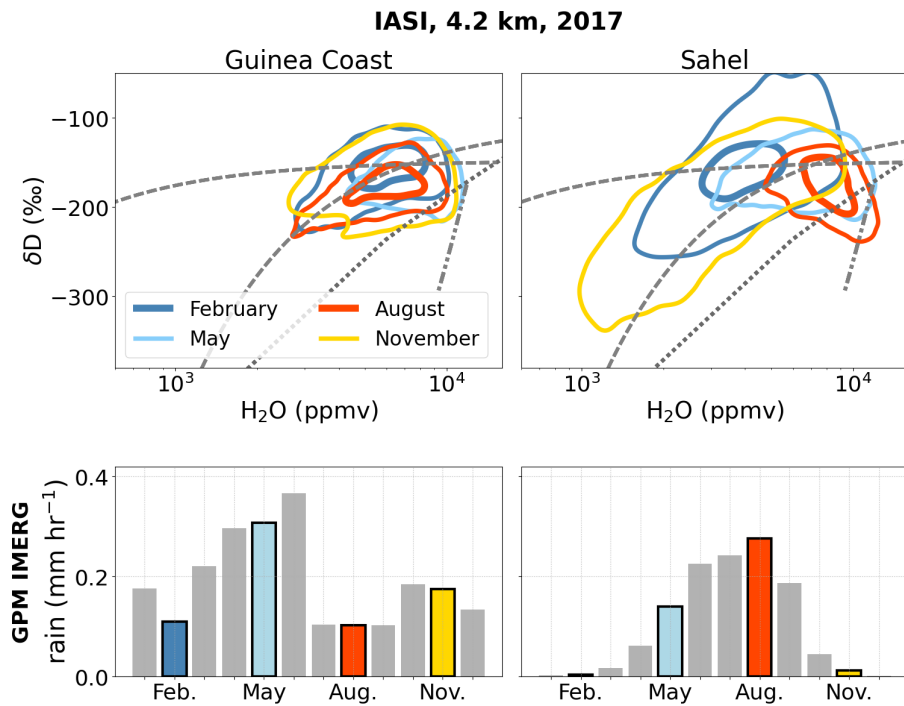


Figure B.3. $\{H_2O, \delta D\}$ pair distributions from MUSICA IASI and monthly averaged precipitation data from GPM IMERG (RGPM2), for the year 2017.

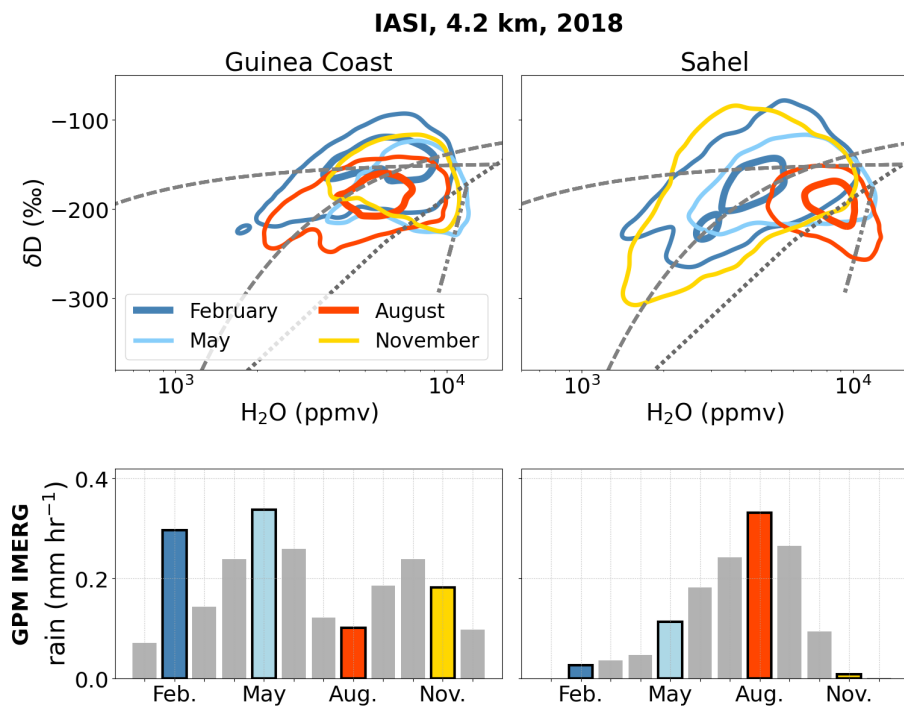


Figure B.4. $\{H_2O, \delta D\}$ pair distributions from MUSICA IASI and monthly averaged precipitation data from GPM IMERG (RGPM2), for the year 2018.

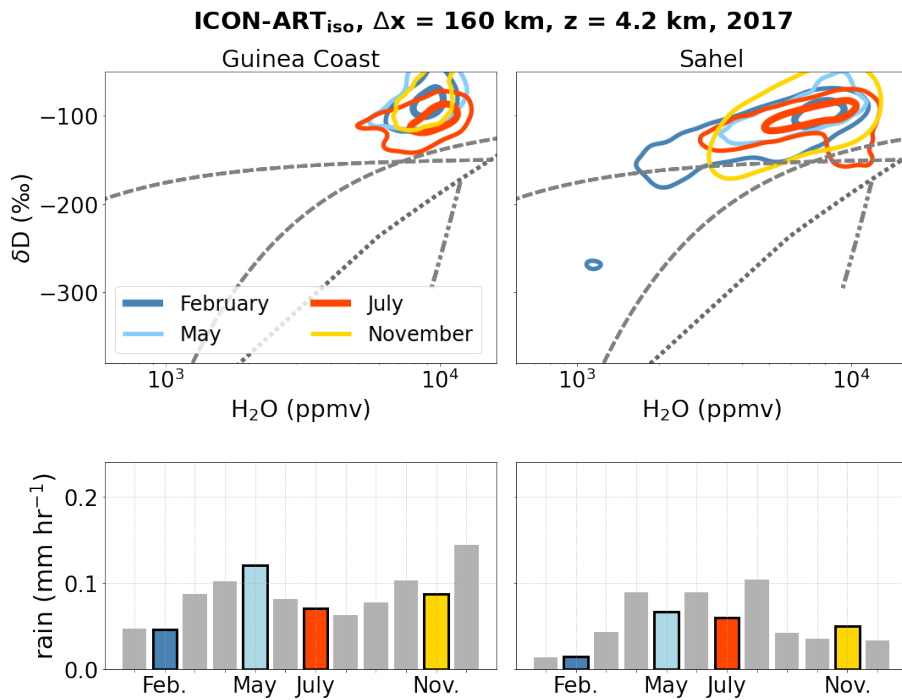


Figure B.5. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO1 for the year 2017.

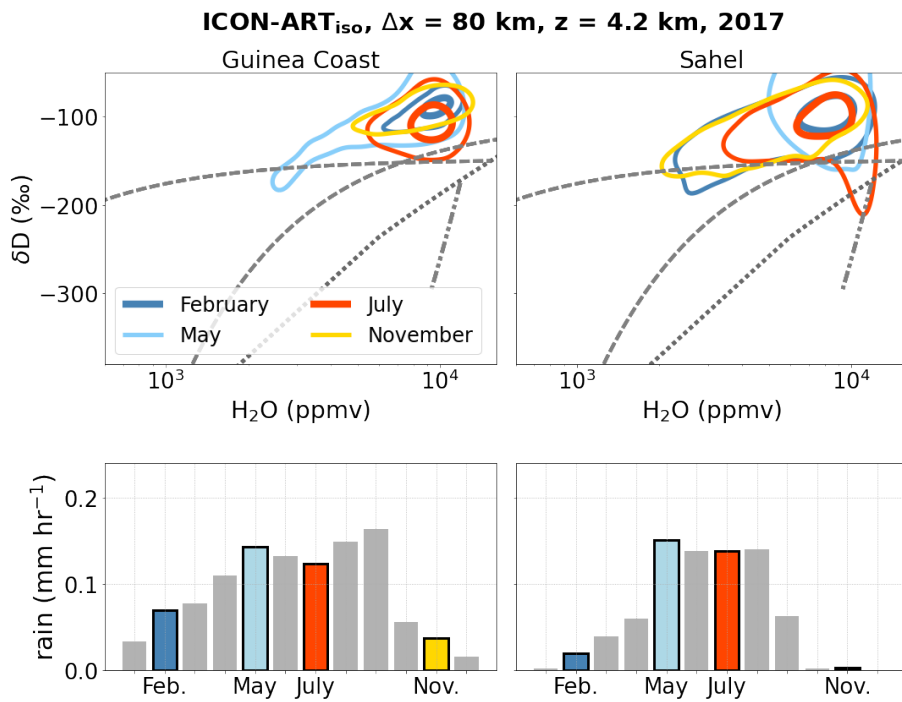


Figure B.6. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO2 for the year 2017.

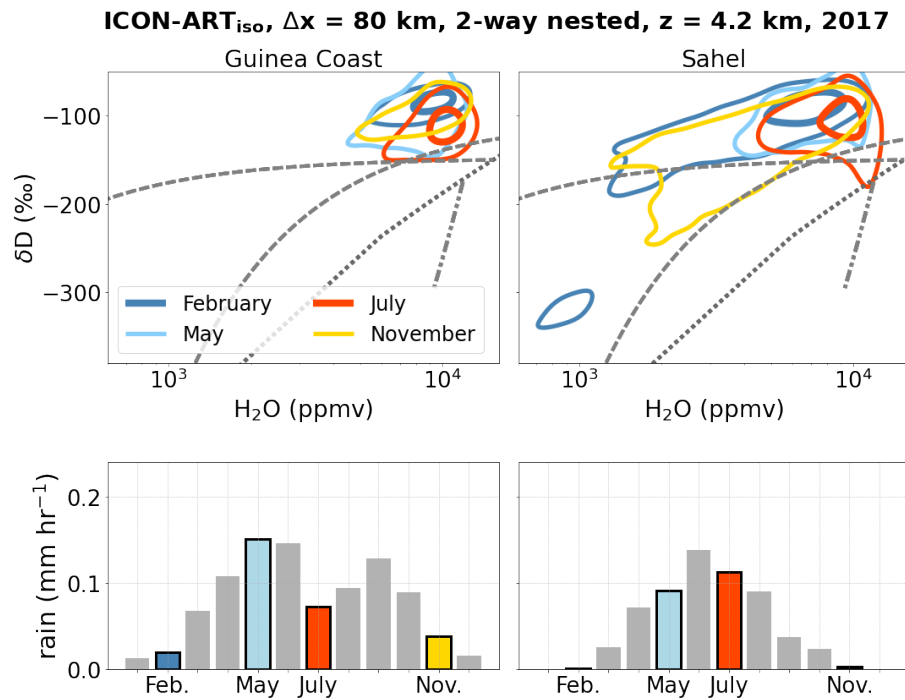


Figure B.7. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO3 for the year 2017.

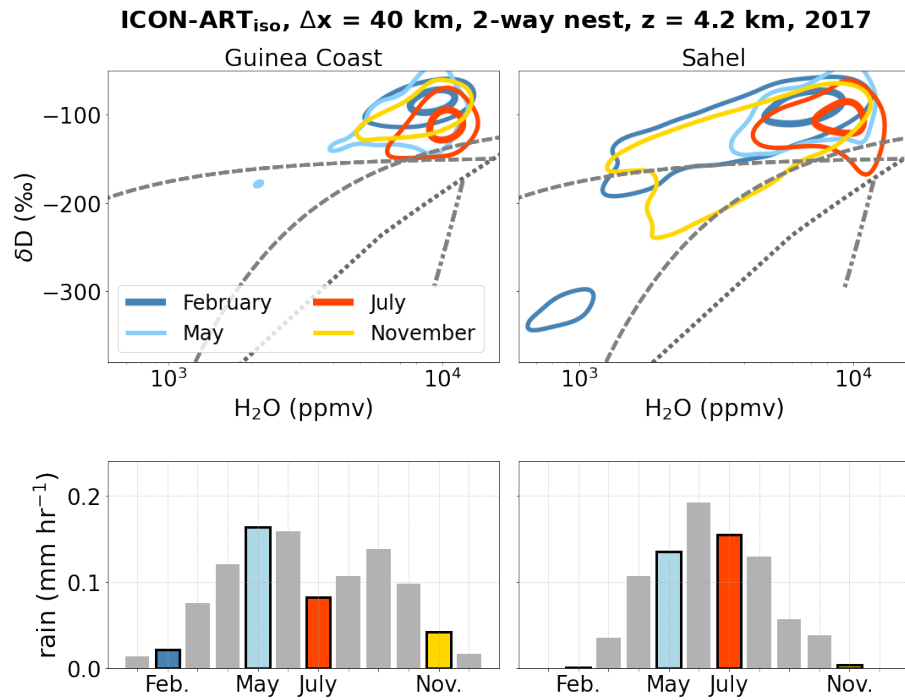


Figure B.8. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO4 for the year 2017.

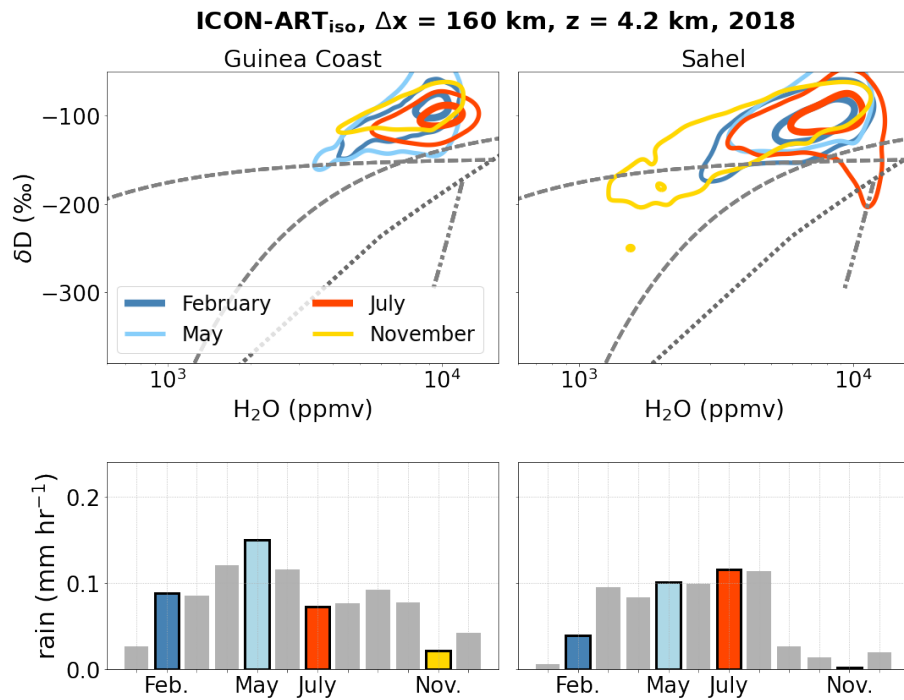


Figure B.9. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO1 for the year 2018.

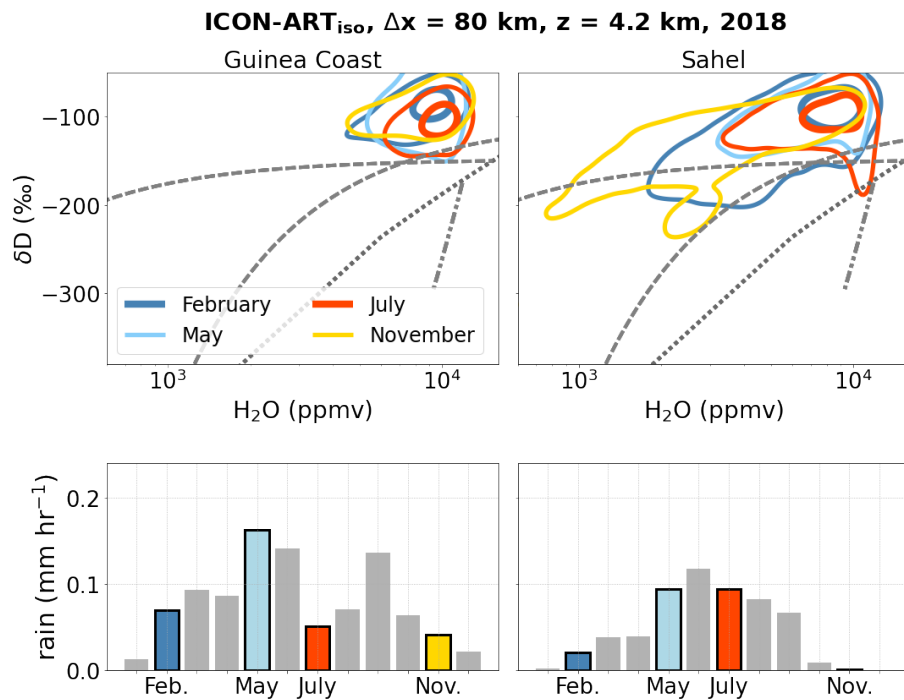


Figure B.10. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO2 for the year 2018.

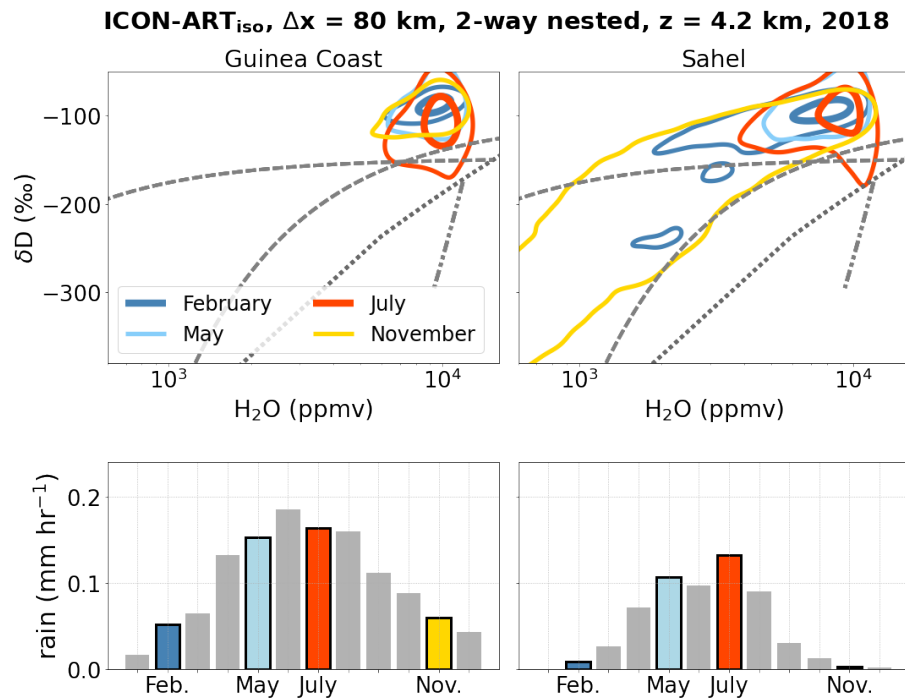


Figure B.11. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO3 for the year 2018.

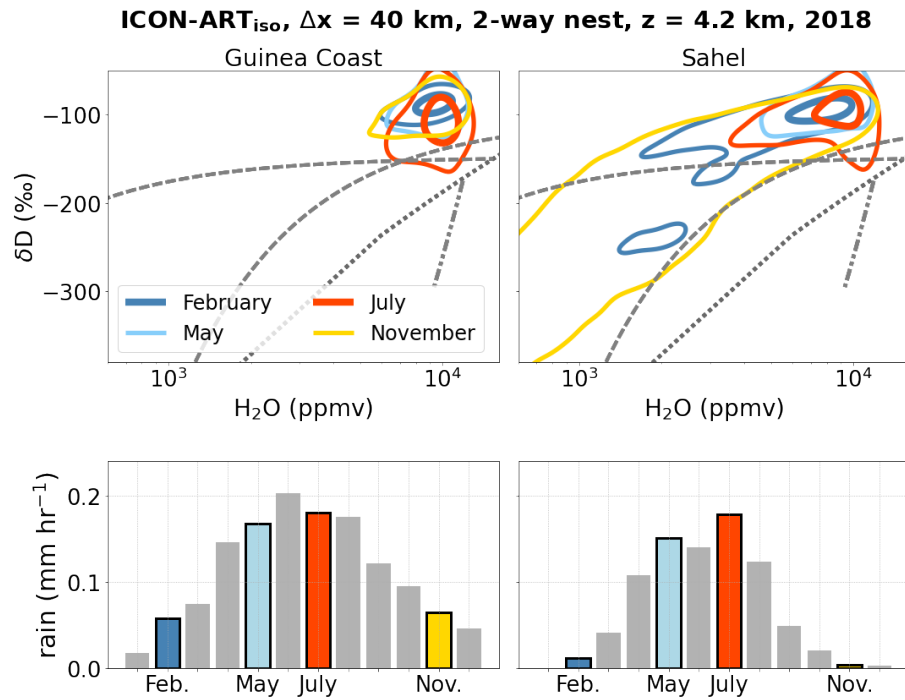


Figure B.12. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO4 for the year 2018.

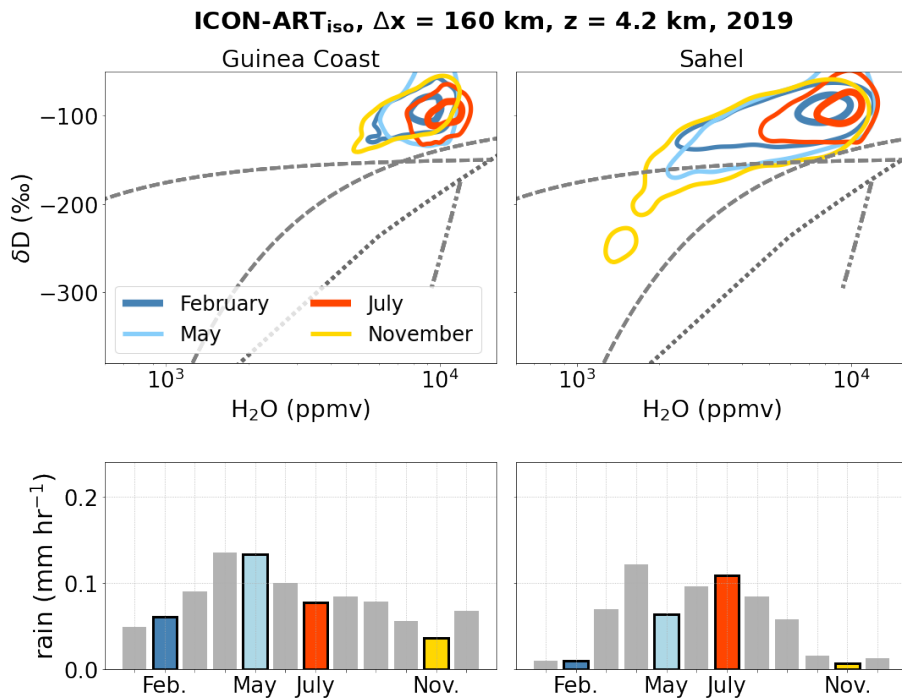


Figure B.13. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO1 for the year 2019.

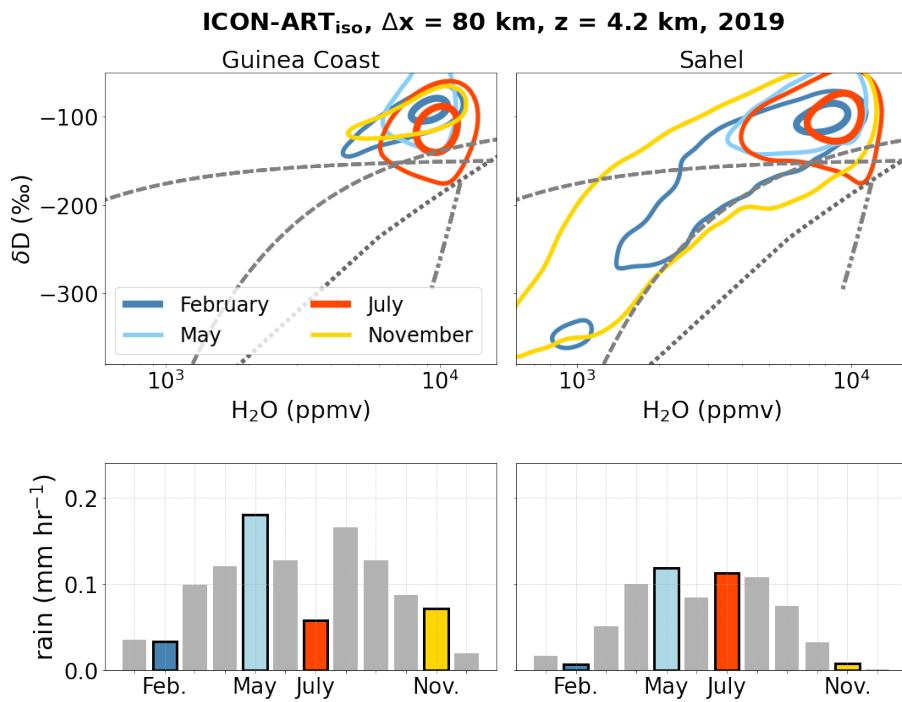


Figure B.14. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO2 for the year 2019.

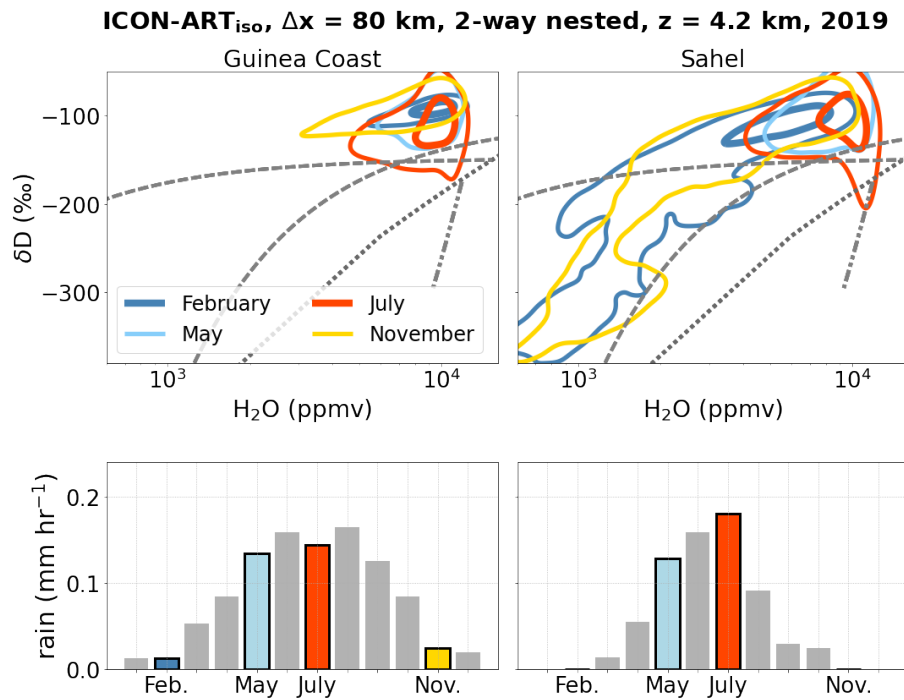


Figure B.15. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO3 for the year 2019.

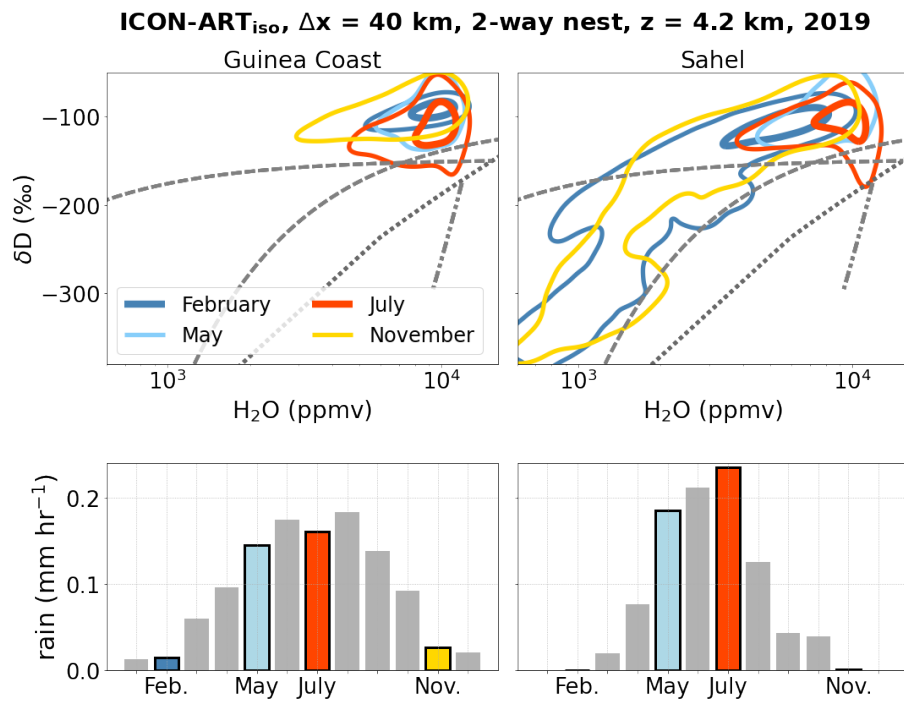


Figure B.16. $\{H_2O, \delta D\}$ pair distributions and monthly averaged precipitation data from the ICON-ART_{iso} simulation IISO4 for the year 2019.

C | List of Acronyms

ACT	Atlantic Cold Tongue
AEJ	African Easterly Jet
AIRS	Atmospheric Infrared Sounder
AOD	Aerosols Optical Depth
AERONET	Aerosol Robotic Network
AMIP	Atmospheric Model Intercomparison Project
AMMA	African Monsoon Multidisciplinary Analysis
ASTER	Advanced Spaceborne Thermal Emission Reflection Radiometer
CESM1/WACCM	Coupled Chemistry-Climate Model / Whole Atmosphere Community Climate Model
COSMO_{iso}	Consortium for Small-scale Modeling, Isotopologues
DACCIWA	Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa
DOFS	Degree of Freedom for Signal
ECHAM5_{wiso}	ECMWF and Hamburg
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	EUMETSAT Polar System
ERA-interim	ECMWF Reanalysis-Interim
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellite
EUREC4A	Elucidating the ROle of Clouds-Circulation Coupling in Climate
GCM	Global Circulation Model
GNIP	Global Network of Isotopes in Precipitation
GOSAT	Greenhouse Observing Satellite
GPCC	Global Precipitation Climatology Centre
GPM	Global Precipitation Measurement
HITRAN	High-Resolution Transmission Molecular Absorption Database
IAEA	International Atomic Energy Agency
IASI	Infrared Atmospheric Sounding Instrument
IASI-NG	IASI Next Generation
ICON-ART_{iso}	Icosahedral Nonhydrostatic - Aerosols and Reactive Trace Gases - Isotopologues
ITD	Intertropical Discontinuity
IMERG	Integrated Multi-Satellite Retrievals for GPM

C. List of Acronyms

IREMIS	Global Infrared Land SURface Emissivity
JAXA	Japan Aerospace and Exploration Agency
LAGRANTO	Lagrangian Analysis Tool
MCS	Mesoscale Convective Systems
Metop	Meteorological Operational Satellite
MUSICA	Multi-platform remote sensing of isotopologues for investigating the cycle of Atmospheric water
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detection of Atmospheric Composition Change
PAGES	Past Global Changes
PROFFIT	Profile Fit
PRFFWD	PROFFIT Forward Model
RCM	Regional Climate Model
SAL	Saharan Air Layer
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
S5P	Sentinel-5P
SHL	Saharan Heat Low
SWING	Stable Water Isotope Intercomparison Group
SWVID	Stable Water Vapour Isotope Database
TCCON	Total Carbon Column Observing Network
TES	Tropospheric Emission Sounder
TROPOMI	Tropospheric Monitoring Instrument
VSMOW	Vienna Standard Mean Ocean Water
WAM	West African Monsoon
WCRP	World Climate Research Programme

Bibliography

- F. Aemisegger, J. K. Spiegel, S. Pfahl, H. Sodemann, W. Eugster, and H. Wernli. Isotope meteorology of cold front passages: A case study combining observations and modeling. *Geophysical Research Letters*, 42(13):5652–5660, jul 2015. ISSN 19448007. doi: 10.1002/2015GL063988.
- G. Backus and F. Gilbert. Uniqueness in the inversion of inaccurate gross Earth data. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 266(1173):123–192, mar 1970. ISSN 0080-4614. doi: 10.1098/rsta.1970.0005.
- A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113(4):711–715, apr 2009. ISSN 00344257. doi: 10.1016/j.rse.2008.11.007.
- P. Baron, P. Ricaud, J. De la Noë, J. E. Eriksson, F. Merino, M. Ridal, and D. P. Murtagh. Studies for the Odin sub-millimetre radiometer. II. Retrieval methodology. *Canadian Journal of Physics*, 80(4):341–356, 2002. ISSN 00084204. doi: 10.1139/P01-150.
- S. Barthlott, M. Schneider, F. Hase, T. Blumenstock, M. Kiel, D. Dubravica, O. E. García, E. Sepúlveda, G. Mengistu Tsidu, S. Takele Kenea, M. Grutter, E. F. Plaza-Medina, W. Stremme, K. Strong, D. Weaver, M. Palm, T. Warneke, J. Notholt, E. Mahieu, C. Servais, N. Jones, D. W. Griffith, D. Smale, and J. Robinson. Tropospheric water vapour isotopologue data (H₂16O, H₂18O, and HD₁₆O) as obtained from NDACC/FTIR solar absorption spectra. *Earth System Science Data*, 9(1):15–29, jan 2017. ISSN 18663516. doi: 10.5194/essd-9-15-2017.
- E. Berntell, Q. Zhang, L. Chafik, and H. Körnich. Representation of Multidecadal Sahel Rainfall Variability in 20th Century Reanalyses. *Scientific Reports*, 8(1):10937, dec 2018. ISSN 20452322. doi: 10.1038/s41598-018-29217-9.
- G. J. Berry and C. Thorncroft. Case study of an intense African easterly wave. *Monthly Weather Review*, 133(4): 752–766, apr 2005. ISSN 00270644. doi: 10.1175/MWR2884.1.
- G. J. Berry and C. D. Thorncroft. African easterly wave dynamics in a mesoscale numerical model: The upscale role of convection. *Journal of the Atmospheric Sciences*, 69(4):1267–1283, apr 2012. ISSN 00224928. doi: 10.1175/JAS-D-11-099.1.
- S. Berthou, D. P. Rowell, E. J. Kendon, M. J. Roberts, R. A. Stratton, J. A. Crook, and C. Wilcox. Improved climatological precipitation characteristics over West Africa at convection-permitting scales. *Climate Dynamics*, 53(3-4):1991–2011, aug 2019. ISSN 14320894. doi: 10.1007/s00382-019-04759-4.
- S. Bielli, H. Douville, and B. Pohl. Understanding the West African monsoon variability and its remote effects: An illustration of the grid point nudging methodology. *Climate Dynamics*, 35(1):159–174, 2010. ISSN 09307575. doi: 10.1007/s00382-009-0667-8.
- J. Bigeleisen. Statistical Mechanics of Isotope Effects on the Thermodynamic Properties of Condensed Systems. *The Journal of Chemical Physics*, 34(5):1485–1493, 1961. ISSN 10897690. doi: 10.1063/1.1701033.

- H. Boesch, N. M. Deutscher, T. Warneke, K. Byckling, A. J. Cogan, D. W. Griffith, J. Notholt, R. J. Parker, and Z. Wang. HDO/H₂O ratio retrievals from GOSAT. *Atmospheric Measurement Techniques*, 6(3):599–612, 2013. ISSN 18671381. doi: 10.5194/amt-6-599-2013.
- M. Bolot, B. Legras, and E. J. Moyer. Modelling and interpreting the isotopic composition of water vapour in convective updrafts. *Atmospheric Chemistry and Physics*, 13(16):7903–7935, 2013. ISSN 16807316. doi: 10.5194/acp-13-7903-2013.
- S. Bony, C. Risi, and F. Vimeux. Influence of convective processes on the isotopic composition ($\delta^{18}\text{O}$ and δD) of precipitation and water vapor in the tropics: 1. Radiative-convective equilibrium and Tropical Ocean-Global Atmosphere-Coupled Ocean-Atmosphere Response Experiment (TOGAR-COARE) simulation. *Journal of Geophysical Research Atmospheres*, 113(19):1–21, 2008. ISSN 01480227. doi: 10.1029/2008JD009942.
- C. Borger, M. Schneider, B. Ertl, F. Hase, O. E. García, M. Sommer, M. Höpfner, S. A. Tjemkes, and X. Calbet. Evaluation of MUSICA IASI tropospheric water vapour profiles using theoretical error assessments and comparisons to GRUAN Vaisala RS92 measurements. *Atmospheric Measurement Techniques*, 11(9):4981–5006, sep 2018. ISSN 18678548. doi: 10.5194/amt-11-4981-2018.
- C. A. Brenninkmeijer, P. Crutzen, F. Boumard, T. Dauer, B. Dix, R. Ebinghaus, D. Filippi, H. Fischer, H. Franke, U. Friß, J. Heintzenberg, F. Helleis, M. Hermann, H. H. Kock, C. Koeppel, J. Lelieveld, M. Leuenberger, B. G. Martinsson, S. Miemczyk, H. P. Moret, H. N. Nguyen, P. Nyfeler, D. Oram, D. O’Sullivan, S. Penkett, U. Platt, M. Pucek, M. Ramonet, B. Randa, M. Reichelt, T. S. Rhee, J. Rohwer, K. Rosenfeld, D. Scharffe, H. Schlager, U. Schumann, F. Slemr, D. Sprung, P. Stock, R. Thaler, F. Valentino, P. Van Velthoven, A. Waibel, A. Wandel, K. Waschitschek, A. Wiedensohler, I. Xueref-Remy, A. Zahn, U. Zech, and H. Ziereis. Civil Aircraft for the regular investigation of the atmosphere based on an instrumented container: The new CARIBIC system. *Atmospheric Chemistry and Physics*, 7(18):4953–4976, sep 2007. ISSN 16807324. doi: 10.5194/acp-7-4953-2007.
- R. W. Burpee. The Origin and Structure of Easterly Waves in the Lower Troposphere of North Africa. *Journal of the Atmospheric Sciences*, 29(1):77–90, jan 1972. ISSN 0022-4928. doi: 10.1175/1520-0469(1972)029<0077:toasoe>2.0.co;2.
- D. L. Cadet and N. O. Nnoli. Water Vapour Transport Over Africa and the Atlantic Ocean During Summer 1979. *Quarterly Journal of the Royal Meteorological Society*, 113(476):581–602, jul 1987. ISSN 1477870X. doi: 10.1002/qj.49711347609.
- G. Caniaux, H. Giordani, J. L. Redelsperger, F. Guichard, E. Key, and M. Wade. Coupling between the Atlantic cold tongue and the West African monsoon in boreal spring and summer. *Journal of Geophysical Research: Oceans*, 116(4), apr 2011. ISSN 21699291. doi: 10.1029/2010JC006570.
- T. N. Carlson. Some remarks on african disturbances and their progress over the tropical Atlantic. *Monthly Weather Review*, 97(10):716–726, oct 1969. ISSN 0027-0644. doi: 10.1175/1520-0493(1969)097<0716:sroada>2.3.co;2.
- H. Celle-Jeanton, R. Gonfiantini, Y. Travi, and B. Sol. Oxygen-18 variations of rainwater during precipitation: Application of the Rayleigh model to selected rainfalls in Southern France. *Journal of Hydrology*, 289(1-4):165–177, apr 2004. ISSN 00221694. doi: 10.1016/j.jhydrol.2003.11.017.
- T. C. Chen. Maintenance of the midtropospheric North African summer circulation; Saharan high and African easterly jet. *Journal of Climate*, 18(15):2943–2962, aug 2005. ISSN 08948755. doi: 10.1175/JCLI3446.1.

- E. Christner, F. Aemisegger, S. Pfahl, M. Werner, A. Cauquoin, M. Schneider, F. Hase, S. Barthlott, and G. Schädler. The Climatological Impacts of Continental Surface Evaporation, Rainout, and Subcloud Processes on δD of Water Vapor and Precipitation in Europe. *Journal of Geophysical Research: Atmospheres*, 123(8):4390–4409, apr 2018. ISSN 21698996. doi: 10.1002/2017JD027260.
- P. Ciais and J. Jouzel. Deuterium and oxygen 18 in precipitation: isotopic model, including mixed cloud processes. *Journal of Geophysical Research*, 99(D8), 1994. ISSN 01480227. doi: 10.1029/94jd00412.
- C. Clerbaux, A. Boynard, L. Clarisse, M. George, J. Hadji-Lazaro, H. Herbin, D. Hurtmans, M. Pommier, A. Razavi, S. Turquety, C. Wespes, and P. F. Coheur. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmospheric Chemistry and Physics*, 9(16):6041–6054, 2009. ISSN 16807324. doi: 10.5194/acp-9-6041-2009.
- K. H. Cook. Generation of the African easterly jet and its role in determining West African precipitation. *Journal of Climate*, 12(5 I):1165–1184, 1999. ISSN 08948755. doi: 10.1175/1520-0442(1999)012<1165:GOTAEJ>2.0.CO;2.
- H. Craig and L. Gordon. Deuterium and oxygen 18 variations in the ocean and the marine atmosphere. *Stable Isotopes in Oceanographic Studies and Paleotemperatures*, (Cons. Naz. Ric., Lab. Geol. Nucl. Pisa):9–130, 1965.
- J. Crétat, E. K. Vizy, and K. H. Cook. The relationship between African easterly waves and daily rainfall over West Africa: observations and regional climate simulations. *Climate Dynamics*, 44(1-2):385–404, jan 2015. ISSN 14320894. doi: 10.1007/s00382-014-2120-x.
- J. Crook, C. Klein, S. Folwell, C. M. Taylor, D. J. Parker, R. Stratton, and T. Stein. Assessment of the Representation of West African Storm Lifecycles in Convection-Permitting Simulations. *Earth and Space Science*, 6(5):818–835, may 2019. ISSN 23335084. doi: 10.1029/2018EA000491.
- F. Dahinden, F. Aemisegger, H. Wernli, M. Schneider, C. J. Diekmann, B. Ertl, P. Knippertz, and S. Pfahl. Disentangling different moisture transport pathways over the eastern subtropical North Atlantic using multi-platform isotope observations and high-resolution numerical modelling. *in prep. for Atmospheric Chemistry and Physics*, 2021.
- W. Dansgaard. Stable isotopes in precipitation. *Tellus*, 16(4):436–468, nov 1964. ISSN 00402826. doi: 10.1111/j.2153-3490.1964.tb00181.x.
- J. Davis, P. Knippertz, and A. H. Fink. The predictability of precipitation episodes during the West African dry season. *Quarterly Journal of the Royal Meteorological Society*, 139(673):1047–1058, apr 2013. ISSN 00359009. doi: 10.1002/qj.2014.
- G. Dhonneur. Les amas nuageux mobiles, principale composante de la météorologie du Sahel. *La Météorologie Paris*, 6(27):75–82, 1981.
- C. J. Diekmann, M. Schneider, and B. Ertl. MUSICA IASI water isotopologue pair product (a posteriori processing version 2), 2021a. dataset, Institute of Meteorology and Climate Research, Atmospheric Trace Gases and Remote Sensing (IMK-ASF), Karlsruhe Institute of Technology (KIT), doi: 10.35097/415.
- C. J. Diekmann, M. Schneider, B. Ertl, F. Hase, O. E. García, F. Khosrawi, E. Sepulveda, P. Knippertz, and P. Braesicke. The MUSICA IASI H₂O, δD pair product. *Earth Syst. Sci. Data Discuss. [preprint]*, 2021b. doi: 10.5194/essd-2021-87.

- C. J. Diekmann, M. Schneider, P. Knippertz, A. de Vries, F. Aemisegger, F. Dahinden, B. Ertl, F. Khosrawi, S. Pfahl, H. Wernli, and P. Braesicke. A Lagrangian perspective on stable water isotopes during the West African Monsoon. *subm. to Journal of Geophysical Research Atmospheres*, 2021c. doi: 10.1002/ESSOAR.10506628.1.
- L. M. Druyan and M. Fulakeza. The impact of the Atlantic cold tongue on West African monsoon onset in regional model simulations for 1998-2002. *International Journal of Climatology*, 35(2):275–287, feb 2014. ISSN 10970088. doi: 10.1002/joc.3980.
- M. Dütsch, S. Pfahl, M. Meyer, and H. Wernli. Lagrangian process attribution of isotopic variations in near-surface water vapour in a 30-year regional climate simulation over Europe. *Atmospheric Chemistry and Physics*, 18(3): 1653–1669, 2018. ISSN 16807324. doi: 10.5194/acp-18-1653-2018.
- C. Dyroff, S. Sanati, E. Christner, A. Zahn, M. Balzer, H. Bouquet, J. B. McManus, Y. González-Ramos, and M. Schneider. Airborne in situ vertical profiling of HDO/H₂¹⁶O in the subtropical troposphere during the MUSICA remote sensing validation campaign. *Atmospheric Measurement Techniques*, 8(5):2037–2049, 2015. ISSN 18678548. doi: 10.5194/amt-8-2037-2015.
- J. Eckstein. *Atmospheric Models and Aircraft Measurements: Representativeness and Model Validation*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2017.
- J. Eckstein, R. Ruhnke, S. Pfahl, E. Christner, C. J. Diekmann, C. Dyroff, D. Reinert, D. Rieger, M. Schneider, J. Schröter, A. Zahn, and P. Braesicke. From climatological to small-scale applications: Simulating water isotopologues with ICON-ART-Iso (version 2.3). *Geoscientific Model Development*, 11(12):5113–5133, 2018. ISSN 19919603. doi: 10.5194/gmd-11-5113-2018.
- P. Eriksson. Analysis and comparison of two linear regularization methods for passive atmospheric observations. *Journal of Geophysical Research Atmospheres*, 105(D14):18157–18167, jul 2000. ISSN 01480227. doi: 10.1029/2000JD900172.
- B. Ertl, J. Meyer, M. Schneider, C. J. Diekmann, and A. Streit. A Semi-Supervised Approach for Trajectory Segmentation to Identify Different Processes in the Atmosphere. *subm. to International Conference on Computational Science*, 2021.
- R. D. Field, C. Risi, G. A. Schmidt, J. Worden, A. Voulgarakis, A. N. Legrande, A. H. Sobel, and R. J. Healy. A Tropospheric Emission Spectrometer HDO/H₂O retrieval simulator for climate models. *Atmospheric Chemistry and Physics*, 12(21):10485–10504, 2012. ISSN 16807316. doi: 10.5194/acp-12-10485-2012.
- A. H. Fink and A. Reiner. Spatiotemporal variability of the relation between African Easterly Waves and West African Squall Lines in 1998 and 1999. *Journal of Geophysical Research: Atmospheres*, 108(11), jun 2003. ISSN 01480227. doi: 10.1029/2002jd002816.
- A. H. Fink, D. G. Vincent, and V. Ermert. Rainfall types in the West African Sudanian zone during the summer monsoon 2002. *Monthly Weather Review*, 134(8):2143–2164, aug 2006. ISSN 00270644. doi: 10.1175/MWR3182.1.
- A. H. Fink, T. Engel, V. Ermert, R. Van Der Linden, M. Schneidewind, R. Redl, E. Afiesimama, W. M. Thiaw, C. Yorke, M. Evans, and S. Janicot. Mean climate and seasonal cycle. In *Meteorology of Tropical West Africa: The Forecasters' Handbook*, pages 1–39. John Wiley & Sons, Ltd, Chichester, UK, feb 2017. ISBN 9781118391297. doi: 10.1002/9781118391297.ch1.

- R. G. Fitzpatrick, C. L. Bain, P. Knippertz, J. H. Marsham, and D. J. Parker. The West African monsoon onset: A concise comparison of definitions. *Journal of Climate*, 28(22):8673–8694, nov 2015. ISSN 08948755. doi: 10.1175/JCLI-D-15-0265.1.
- C. Flamant, P. Knippertz, A. H. Fink, A. Akpo, B. Brooks, C. J. Chiu, H. Coe, S. Danuor, M. Evans, O. Jegede, N. Kalthoff, A. Konaré, C. Liousse, F. Lohou, C. Mari, H. Schlager, A. Schwarzenboeck, B. Adler, L. Amekudzi, J. Aryee, M. Ayoola, A. M. Batenburg, G. Bessardon, S. Borrmann, J. Brito, K. Bower, F. Burnet, V. Catoire, A. Colomb, C. Denjean, K. Fosu-Amankwah, P. G. Hill, J. Lee, M. Lothon, M. Maranan, J. Marsham, R. Meynadier, J. B. Ngamini, P. Rosenberg, D. Sauer, V. Smith, G. Stratmann, J. W. Taylor, C. Voigt, and V. Yoboué. The dynamics-aerosol-chemistry-cloud interactions in West Africa field campaign overview and research highlights. *Bulletin of the American Meteorological Society*, 99(1):83–104, jan 2018. ISSN 00030007. doi: 10.1175/BAMS-D-16-0256.1.
- B. Fontaine, P. Roucou, and S. Trzaska. Atmospheric water cycle and moisture fluxes in the West African monsoon: Mean annual cycles and relationship using NCEP/NCAR reanalysis. *Geophysical Research Letters*, 30(3):1117, feb 2003. ISSN 00948276. doi: 10.1029/2002GL015834.
- C. Frankenberg, K. Yoshimura, T. Warneke, I. Aben, A. Butz, N. Deutscher, D. Griffith, F. Hase, J. Notholt, M. Schneider, H. Schrijver, and T. Röckmann. Dynamic processes governing lower-tropospheric HDO/H₂O Ratios as Observed from Space and Ground. *Science*, 325(5946):1374–1377, 2009. ISSN 00368075. doi: 10.1126/science.1173791.
- C. Frankenberg, D. Wunch, G. Toon, C. Risi, R. Scheepmaker, J. E. Lee, P. Wennberg, and J. Worden. Water vapor isotopologue retrievals from high-resolution GOSAT shortwave infrared spectra. *Atmospheric Measurement Techniques*, 6(2):263–274, 2013. ISSN 18671381. doi: 10.5194/amt-6-263-2013.
- M. Gaetani, B. Pohl, H. Douville, and B. Fontaine. West African Monsoon influence on the summer Euro-Atlantic circulation. *Geophysical Research Letters*, 38(9):n/a–n/a, may 2011. ISSN 00948276. doi: 10.1029/2011GL047150.
- J. Galewsky, H. C. Steen-Larsen, R. D. Field, J. Worden, C. Risi, and M. Schneider. Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle. *Reviews of Geophysics*, 54(4):809–865, 2016. ISSN 19449208. doi: 10.1002/2015RG000512.
- O. E. García, M. Schneider, B. Ertl, E. Sepúlveda, C. Borger, C. Diekmann, A. Wiegeler, F. Hase, S. Barthlott, T. Blumenstock, U. Raffalski, A. Gómez-Peláez, M. Steinbacher, L. Ries, and A. M. De Frutos. The MUSICA IASI CH₄ and N₂O products and their comparison to HIPPO, GAW and NDACC FTIR references. *Atmospheric Measurement Techniques*, 11(7):4171–4215, jul 2018. ISSN 18678548. doi: 10.5194/amt-11-4171-2018.
- O. E. García, M. Schneider, B. Ertl, E. Sepúlveda, C. Borger, C. Diekmann, F. Hase, F. Khosrawi, A. Cansado, and M. Aullé. Monitorización de las concentraciones atmosféricas de metano y óxido nítrico a partir del Metop/IASI. *Revista de Teledetección*, 57(57):1, dec 2020. ISSN 1133-0953. doi: 10.4995/raet.2020.13290.
- D. M. Giles, A. Sinyuk, M. G. Sorokin, J. S. Schafer, A. Smirnov, I. Slutsker, T. F. Eck, B. N. Holben, J. R. Lewis, J. R. Campbell, E. J. Welton, S. V. Korkin, and A. I. Lyapustin. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database - Automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmospheric Measurement Techniques*, 12(1):169–209, jan 2019. ISSN 18678548. doi: 10.5194/amt-12-169-2019.
- C. Gong and E. Eltahir. Sources of moisture for rainfall in west Africa. *Water Resources Research*, 32(10):3115–3121, oct 1996. ISSN 00431397. doi: 10.1029/96WR01940.

- Y. González, M. Schneider, C. Dyroff, S. Rodríguez, E. Christner, O. E. García, E. Cuevas, J. J. Bustos, R. Ramos, C. Guirado-Fuentes, S. Barthlott, A. Wiegeler, and E. Sepúlveda. Detecting moisture transport pathways to the subtropical North Atlantic free troposphere using paired H₂O- δ D in situ measurements. *Atmospheric Chemistry and Physics*, 16(7):4251–4269, apr 2016. ISSN 16807324. doi: 10.5194/acp-16-4251-2016.
- I. E. Gordon, L. S. Rothman, C. Hill, R. V. Kochanov, Y. Tan, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. V. Chance, B. J. Drouin, J. M. Flaud, R. R. Gamache, J. T. Hodges, D. Jacquemart, V. I. Perevalov, A. Perrin, K. P. Shine, M. A. Smith, J. Tennyson, G. C. Toon, H. Tran, V. G. Tyuterev, A. Barbe, A. G. Császár, V. M. Devi, T. Furtenbacher, J. J. Harrison, J. M. Hartmann, A. Jolly, T. J. Johnson, T. Karman, I. Kleiner, A. A. Kyuberis, J. Loos, O. M. Lyulin, S. T. Massie, S. N. Mikhailenko, N. Moazzen-Ahmadi, H. S. Müller, O. V. Naumenko, A. V. Nikitin, O. L. Polyansky, M. Rey, M. Rotger, S. W. Sharpe, K. Sung, E. Starikova, S. A. Tashkun, J. V. Auwera, G. Wagner, J. Wilzewski, P. Wcisło, S. Yu, and E. J. Zak. The HITRAN2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 203:3–69, dec 2017. ISSN 00224073. doi: 10.1016/j.jqsrt.2017.06.038.
- P. Graf, H. Wernli, S. Pfahl, and H. Sodemann. A new interpretative framework for below-cloud effects on stable water isotopes in vapour and rain. *Atmospheric Chemistry and Physics*, 19(2):747–765, 2019. ISSN 16807324. doi: 10.5194/acp-19-747-2019.
- C. M. Grams, S. C. Jones, J. H. Marsham, D. J. Parker, J. M. Haywood, and V. Heuveline. The Atlantic inflow to the Saharan heat low: Observations and Modelling. *Quarterly Journal of the Royal Meteorological Society*, 136(SUPPL. 1):125–140, jan 2010. ISSN 00359009. doi: 10.1002/qj.429.
- G. Gu, R. F. Adler, G. J. Huffman, and S. Curtis. African easterly waves and their association with precipitation. *Journal of Geophysical Research: Atmospheres*, 109(4):4101, 2004. ISSN 01480227. doi: 10.1029/2003jd003967. URL <https://www.researchgate.net/publication/248803347>.
- R. Hagemann, G. Nief, and E. Roth. Absolute isotopic scale for deuterium analysis of natural waters. Absolute D/H ratio for SMOW. *Tellus*, 22(6):712–715, dec 1970. ISSN 00402826. doi: 10.1111/j.2153-3490.1970.tb00540.x.
- S. M. Hagos and K. H. Cook. Dynamics of the West African monsoon jump. *Journal of Climate*, 20(21):5264–5284, nov 2007. ISSN 08948755. doi: 10.1175/2007JCLI1533.1.
- N. M. Hall and P. Peyrillé. Dynamics of the West African monsoon. In *Journal De Physique. IV : JP*, volume 139, pages 81–99. EDP Sciences, dec 2006. ISBN 2868839436. doi: 10.1051/jp4:2006139007.
- F. Hase, J. W. Hannigan, M. T. Coffey, A. Goldman, M. Höpfner, N. B. Jones, C. P. Rinsland, and S. W. Wood. Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 87(1):25–52, 2004. ISSN 00224073. doi: 10.1016/j.jqsrt.2003.12.008.
- H. Herbin, D. Hurtmans, C. Clerbaux, L. Clarisse, and P. F. Coheur. H₂16O and HDO measurements with IASI/MetOp. *Atmospheric Chemistry and Physics*, 9(24):9433–9447, 2009. ISSN 16807324. doi: 10.5194/acp-9-9433-2009.
- J. Horita and D. J. Wesolowski. Liquid-vapor fractionation of oxygen and hydrogen isotopes of water from the freezing to the critical temperature. *Geochimica et Cosmochimica Acta*, 58(16):3425–3437, 1994. ISSN 00167037. doi: 10.1016/0016-7037(94)90096-5.
- R. A. Houze. Structure and Dynamics of a Tropical Squall–Line System. *Monthly Weather Review*, 105(12):1540–1567, dec 1977. ISSN 0027-0644. doi: 10.1175/1520-0493(1977)105<1540:sadoat>2.0.co;2.

- R. A. Houze, S. A. Rutledge, M. I. Biggerstaff, and B. F. Smull. Interpretation of Doppler weather radar displays of midlatitude mesoscale convective systems. *Bulletin - American Meteorological Society*, 70(6):608–619, jun 1989. ISSN 0003-0007. doi: 10.1175/1520-0477(1989)070<0608:IODWRD>2.0.CO;2. URL <https://journals.ametsoc.org/view/journals/bams/70/6/1520-0477{-}1989{-}070{-}0608{-}iodwrd{-}2{-}0{-}co{-}2.xml>.
- J. S. Hsieh and K. H. Cook. On the instability of the African easterly jet and the generation of African waves: Reversals of the potential vorticity gradient. *Journal of the Atmospheric Sciences*, 65(7):2130–2151, jul 2008. ISSN 00224928. doi: 10.1175/2007JAS2552.1.
- G. Huffman, D. Bolvin, E. Nelkin, and J. Tan. Integrated Multi-satellite Retrievals for GPM (IMERG) Technical Documentation. Technical report, NASA Goddard Centre, 2019. <https://gpm.nasa.gov/data/imerg>, accessed on 2021-02-15.
- G. J. Huffman, D. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, and P. Xie. Integrated Multi-satellite Retrievals for GPM (IMERG), version 4.4, 2014. <https://gpm.nasa.gov/data/directory>, accessed on 2021-02-07.
- M. Hulme. Climatic perspectives on Sahelian desiccation: 1973-1998. *Global Environmental Change*, 11(1):19–29, apr 2001. ISSN 09593780. doi: 10.1016/S0959-3780(00)00042-X.
- D. G. Johnson, K. W. Jucks, W. A. Traub, and K. V. Chance. Isotopic composition of stratospheric water vapor: Implications for transport. *Journal of Geophysical Research Atmospheres*, 106(D11):12219–12226, 2001. ISSN 01480227. doi: 10.1029/2000JD900764.
- J. Jouzel and L. Merlivat. Deuterium and oxygen 18 in precipitation: Modeling of the isotopic effects during snow formation. *Journal of Geophysical Research*, 89(D7):11749, 1984. ISSN 0148-0227. doi: 10.1029/JD089iD07p11749.
- A. Keppens, J. C. Lambert, J. Granville, G. Miles, R. Siddans, J. C. Van Peet, R. J. Van Der A, D. Hubert, T. Verhoelst, A. Delcloo, S. Godin-Beekmann, R. Kivi, R. Stübi, and C. Zehner. Round-robin evaluation of nadir ozone profile retrievals: Methodology and application to MetOp-A GOME-2. *Atmospheric Measurement Techniques*, 8(5): 2093–2120, may 2015. ISSN 18678548. doi: 10.5194/amt-8-2093-2015.
- F. Khosrawi, K. Toride, K. Yoshimura, C. J. Diekmann, B. Ertl, F. Hase, and M. Schneider. Can the assimilation of water isotopologue observation improve the quality of meteorological analysis fields? *in preparation*, 2021.
- A. Kniffka, P. Knippertz, and A. H. Fink. The role of low-level clouds in the West African monsoon system. *Atmospheric Chemistry and Physics*, 19(3):1623–1647, 2019. ISSN 16807324. doi: 10.5194/acp-19-1623-2019.
- P. Knippertz and A. H. Fink. Synoptic and dynamic aspects of an extreme springtime Saharan dust outbreak. *Quarterly Journal of the Royal Meteorological Society*, 132(617):1153–1177, apr 2006. ISSN 00359009. doi: 10.1256/qj.05.109.
- P. Knippertz and A. H. Fink. Prediction of dry-season precipitation in tropical West Africa and its relation to forcing from the extratropics. *Weather and Forecasting*, 24(4):1064–1084, aug 2009. ISSN 08828156. doi: 10.1175/2009WAF2222221.1.
- P. Knippertz and J. E. Martin. Tropical plumes and extreme precipitation in subtropical and tropical West Africa. *Quarterly Journal of the Royal Meteorological Society*, 131(610 B):2337–2365, jul 2005. ISSN 00359009. doi: 10.1256/qj.04.148.

- P. Knippertz, A. H. Fink, A. Deroubaix, E. Morris, F. Tocquer, M. J. Evans, C. Flamant, M. Gaetani, C. Lavaysse, C. Mari, J. H. Marsham, R. Meynadier, A. Affo-Dogo, T. Bahaga, F. Brosse, K. Deetz, R. Guebsi, I. Latifou, M. Maranan, P. D. Rosenberg, and A. Schlueter. A meteorological and chemical overview of the DACCIWA field campaign in West Africa in June-July 2016. *Atmospheric Chemistry and Physics*, 17(17):10893–10918, sep 2017. ISSN 16807324. doi: 10.5194/acp-17-10893-2017.
- B. L. Konecky, N. P. McKay, O. V. Churakova, L. Comas-Bru, E. P. Dassié, K. L. DeLong, G. M. Falster, M. J. Fischer, M. D. Jones, L. Jonkers, D. S. Kaufman, G. Leduc, S. R. Managave, B. Martrat, T. Opel, A. J. Orsi, J. W. Partin, H. R. Sayani, E. K. Thomas, D. M. Thompson, J. J. Tyler, N. J. Abram, A. R. Atwood, O. Cartapanis, J. L. Conroy, M. A. Curran, S. G. Dee, M. Deininger, D. V. Divine, Z. Kern, T. J. Porter, S. L. Stevenson, L. von Gunten, K. Braun, M. Carré, A. Incarbona, N. Kaushal, R. M. Klabe, H. R. Kolus, P. G. Mortyn, A. D. Moy, H. A. Roop, M. A. Sicre, and K. Yoshimura. The Iso2k database: A global compilation of paleo- $\delta^{18}\text{O}$ and $\delta^2\text{H}$ records to aid understanding of Common Era climate. *Earth System Science Data*, 12(3):2261–2288, sep 2020. ISSN 18663516. doi: 10.5194/essd-12-2261-2020.
- R. D. Koster, P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C. H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada. Regions of strong coupling between soil moisture and precipitation. *Science*, 305(5687):1138–1140, aug 2004. ISSN 00368075. doi: 10.1126/science.1100217.
- N. Kurita. Water isotopic variability in response to mesoscale convective system over the tropical ocean. *Journal of Geophysical Research Atmospheres*, 118(18):10,376–10,390, 2013. ISSN 21698996. doi: 10.1002/jgrd.50754.
- J. L. Lacour, C. Risi, L. Clarisse, S. Bony, D. Hurtmans, C. Clerbaux, and P. F. Coheur. Mid-tropospheric δd observations from IASI/MetOp at high spatial and temporal resolution. *Atmospheric Chemistry and Physics*, 12(22):10817–10832, 2012. ISSN 16807316. doi: 10.5194/acp-12-10817-2012.
- J. L. Lacour, C. Risi, J. Worden, C. Clerbaux, and P. F. Coheur. Importance of depth and intensity of convection on the isotopic composition of water vapor as seen from IASI and TES δD observations. *Earth and Planetary Science Letters*, 481:387–394, 2018. ISSN 0012821X. doi: 10.1016/j.epsl.2017.10.048.
- C. Lavaysse, C. Flamant, S. Janicot, D. J. Parker, J. P. Lafore, B. Sultan, and J. Pelon. Seasonal evolution of the West African heat low: A climatological perspective. *Climate Dynamics*, 33(2-3):313–330, apr 2009. ISSN 09307575. doi: 10.1007/s00382-009-0553-4.
- J. R. Lawrence, S. D. Gedzelman, D. Dexheimer, H. K. Cho, G. D. Carrie, R. Gasparini, C. R. Anderson, K. P. Bowman, and M. I. Biggerstaff. Stable isotopic composition of water vapor in the tropics. *Journal of Geophysical Research: Atmospheres*, 109(6), mar 2004. ISSN 01480227. doi: 10.1029/2003jd004046.
- J. E. Lee and I. Fung. "Amount effect" of water isotopes and quantitative analysis of post-condensation processes. *Hydrological Processes*, 22(1):1–8, jan 2008. ISSN 08856087. doi: 10.1002/hyp.6637.
- A. N. LeGrande and G. A. Schmidt. Global gridded data set of the oxygen isotopic composition in seawater. *Geophysical Research Letters*, 33(12):1–5, 2006. doi: 10.1029/2006GL026011.
- M. I. Lélé and P. J. Lamb. Variability of the Intertropical Front (ITF) and rainfall over the West African Sudan-Sahel zone. *Journal of Climate*, 23(14):3984–4004, jul 2010. ISSN 08948755. doi: 10.1175/2010JCLI3277.1.

- M. I. Lélé and L. M. Leslie. Intraseasonal variability of low-level moisture transport over West Africa. *Climate Dynamics*, 47(11):3575–3591, dec 2016. ISSN 14320894. doi: 10.1007/s00382-016-3334-x.
- S. Leroux and N. M. Hall. On the relationship between African easterly waves and the African easterly jet. *Journal of the Atmospheric Sciences*, 66(8):2303–2316, 2009. ISSN 00224928. doi: 10.1175/2009JAS2988.1.
- M. Majoube. Fractionnement en oxygène 18 et en deutérium entre l’eau et sa vapeur. *Journal de Chimie Physique*, 1971. ISSN 0021-7689. doi: 10.1051/jcp/1971681423.
- M. Maranan, A. H. Fink, P. Knippertz, S. D. Francis, A. B. Akpo, G. Jegede, and C. Yorke. Interactions between convection and a moist vortex associated with an extreme rainfall event over southern West Africa. *Monthly Weather Review*, 147(7):2309–2328, jul 2019. ISSN 15200493. doi: 10.1175/MWR-D-18-0396.1.
- D. R. Marsh, M. J. Mills, D. E. Kinnison, J. F. Lamarque, N. Calvo, and L. M. Polvani. Climate change from 1850 to 2005 simulated in CESM1(WACCM). *Journal of Climate*, 26(19):7372–7391, oct 2013. ISSN 08948755. doi: 10.1175/JCLI-D-12-00558.1.
- J. H. Marsham, N. S. Dixon, L. Garcia-Carreras, G. M. Lister, D. J. Parker, P. Knippertz, and C. E. Birch. The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophysical Research Letters*, 40(9):1843–1849, may 2013. ISSN 00948276. doi: 10.1002/grl.50347.
- I. R. Martínez and J. P. Chaboureaud. Precipitation and Mesoscale convective systems: Explicit versus parameterized convection over Northern Africa. *Monthly Weather Review*, 146(3):797–812, mar 2018. ISSN 15200493. doi: 10.1175/MWR-D-17-0202.1.
- K. Masuda, T. Takashima, and Y. Takayama. Emissivity of pure and sea waters for the model sea surface in the infrared window regions. *Remote Sensing of Environment*, 24(2):313–329, 1988. ISSN 00344257. doi: 10.1016/0034-4257(88)90032-6.
- V. Maurer, I. Bischoff-Gauß, N. Kalthoff, L. Gantner, R. Roca, and H. J. Panitz. Initiation of deep convection in the Sahel in a convection-permitting climate simulation for northern Africa. *Quarterly Journal of the Royal Meteorological Society*, 143(703):806–816, jan 2017. ISSN 1477870X. doi: 10.1002/qj.2966.
- A. Mekonnen, C. D. Thorncroft, and A. R. Aiyer. Analysis of convection and its association with African easterly waves. *Journal of Climate*, 19(20):5405–5421, oct 2006. ISSN 08948755. doi: 10.1175/JCLI3920.1.
- L. Merlivat and G. Nief. Fractionnement isotopique lors des changements d’état solide-vapeur et liquide-vapeur de l’eau à des températures inférieures à 0°C. *Tellus*, 19(1):122–127, 1967. ISSN 00402826. doi: 10.1111/j.2153-3490.1967.tb01465.x.
- R. Meynadier, O. Bock, S. Gervois, F. Guichard, J. L. Redelsperger, A. Agust-Panareda, and A. Beljaars. West African Monsoon water cycle: 2. Assessment of numerical weather prediction water budgets. *Journal of Geophysical Research Atmospheres*, 115(19):D19107, oct 2010. ISSN 01480227. doi: 10.1029/2010JD013919.
- A. K. Miltenberger, S. Pfahl, and H. Wernli. An online trajectory module (version 1.0) for the nonhydrostatic numerical weather prediction model COSMO. *Geoscientific Model Development*, 6(6):1989–2004, 2013. ISSN 1991959X. doi: 10.5194/gmd-6-1989-2013.

- C. Niang, A. M. Mancho, V. J. García-Garrido, E. Mohino, B. Rodriguez-Fonseca, and J. Curbelo. Transport pathways across the West African Monsoon as revealed by Lagrangian Coherent Structures. *Scientific Reports*, 10(1), dec 2020. ISSN 20452322. doi: 10.1038/s41598-020-69159-9.
- R. Nieto, L. Gimeno, and R. M. Trigo. A Lagrangian identification of major sources of Sahel moisture. *Geophysical Research Letters*, 33(18):n/a–n/a, sep 2006. ISSN 00948276. doi: 10.1029/2006GL027232.
- E. G. Njoku. *Encyclopedia of Remote Sensing*. Encyclopedia of Earth Sciences Series. Springer New York, New York, NY, 2014. ISBN 978-0-387-36698-2. doi: 10.1007/978-0-387-36699-9.
- B. Nlend, H. Celle-Jeanton, C. Risi, B. Pohl, F. Huneau, S. Ngo Boum-Nkot, G. Seze, P. Roucou, P. Camberlin, J. Etame, and B. Ketchemen-Tandia. Identification of processes that control the stable isotope composition of rainwater in the humid tropical West-Central Africa. *Journal of Hydrology*, 584:124650, may 2020. ISSN 00221694. doi: 10.1016/j.jhydrol.2020.124650.
- D. Noone. Pairing measurements of the water vapor isotope ratio with humidity to deduce atmospheric moistening and dehydration in the tropical midtroposphere. *Journal of Climate*, 25(13):4476–4494, jul 2012. ISSN 08948755. doi: 10.1175/JCLI-D-11-00582.1.
- D. Noone, J. Galewsky, Z. D. Sharp, J. Worden, J. Barnes, D. Baer, A. Bailey, D. P. Brown, L. Christensen, E. Crosson, F. Dong, J. V. Hurley, L. R. Johnson, M. Strong, D. Toohey, A. Van Pelt, and J. S. Wright. Properties of air mass mixing and humidity in the subtropics from measurements of the D/H isotope ratio of water vapor at the Mauna Loa Observatory. *Journal of Geophysical Research Atmospheres*, 116(22), 2011. ISSN 01480227. doi: 10.1029/2011JD015773.
- G. Pante and P. Knippertz. Resolving Sahelian thunderstorms improves mid-latitude weather forecasts. *Nature Communications*, 10(1):1–9, dec 2019. ISSN 20411723. doi: 10.1038/s41467-019-11081-4.
- D. J. Parker, C. D. Thorncroft, R. R. Burton, and A. Diongue-Niang. Analysis of the African easterly jet, using aircraft observations from the JET2000 experiment. *Quarterly Journal of the Royal Meteorological Society*, 131(608):1461–1482, apr 2005. ISSN 00359009. doi: 10.1256/qj.03.189.
- D. J. Parker, A. Fink, S. Janicot, J. B. Ngamini, M. Douglas, E. Afiesimama, A. Agusti-Panareda, A. Beljaars, F. Dide, A. Diedhiou, T. Lebel, J. Polcher, J. L. Redelsperger, C. Thorncroft, and G. A. Wilson. The Amma radiosonde program and its implications for the future of atmospheric monitoring over Africa. *Bulletin of the American Meteorological Society*, 89(7):1015–1027, 2008. ISSN 00030007. doi: 10.1175/2008BAMS2436.1.
- S. Pfahl and H. Wernli. Air parcel trajectory analysis of stable isotopes in water vapor in the eastern Mediterranean. *Journal of Geophysical Research Atmospheres*, 113(20), oct 2008. ISSN 01480227. doi: 10.1029/2008JD009839.
- S. Pfahl, H. Wernli, and K. Yoshimura. The isotopic composition of precipitation from a winter storm—a case study with the limited-area model COSMOiso. *Atmospheric Chemistry and Physics*, 12(3):1629–1648, 2012. ISSN 16807316. doi: 10.5194/acp-12-1629-2012.
- D. L. Phillips. A Technique for the Numerical Solution of Certain Integral Equations of the First Kind. *Journal of the ACM (JACM)*, 9(1):84–97, jan 1962. ISSN 1557735X. doi: 10.1145/321105.321114.
- R. J. Purser and H. L. Huang. Estimating effective data density in a satellite retrieval or an objective analysis. *Journal of Applied Meteorology*, 32(6):1092–1107, jun 1993. ISSN 08948763. doi: 10.1175/1520-0450(1993)032<1092:EEDDIA>2.0.CO;2.

- L. Rayleigh. On the distillation of binary mixtures. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 4(23):521–537, 1902. ISSN 1941-5982. doi: 10.1080/14786440209462876.
- J. L. Redelsperger, C. D. Thorncroft, A. Diedhiou, T. Lebel, D. J. Parker, and J. Polcher. African Monsoon Multidisciplinary Analysis: An international research project and field campaign. *Bulletin of the American Meteorological Society*, 87(12):1739–1746, 2006. ISSN 00030007. doi: 10.1175/BAMS-87-12-1739.
- D. Reinert, F. Prill, G. Zängl, D. Rieger, J. Schröter, J. Förstner, S. Werchner, M. Weimer, R. Ruhnke, and B. Vogel. Working with the ICON Model, Tech. rep. Technical report, Deutscher Wetterdienst, Offenbach, 2019. https://code.mpimet.mpg.de/attachments/download/19568/ICON_tutorial_2019.pdf.
- A. Renfrew, R. S. Pickart, K. Våge, G. W. Moore, T. J. Bracegirdle, A. D. Elvidge, E. Jeansson, T. Lachlan-Cope, L. T. McRaven, L. Papritz, J. Reuder, H. Sodemann, A. Terpstra, S. Waterman, H. Valdimarsson, A. Weiss, M. Almansi, F. Bahr, A. Brakstad, C. Barrell, J. K. Brooke, B. J. Brooks, I. M. Brooks, M. E. Brooks, E. M. Bruvik, C. Duscha, I. Fer, H. M. Golid, M. Hallerstig, I. Hessevik, J. Huang, L. Houghton, S. Jónsson, M. Jonassen, K. Jackson, K. Kvalsund, E. W. Kolstad, K. Konstali, J. Kristiansen, R. Ladkin, P. Lin, A. MacRander, A. Mitchell, H. Olafsson, A. Pacini, C. Payne, B. Palmason, M. D. Pérez-Hernández, A. K. Peterson, G. N. Petersen, M. N. Pisareva, J. O. Pope, A. Seidl, S. Semper, D. Sergeev, S. Skjelsvik, H. Søiland, D. Smith, M. A. Spall, T. Spengler, A. Touzeau, G. Tupper, Y. Weng, K. D. Williams, X. Yang, and S. Zhou. The Iceland Greenland seas project. *Bulletin of the American Meteorological Society*, 100(9):1795–1817, sep 2019. ISSN 00030007. doi: 10.1175/BAMS-D-18-0217.1.
- D. Rieger, M. Bangert, I. Bischoff-Gauss, J. Förstner, K. Lundgren, D. Reinert, J. Schröter, H. Vogel, G. Zängl, R. Ruhnke, and B. Vogel. ICON-ART 1.0 - A new online-coupled model system from the global to regional scale. *Geoscientific Model Development*, 8(6):1659–1676, 2015. ISSN 19919603. doi: 10.5194/gmd-8-1659-2015.
- M. Rindsberger, M. Magaritz, I. Carmi, and D. Gilad. The relation between air mass trajectories and the water isotope composition of rain in the Mediterranean Sea area. *Geophysical Research Letters*, 10(1):43–46, jan 1983. ISSN 19448007. doi: 10.1029/GL010i001p00043.
- C. Risi, S. Bony, and F. Vimeux. Influence of convective processes on the isotopic composition ($\delta^{18}\text{O}$ and δD) of precipitation and water vapor in the tropics: 2. Physical interpretation of the amount effect. *Journal of Geophysical Research Atmospheres*, 113(19), oct 2008a. ISSN 01480227. doi: 10.1029/2008JD009943.
- C. Risi, S. Bony, F. Vimeux, L. Descroix, B. Ibrahim, E. Lebreton, I. Mamadou, and B. Sultan. What controls the isotopic composition of the African monsoon precipitation? Insights from event-based precipitation collected during the 2006 AMMA field campaign. *Geophysical Research Letters*, 35(24):1–6, 2008b. ISSN 00948276. doi: 10.1029/2008GL035920.
- C. Risi, S. Bony, F. Vimeux, M. Chongd, and L. Descroix. Evolution of the stable water isotopic composition of the rain sampled along sahelian squall lines. *Quarterly Journal of the Royal Meteorological Society*, 136(SUPPL. 1): 227–242, jan 2010a. ISSN 00359009. doi: 10.1002/qj.485.
- C. Risi, S. Bony, F. Vimeux, C. Frankenberg, D. Noone, and J. Worden. Understanding the Sahelian water budget through the isotopic composition of water vapor and precipitation. *Journal of Geophysical Research Atmospheres*, 115(24):1–23, 2010b. ISSN 01480227. doi: 10.1029/2010JD014690.
- C. Risi, S. Bony, F. Vimeux, and J. Jouzel. Water-stable isotopes in the LMDZ4 general circulation model: Model evaluation for present-day and past climates and applications to climatic interpretations of tropical isotopic

- records. *Journal of Geophysical Research Atmospheres*, 115(12):1–27, 2010c. ISSN 01480227. doi: 10.1029/2009JD013255.
- C. Risi, D. Noone, J. Worden, C. Frankenberg, G. Stiller, M. Kiefer, B. Funke, K. Walker, P. Bernath, M. Schneider, D. Wunch, V. Sherlock, N. Deutscher, D. Griffith, P. O. Wennberg, K. Strong, D. Smale, E. Mahieu, S. Barthlott, F. Hase, O. García, J. Notholt, T. Warneke, G. Toon, D. Sayres, S. Bony, J. Lee, D. Brown, R. Uemura, and C. Sturm. Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopologues: 1. Comparison between models and observations. *Journal of Geophysical Research Atmospheres*, 117(5):1–26, 2012. ISSN 01480227. doi: 10.1029/2011JD016621.
- R. Roca, J. P. Lafore, C. Piriou, and J. L. Redelsperger. Extratropical dry-air intrusions into the West African monsoon midtroposphere: An important factor for the convective activity over the Sahel. *Journal of the Atmospheric Sciences*, 62(2):390–407, feb 2005. ISSN 00224928. doi: 10.1175/JAS-3366.1.
- C. D. Rodgers. Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics–Vol. 2., *Series on Atmospheric, Oceanic and Planetary Physics–Vol. 2*, 2:256, 2000.
- C. D. Rodgers and B. J. Connor. Intercomparison of remote sounding instruments. *Journal of Geophysical Research: Atmospheres*, 108(3):4116, feb 2003. ISSN 01480227. doi: 10.1029/2002jd002299.
- S. Rodríguez, A. Alastuey, S. Alonso-Pérez, X. Querol, E. Cuevas, J. Abreu-Afonso, M. Viana, N. Pérez, M. Pandolfi, and J. De La Rosa. Transport of desert dust mixed with North African industrial pollutants in the subtropical Saharan Air Layer. *Atmospheric Chemistry and Physics*, 11(13):6663–6685, 2011. ISSN 16807316. doi: 10.5194/acp-11-6663-2011.
- R. Roehrig, D. Bouniol, F. Guichard, F. Hourdin, and J. L. Redelsperger. The present and future of the west african monsoon: A process-oriented assessment of CMIP5 simulations along the AMMA transect. *Journal of Climate*, 26(17):6471–6505, sep 2013. ISSN 08948755. doi: 10.1175/JCLI-D-12-00505.1.
- N. V. Rokotyan, V. I. Zakharov, K. G. Griбанov, M. Schneider, F. M. Bréon, J. Jouzel, R. Imasu, M. Werner, M. Butzin, C. Petri, T. Warneke, and J. Notholt. A posteriori calculation of $\delta^{18}\text{O}$ and δD in atmospheric water vapour from ground-based near-infrared FTIR retrievals of H_2^{16}O , H_2^{18}O , and HD^{16}O . *Atmospheric Measurement Techniques*, 7(8):2567–2580, aug 2014. ISSN 18678548. doi: 10.5194/amt-7-2567-2014.
- D. P. Rowell. The impact of Mediterranean SSTs on the Sahelian rainfall season. *Journal of Climate*, 16(5):849–862, mar 2003. ISSN 08948755. doi: 10.1175/1520-0442(2003)016<0849:TIOMSO>2.0.CO;2.
- K. Rozanski, L. Araguães-Araguães, and R. Gonfiantini. Relation between long-term trends of oxygen-18 isotope composition of precipitation and climate. *Science*, 258(5084):981–985, 1992. ISSN 00368075. doi: 10.1126/science.258.5084.981.
- V. Salamalikis, A. A. Argiriou, and E. Dotsika. Stable isotopic composition of atmospheric water vapor in Patras, Greece: A concentration weighted trajectory approach. *Atmospheric Research*, 152:93–104, jan 2015. ISSN 01698095. doi: 10.1016/j.atmosres.2014.02.021.
- A. A. Salih, Q. Zhang, and M. Tjernström. Lagrangian tracing of Sahelian Sudan moisture sources. *Journal of Geophysical Research*, 120(14):6793–6808, jul 2015. ISSN 21562202. doi: 10.1002/2015JD023238.

- R. A. Scheepmaker, C. Frankenberg, A. Galli, A. Butz, H. Schrijver, N. M. Deutscher, D. Wunch, T. Warneke, S. Fally, and I. Aben. Improved water vapour spectroscopy in the 4174–4300 cm⁻¹ region and its impact on SCIAMACHY HDO/H₂O measurements. *Atmospheric Measurement Techniques*, 6(4):879–894, 2013. ISSN 18671381. doi: 10.5194/amt-6-879-2013.
- S. Schemm, A. Nummelina, N. G. Kvamstø, and Ø. Breivik. The ocean version of the Lagrangian analysis tool LAGRANTO. *Journal of Atmospheric and Oceanic Technology*, 34(8):1723–1741, aug 2017. ISSN 15200426. doi: 10.1175/JTECH-D-16-0198.1.
- A. Schneider, T. Borsdorff, J. Aan De Brugh, H. Hu, and J. Landgraf. A full-mission data set of H₂O and HDO columns from SCIAMACHY 2.3 μm reflectance measurements. *Atmospheric Measurement Techniques*, 11(6):3339–3350, jun 2018. ISSN 18678548. doi: 10.5194/amt-11-3339-2018.
- A. Schneider, T. Borsdorff, J. aan de Brugh, F. Aemisegger, D. G. Feist, R. Kivi, F. Hase, M. Schneider, and J. Landgraf. First data set of H₂O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI). *Atmospheric Measurement Techniques*, 13(1):85–100, jan 2020. ISSN 1867-8548. doi: 10.5194/amt-13-85-2020.
- M. Schneider and F. Hase. Improving spectroscopic line parameters by means of atmospheric spectra: Theory and example for water vapor and solar absorption spectra. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(17):1825–1839, nov 2009. ISSN 00224073. doi: 10.1016/j.jqsrt.2009.04.011.
- M. Schneider and F. Hase. Optimal estimation of tropospheric H₂O and δD with IASI/METOP. *Atmospheric Chemistry and Physics*, 11(21):11207–11220, 2011. ISSN 16807316. doi: 10.5194/acp-11-11207-2011.
- M. Schneider, B. Ertl, and C. Diekmann. MUSICA IASI full retrieval product extended output (processing version 3.2.1), jan . dataset, Institute of Meteorology and Climate Research, Atmospheric Trace Gases and Remote Sensing (IMK-ASF), Karlsruhe Institute of Technology (KIT), doi: 10.35097/412.
- M. Schneider, F. Hase, and T. Blumenstock. Ground-based remote sensing of HDO/H₂O ratio profiles: introduction and validation of an innovative retrieval approach. *Atmospheric Chemistry and Physics Discussions*, 6:5269–5327, 2006. ISSN 16807375. doi: 10.5194/acpd-6-5269-2006.
- M. Schneider, S. Barthlott, F. Hase, Y. González, K. Yoshimura, O. E. García, E. Sepúlveda, A. Gomez-Pelaez, M. Gisi, R. Kohlhepp, S. Dohe, T. Blumenstock, A. Wiegeler, E. Christner, K. Strong, D. Weaver, M. Palm, N. M. Deutscher, T. Warneke, J. Notholt, B. Lejeune, P. Demoulin, N. Jones, D. W. Griffith, D. Smale, and J. Robinson. Ground-based remote sensing of tropospheric water vapour isotopologues within the project MUSICA. *Atmospheric Measurement Techniques*, 5(12):3007–3027, 2012. ISSN 18671381. doi: 10.5194/amt-5-3007-2012.
- M. Schneider, Y. González, C. Dyroff, E. Christner, A. Wiegeler, S. Barthlott, O. E. García, E. Sepúlveda, F. Hase, J. Andrey, T. Blumenstock, C. Guirado, R. Ramos, and S. Rodríguez. Empirical validation and proof of added value of MUSICA’s tropospheric δD remote sensing products. *Atmospheric Measurement Techniques*, 8(1):483–503, jan 2015. ISSN 18678548. doi: 10.5194/amt-8-483-2015.
- M. Schneider, A. Wiegeler, S. Barthlott, Y. González, E. Christner, C. Dyroff, E. O. García, F. Hase, T. Blumenstock, E. Sepúlveda, G. Mengistu Tsidu, S. Takele Kenea, S. Rodríguez, and J. Andrey. Accomplishments of the MUSICA project to provide accurate, long-term, global and high-resolution observations of tropospheric H₂O, δD pairs - A review. *Atmospheric Measurement Techniques*, 9(7):2845–2875, 2016. ISSN 18678548. doi: 10.5194/amt-9-2845-2016.

- M. Schneider, C. Borger, A. Wiegele, F. Hase, O. E. García, E. Sepúlveda, and M. Werner. MUSICA MetOp/IASI H₂O, δ D pair retrieval simulations for validating tropospheric moisture pathways in atmospheric models. *Atmospheric Measurement Techniques*, 10(2):507–525, 2017. ISSN 18678548. doi: 10.5194/amt-10-507-2017.
- M. Schneider, B. Ertl, and C. Diekmann. MUSICA IASI full retrieval product standard output (processing version 3.2.1), 2021a. dataset, Institute of Meteorology and Climate Research, Atmospheric Trace Gases and Remote Sensing (IMK-ASF), Karlsruhe Institute of Technology (KIT), doi: 10.35097/408.
- M. Schneider, B. Ertl, C. Diekmann, F. Khosrawi, A. Röhling, F. Hase, D. Dubravica, O. García, E. Sepúlveda, T. Borsdorff, J. Landgraf, A. Lorente, H. Chen, R. Kivi, T. Laemmle, M. Ramonet, C. Crevoisier, J. Pernin, M. Steinbacher, F. Meinhardt, N. Deutscher, D. Griffith, V. Velasco, and D. Pollard. Synergetic use of IASI and TROPOMI space borne sensors for generating a tropospheric methane profile product. *Atmospheric Measurement Techniques Discussions*, pages 1–37, 2021b. ISSN 1867-1381. doi: 10.5194/amt-2021-31.
- M. Schneider, B. Ertl, F. Khosrawi, A. Weber, F. Hase, O. E. Garcia, E. Sepulveda, and D. Kinnison. Design and description of the MUSICA IASI full retrieval product. *submitted to Earth System Science Data*, 2021c.
- J. Schröter, D. Rieger, C. Stassen, H. Vogel, M. Weimer, S. Werchner, J. Förstner, F. Prill, D. Reinert, M. Giorgetta, R. Ruhnke, B. Vogel, and P. Braesicke. ICON-ART 2.1: A flexible tracer framework and its application for composition studies in numerical weather forecasting and climate simulations. *Geoscientific Model Development*, 11(10):4043–4068, oct 2018. ISSN 19919603. doi: 10.5194/gmd-11-4043-2018.
- S. W. Seemann, E. E. Borbas, R. O. Knuteson, G. R. Stephenson, and H. L. Huang. Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements. *Journal of Applied Meteorology and Climatology*, 47(1):108–123, 2008. ISSN 15588424. doi: 10.1175/2007JAMC1590.1.
- C. B. Skinner and N. S. Diffenbaugh. The contribution of African easterly waves to monsoon precipitation in the CMIP3 ensemble. *Journal of Geophysical Research: Atmospheres*, 118(9):3590–3609, 2013. doi: 10.1002/jgrd.50363.
- H. Sodemann. Beyond turnover time: Constraining the lifetime distribution of water vapor from simple and complex approaches. *Journal of the Atmospheric Sciences*, 77(2):413–433, 2020. ISSN 15200469. doi: 10.1175/JAS-D-18-0336.1.
- H. Sodemann, C. Schwierz, and H. Wernli. Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation influence. *Journal of Geophysical Research Atmospheres*, 113(3), feb 2008. ISSN 01480227. doi: 10.1029/2007JD008503.
- K. Soderberg, S. P. Good, M. O’connor, L. Wang, K. Ryan, and K. K. Caylor. Using atmospheric trajectories to model the isotopic composition of rainfall in central Kenya. *Ecosphere*, 4(3):1–18, mar 2013. ISSN 21508925. doi: 10.1890/ES12-00160.1.
- R. W. Spencer and W. D. Braswell. How Dry is the Tropical Free Troposphere? Implications for Global Warming Theory. *Bulletin of the American Meteorological Society*, 78(6):1097–1106, jun 1997. ISSN 00030007. doi: 10.1175/1520-0477(1997)078<1097:HDITTF>2.0.CO;2.
- M. Sprenger and H. Wernli. The LAGRANTO Lagrangian analysis tool - Version 2.0. *Geoscientific Model Development*, 8(8):2569–2586, aug 2015. ISSN 19919603. doi: 10.5194/gmd-8-2569-2015.

- T. Steck. Methods for determining regularization for atmospheric retrieval problems. *Applied Optics*, 41(9):1788, mar 2002. ISSN 0003-6935. doi: 10.1364/AO.41.001788.
- J. Steppeler, G. Doms, U. Scattler, H. W. Bitzer, A. Gassmann, U. Damrath, and G. Gregoric. Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorology and Atmospheric Physics*, 82(1-4):75–96, jan 2003. ISSN 01777971. doi: 10.1007/s00703-001-0592-9.
- M. K. Stewart. Stable isotope fractionation due to evaporation and isotopic exchange of falling waterdrops: Applications to atmospheric processes and evaporation of lakes. *Journal of Geophysical Research*, 80(9):1133, 1975. ISSN 0148-0227. doi: 10.1029/JC080i009p01133.
- A. Stohl and P. James. A Lagrangian analysis of the atmospheric branch of the global water cycle: Part 1: Method description, validation, and demonstration for the August 2002 flooding in central Europe. *Journal of Hydrometeorology*, 5(4):656–678, aug 2004. ISSN 1525755X. doi: 10.1175/1525-7541(2004)005<0656:ALAOTA>2.0.CO;2.
- K. Sturm, G. Hoffmann, B. Langmann, and W. Stichler. Simulation of $\delta^{18}\text{O}$ in precipitation by the regional circulation model REMO iso. *Hydrological Processes*, 19(17):3425–3444, nov 2005. ISSN 08856087. doi: 10.1002/hyp.5979.
- B. Sultan and S. Janicot. The West African monsoon dynamics. Part II: The "preonset" and "onset" of the summer monsoon. *Journal of Climate*, 16(21):3407–3427, nov 2003. ISSN 08948755. doi: 10.1175/1520-0442(2003)016<3407:TWAMDP>2.0.CO;2.
- B. Sultan, C. Baron, M. Dingkuhn, B. Sarr, and S. Janicot. Agricultural impacts of large-scale variability of the West African monsoon. *Agricultural and Forest Meteorology*, 128(1-2):93–110, jan 2005. ISSN 01681923. doi: 10.1016/j.agrformet.2004.08.005.
- A. Sy, B. Diop, J. Van Baelen, C. Duroure, Y. Gour, and J. L. Baray. Upper tropospheric water vapor transport from Indian to Sahelian Regions. *Atmosphere*, 9(10):403, oct 2018. ISSN 20734433. doi: 10.3390/atmos9100403.
- S. Terzer, L. I. Wassenaar, L. J. Araguás-Araguás, and P. K. Aggarwal. Global isoscapes for $\delta^{18}\text{O}$ and $\delta^2\text{H}$ in precipitation: Improved prediction using regionalized climatic regression models. *Hydrology and Earth System Sciences*, 17(11):4713–4728, nov 2013. ISSN 10275606. doi: 10.5194/hess-17-4713-2013.
- T. Tharammal, G. Bala, and D. Noone. Impact of deep convection on the isotopic amount effect in tropical precipitation. *Journal of Geophysical Research*, 122(3):1505–1523, 2017. ISSN 21562202. doi: 10.1002/2016JD025555.
- C. D. Thorncroft and M. Blackburn. Maintenance of the African easterly jet. *Quarterly Journal of the Royal Meteorological Society*, 125(555):763–786, apr 1999. ISSN 00359009. doi: 10.1002/qj.49712555502.
- C. D. Thorncroft and B. J. Hoskins. An idealized study of African easterly waves. II: A nonlinear view. *Quarterly Journal of the Royal Meteorological Society*, 120(518):983–1015, jul 1994. ISSN 1477870X. doi: 10.1002/qj.49712051810.
- C. D. Thorncroft, N. M. Hall, and G. N. Kiladis. Three-dimensional structure and dynamics of African easterly waves. Part III: Genesis. *Journal of the Atmospheric Sciences*, 65(11):3596–3607, nov 2008. ISSN 00224928. doi: 10.1175/2008JAS2575.1.
- C. D. Thorncroft, H. Nguyen, C. Zhang, and P. Peyrille. Annual cycle of the West African monsoon: Regional circulations and associated water vapour transport. *Quarterly Journal of the Royal Meteorological Society*, 137(654):129–147, jan 2011. ISSN 00359009. doi: 10.1002/qj.728.

- A. Tikhonov. On the solution of improperly posed problems and the method of regularization. *Dokl. Acad. Nauk SSSR*, 151:501, 1963.
- K. Toride, K. Yoshimura, M. Tada, C. Diekmann, B. Ertl, F. Khosrawi, and M. Schneider. Potential of Mid-tropospheric Water Vapor Isotopes to Improve Large-Scale Circulation and Weather Predictability. *Geophysical Research Letters*, 48(5):e2020GL091698, mar 2021. ISSN 0094-8276. doi: 10.1029/2020GL091698. URL <https://onlinelibrary.wiley.com/doi/10.1029/2020GL091698>.
- R. A. Toth. HDO and D2O low pressure, long path spectra in the 600-3100 cm⁻¹ region II. D2O line positions and strengths. *Journal of Molecular Spectroscopy*, 195(1):98–122, may 1999. ISSN 00222852. doi: 10.1006/jmsp.1999.7815.
- G. Tremoy, F. Vimeux, S. Mayaki, I. Souley, O. Cattani, C. Risi, G. Favreau, and M. Oi. A 1-year long $\delta^{18}\text{O}$ record of water vapor in Niamey (Niger) reveals insightful atmospheric processes at different timescales. *Geophysical Research Letters*, 39(8):L08805, apr 2012. ISSN 00948276. doi: 10.1029/2012GL051298.
- G. Tremoy, F. Vimeux, S. Soumana, I. Souley, C. Risi, G. Favreau, and M. Oi. Clustering mesoscale convective systems with laser-based water vapor $\delta^{18}\text{O}$ monitoring in niamey (Niger). *Journal of Geophysical Research*, 119(9):5079–5103, may 2014. ISSN 21562202. doi: 10.1002/2013JD020968.
- E. Tyrlis and J. Lelieveld. Climatology and dynamics of the summer Etesian winds over the eastern Mediterranean. *Journal of the Atmospheric Sciences*, 70(11):3374–3396, nov 2013. ISSN 00224928. doi: 10.1175/JAS-D-13-035.1.
- H. C. Urey. The thermodynamic properties of isotopic substances. Livensidge lecture, delivered before the Chemical Society in the Royal Institution on December 18th, 1946. *Journal of the Chemical Society (Resumed)*, pages 562–581, 1947. ISSN 03681769. doi: 10.1039/jr9470000562.
- J. Vergara-Temprado, N. Ban, D. Panosetti, L. Schlemmer, and C. Schär. Climate models permit convection at much coarser resolutions than previously considered. *Journal of Climate*, 33(5):1915–1933, mar 2020. ISSN 08948755. doi: 10.1175/JCLI-D-19-0286.1.
- P. Vogel, P. Knippertz, A. H. Fink, A. Schlueter, and T. Gneiting. Skill of global raw and postprocessed ensemble predictions of rainfall over Northern Tropical Africa. *Weather and Forecasting*, 33(2):369–388, apr 2018. ISSN 15200434. doi: 10.1175/WAF-D-17-0127.1.
- T. von Clarmann, D. A. Degenstein, N. J. Livesey, S. Bender, A. Braverman, A. Butz, S. Compornolle, R. Damadeo, S. Dueck, P. Eriksson, B. Funke, M. C. Johnson, Y. Kasai, A. Keppens, A. Kleinert, N. A. Kramarova, A. Laeng, B. Langerock, V. H. Payne, A. Rozanov, T. O. Sato, M. Schneider, P. Sheese, V. Sofieva, G. P. Stiller, C. von Savigny, and D. Zawada. Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature. *Atmospheric Measurement Techniques*, 13(8):4393–4436, aug 2020. ISSN 18678548. doi: 10.5194/amt-13-4393-2020.
- K. Y. Wang and C. H. Liu. Transport of the Saharan dust air plumes over the tropical North Atlantic from FORMOSAT-3/COSMIC observation. *Atmospheric Pollution Research*, 5(3):539–553, jul 2014. ISSN 13091042. doi: 10.5094/APR.2014.063.
- A. Weber. Storage-Efficient Analysis of Spatio-Temporal Data with Application to Climate Research. *Zenodo*, aug 2019. doi: 10.5281/ZENODO.3360021. URL <https://zenodo.org/record/3360021>.

- Z. Wei, X. Lee, F. Aemisegger, M. Benetti, M. Berkelhammer, M. Casado, K. Caylor, E. Christner, C. Dyroff, O. García, Y. González, T. Griffis, N. Kurita, J. Liang, M. C. Liang, G. Lin, D. Noone, K. Gribanov, N. C. Munksgaard, M. Schneider, F. Ritter, H. C. Steen-Larsen, C. Vallet-Coulomb, X. Wen, J. S. Wright, W. Xiao, and K. Yoshimura. A global database of water vapor isotopes measured with high temporal resolution infrared laser spectroscopy. *Scientific Data*, 6(1):1–15, jan 2019. ISSN 20524463. doi: 10.1038/sdata.2018.302.
- M. Weimer, J. Buchmüller, L. Hoffmann, O. Kirner, B. Luo, R. Ruhnke, M. Steiner, I. Tritscher, and P. Braesicke. Mountain-wave induced polar stratospheric clouds and their representation in the global chemistry model ICON-ART. *Atmospheric Chemistry and Physics Discussions*, pages 1–42, 2020. ISSN 1680-7316. doi: 10.5194/acp-2020-1156.
- M. Werner, P. M. Langebroek, T. Carlsen, M. Herold, and G. Lohmann. Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale. *Journal of Geophysical Research*, 116(D15):D15109, aug 2011. ISSN 0148-0227. doi: 10.1029/2011JD015681.
- H. Wernli and H. C. Davies. A lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Quarterly Journal of the Royal Meteorological Society*, 123(538):467–489, jan 1997. ISSN 00359009. doi: 10.1002/qj.49712353811.
- A. Wiegeler, M. Schneider, F. Hase, S. Barthlott, O. E. Garcia, E. Sepúlveda, Y. González, T. Blumenstock, U. Raffalski, M. Gisi, and R. Kohlhepp. The MUSICA MetOp/IASI H₂O and δd products: Characterisation and long-term comparison to NDACC/FTIR data. *Atmospheric Measurement Techniques*, 7(8):2719–2732, aug 2014. ISSN 18678548. doi: 10.5194/amt-7-2719-2014.
- J. Worden, K. Bowman, D. Noone, R. Beer, S. Clough, A. Eldering, B. Fisher, A. Goldman, M. Gunson, R. Herman, S. S. Kulawik, M. Lampel, M. Luo, G. Osterman, C. Rinsland, C. Rodgers, S. Sander, M. Shephard, and H. Worden. Tropospheric Emission Spectrometer observations of the tropospheric HDO/H₂O ratio: Estimation approach and characterization. *Journal of Geophysical Research*, 111(D16):D16309, 2006. ISSN 0148-0227. doi: 10.1029/2005JD006606.
- J. Worden, D. Noone, K. Bowman, R. Beer, A. Eldering, B. Fisher, M. Gunson, A. Goldman, R. Herman, S. S. Kulawik, M. Lampel, G. Osterman, C. Rinsland, C. Rodgers, S. Sander, M. Shephard, C. R. Webster, and H. Worden. Importance of rain evaporation and continental convection in the tropical water cycle. *Nature*, 445(7127):528–532, feb 2007. ISSN 14764687. doi: 10.1038/nature05508.
- J. Worden, S. Kulawik, C. Frankenberg, V. Payne, K. Bowman, K. Cady-Peirara, K. Wecht, J. E. Lee, and D. Noone. Profiles of CH₄, HDO, H₂O, and N₂O with improved lower tropospheric vertical resolution from Aura TES radiances. *Atmospheric Measurement Techniques*, 5(2):397–411, feb 2012. ISSN 18671381. doi: 10.5194/amt-5-397-2012.
- J. R. Worden, S. S. Kulawik, D. Fu, V. H. Payne, A. E. Lipton, I. Polonsky, Y. He, K. Cady-Pereira, J. L. Moncet, R. L. Herman, F. W. Irion, and K. W. Bowman. Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor. *Atmospheric Measurement Techniques*, 12(4):2331–2339, 2019. ISSN 18678548. doi: 10.5194/amt-12-2331-2019.
- H. Yang and R. T. Pierrehumbert. Production of dry air by isentropic mixing. *Journal of the Atmospheric Sciences*, 51(23):3437–3454, dec 1994. ISSN 00224928. doi: 10.1175/1520-0469(1994)051<3437:PODABI>2.0.CO;2.

- K. Yoshimura. Stable Water Isotopes in Climatology, Meteorology, and Hydrology: A Review. *Journal of the Meteorological Society of Japan. Ser. II*, 93(5):513–533, 2015. ISSN 0026-1165. doi: 10.2151/jmsj.2015-036.
- K. Yoshimura, M. Kanamitsu, D. Noone, and T. Oki. Historical isotope simulation using Reanalysis atmospheric data. *Journal of Geophysical Research Atmospheres*, 113(19):1–15, 2008. ISSN 01480227. doi: 10.1029/2008JD010074.
- K. Yoshimura, M. Kanamitsu, and M. Dettinger. Regional downscaling for stable water isotopes: A case study of an atmospheric river event. *Journal of Geophysical Research Atmospheres*, 115(18), 2010. ISSN 01480227. doi: 10.1029/2010JD014032.
- K. Yoshimura, C. Frankenberg, J. Lee, M. Kanamitsu, J. Worden, and T. Röckmann. Comparison of an isotopic atmospheric general circulation model with new quasi-global satellite measurements of water vapor isotopologues. *Journal of Geophysical Research Atmospheres*, 116(19):1–15, 2011. ISSN 01480227. doi: 10.1029/2011JD016035.
- G. Zängl, D. Reinert, P. Rípodas, and M. Baldauf. The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141(687):563–579, 2015. ISSN 1477870X. doi: 10.1002/qj.2378.

Danksagung

Als Abschluss meiner Dissertation möchte ich die Möglichkeit wahrnehmen, mich bei all denjenigen Personen zu bedanken, die mich auf diesem Weg begleitet und unterstützt haben.

Mein herzlicher Dank geht hierbei an Prof. Dr. Peter Braesicke und Prof. Dr. Peter Knippertz, unter deren Leitung ich diese Promotion durchgeführt habe. Seit Beginn dieser Arbeit durfte ich mich sehr kompetent und konstruktiv betreut fühlen, was sowohl fachinterne als auch organisatorische Unterstützung und Beratung miteinschließt. Zudem möchte ich ganz besonders Dr. Matthias Schneider danken, dass er mir nicht nur dieses sehr spannende Thema für meine Promotion zur Verfügung gestellt, sondern meine Promotion über die 3.5 Jahre hinweg stets mit einem offenen Ohr sowie Antworten und neuen Ideen bereichert hat.

Zudem gilt mein Dank dem weiteren MUSICA-Team, insbesondere Benjamin Ertl, der mithilfe der Bereitstellung der IASI Retrieval-Ergebnisse ein wichtiges Fundament für meine Dissertation gelegt hat, und Dr. Farahnaz Khosrawi, eine der unermüdlichsten Korrekturleserinnen, die ich kenne. Der gemeinsame fachliche und kollegiale Austausch mit euch war im Arbeitsalltag stets inspirierend und motivierend. Des Weiteren möchte ich Dr. Thomas Blumenstock und Dr. Frank Hase für die sehr angenehme Aufnahme in die Arbeitsgruppe BOD danken, ebenso wie Dr. Darko Dubravica, Amelie Röhling, Carlos Alberti, Benedikt Herkommer und Dr. Qiansi Tu. Mein weiterer Dank geht an Dr. Roland Ruhnke, Dr. Johannes Eckstein und Dr. Jennifer Buchmüller für den fachlichen und technischen Support in der Arbeit mit ICON-ART_{iso}. Jochen Groß sowie Marc Rische gilt hierbei mein Dank für alle technische Unterstützung, wenn die Computer, Laptops und Server mal wieder nicht so taten, wie sie sollten. Ebenso möchte ich meiner Bürokollegin und Leidensgenossin Marleen Braun danken, für das gemeinsame Entdecken, Durchhalten und (hoffentlich schon bald erfolgreiche) Absolvieren des Doktor-Weges. Auch Christian Scharun, Katerina Kuskova, Simon Weber und der restlichen Doktoranden-Truppe des ASFs danke ich für eine gute Zeit, besonders im Rahmen des neu auf die Beine gestellten PhD-Meetings, welches während meiner Abschlussphase der Dissertation immer wieder ein erfrischender Ausgleich war. Ebenso bedanke ich mich bei Dr. Tobias Kerzenmacher, Dr. Florian Haenel, Dr. Sören Johansson und Dr. Michael Weimer sowie bei Sofia Frank, Esther-Salome Prietzel, Derya Cayiroglu und Regina Hühn für alle weitere Unterstützung in den letzten Jahren.

Weiterhin möchte ich meinen Dank an die Kollegen und Projektpartner richten, die zwar weltweit verstreut sitzen (Zürich, Berlin, Leicester, Tokyo, Washington und Bergen), aber mit inspirierenden Diskussionen in unzähligen Telefon-Konferenzen im Rahmen verschiedener Projekte (MOTIV, TEDDY und S5P+I H2O-ISO) meine Promotion immer wieder bereichern haben. Besonders möchte ich an dieser Stelle Dr. Andries J. de Vries nennen, der die COSMO_{iso}-Simulationen für meine Dissertation erstellt und mir zur Verfügung gestellt hat. Für die herzliche Aufnahme während meines unerwartet kurzen Besuchs in Leicester möchte ich Prof. Dr. Hartmut Boesch, Dr. Tim Trent sowie der Doktoranden-Truppe vor Ort meinen Dank aussprechen.

Ich bedanke mich bei der DFG sowie bei der ESA für die finanzielle Förderung meiner Tätigkeit und Promotion am KIT. Mein weiterer Dank ist an das SCC gerichtet, das die Nutzung von Rechen-, Speicher- und Archivsystemen (ForHLRII, LSDF, RADAR) für meine Promotionstätigkeit ermöglicht hat. In diesem Zuge möchte ich ganz besonders Dr. Felix Bach, Dr. Sabine Barthlott sowie Dr. Uğur Çayoğlu für den tatkräftigen technischen und organisatorischen Support beim Publizieren der MUSICA IASI Datensätze danken. Bei GRACE möchte ich mich für die vielzählige Förderung bedanken, die mir im Rahmen dieses Programms ermöglicht wurde und eine große Bereicherung für meine Doktorarbeit dargestellt hat.

Nicht zuletzt möchte ich all den lieben Leuten danken, die mich abseits der Arbeit begleitet haben und mir täglich Kraft und Inspiration waren. An all meine Freunde, sei es im TdH, im Hauskreis, in der Band oder unabhängig davon, es ist mir ein Privileg, euch kennen zu dürfen! Ganz besonders möchte ich an dieser Stelle meinen Eltern Frank und Dora sowie meinem Bruder Daniel danken, die ihr mich stets auf so unterschiedliche und liebevolle Art und Weise unterstützt. Nicht nur durch ganz praktische Hilfestellungen wie das unermüdliche Korrekturlesen und euer stets offenes Ohr, sondern ganz besonders dadurch, dass ihr Inspiration, Vorbild und Freudeschenker für mich seid, habt ihr einen wichtigen Teil zum Gelingen dieser Arbeit beigetragen. Danke, dass ich mich immer auf euch verlassen kann!

Mein Schlusswort gebührt dir, Hannah, die du nicht nur meine beste Freundin bist, sondern mittlerweile auch meine Frau. Danke für all die Liebe, Geduld und Unterstützung, die du mir bereits in so vielen Lebenslagen und ganz besonders in der Abschlussphase dieser Arbeit geschenkt hast. Es ist ein Geschenk des Himmels, mit dir gemeinsam unterwegs sein zu dürfen!