

Assessing water vapour from
state-of-the-art observations and models
in the central Arctic and the impact of
inversions on downwelling longwave
radiation

INAUGURAL-DISSERTATION

zur

Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln



vorgelegt von

Andreas Walbröl

aus Kirchen (Sieg)

20. September 2024

Berichtersteller/in:

Prof. Dr. Susanne Crewell

Prof. Dr. Ulrich Löhnert

Tag der mündlichen Prüfung: 11. Dezember 2024

Abstract

Observations have revealed that the rapidly warming Arctic is also moistening in certain regions and seasons. As water vapour is the strongest greenhouse gas, it contributes to the enhanced warming of the Arctic via the water vapour feedback. Water vapour estimates are uncertain in the Arctic due to the low amount of ground stations and challenges in satellite remote sensing. Thus, it is not surprising to see uncertainties in water vapour trends across reanalyses, which use these observations. In contrast to lower latitudes, Arctic humidity profiles feature inversions where the specific humidity increases with height. The representation of humidity inversions in current models and satellite products and the radiative effect of humidity inversions is poorly studied. Furthermore, the ability of ground-based microwave radiometers (MWRs) to capture humidity inversions has yet to be analyzed. The year-long Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition in the Arctic Ocean provides excellent reference water vapour observations to evaluate the water vapour products of models and satellites. Radiosonde observations are complemented by two MWRs with complementary water vapour sensitivity.

The first part of this thesis includes two studies to quantify the benefit of the synergy of the two MWRs for water vapour products compared to the use of single MWRs. In the first study, the measurements of each MWR were quality controlled and atmospheric parameters, including coarse humidity profiles and integrated water vapour (IWV), were retrieved using regression and Neural Networks. The single MWR retrievals were evaluated with the MOSAiC radiosondes. In the second study, measurements from both MWRs were combined in a Neural Network approach to exploit their complementary moisture sensitivity. The synergy benefit was determined by comparing the errors computed in the synergy evaluation to those of the single MWR retrievals. The synergy reduces lower tropospheric specific humidity errors by 50 % and the root mean squared error of IWV by 15 % over a wide atmospheric moisture range. Additionally, the vertical resolution of the specific humidity profile is improved by a factor of two in the lower troposphere.

In the second part of the thesis, the water vapour products from the MWR synergy and the radiosondes were used as reference to evaluate the water vapour from four state-of-the-art models (two global reanalyses, a regional and a global weather forecast model) and two satellite products. A particular focus was on analyzing the representation of humidity inversions with respect to radiosondes. Strong negative IWV and specific humidity biases in moist conditions were found for the satellite data. The models underestimate the lower tropospheric specific

humidity in the cold seasons, except for the lowest 100 m. The presence of surface-based inversions is well captured by the models and the MWR synergy but not by the satellite observations. Elevated inversions are missed by both the MWR synergy and satellite observations. Additionally, all tested data sets tend to underestimate the number of inversions per profile and the inversion strength but overestimate the vertical extent (depth) of inversions.

Then, radiative transfer simulations for clear sky cases were used to analyze the sensitivity of downwelling longwave radiation (DLR) to water vapour within humidity inversions. Therefore, the radiative effect of humidity inversions was also quantified. The sensitivity tests with artificial humidity inversion strength modifications revealed that in most cases the radiation emitted from water vapour within an inversion only has a notable impact on DLR when the inversion is located in the lower troposphere. In most of the clear sky cases, the surface-based inversion contributed 60–100 % to the total radiative effect of humidity inversions, which can reach up to 16 W m^{-2} . Additionally, we quantified DLR deviations resulting from using specific humidity profiles of the different models and observations (ground- and space-based) as input to the radiative transfer simulations. With the DLR based on the radiosonde profiles as reference, DLR deviations exceed 5 W m^{-2} in some cases but are mostly lower. The humidity profiles of the MWR synergy resulted in one of the smallest DLR deviations, demonstrating the high quality of the MWR humidity profiles. The deviations could be equally attributed to differences in lower tropospheric specific humidity and IWV. The results suggest that the IWV and the lower tropospheric specific humidity, and therefore the near-surface humidity inversions, are equally important for accurate DLR calculations.

Zusammenfassung

Beobachtungen haben ergeben, dass die sich rapide erwärmende Arktis in gewissen Regionen und Jahreszeiten ebenfalls feuchter wird. Wasserdampf, welches das stärkste Treibhausgas ist, trägt zur schnellen Erwärmung der Arktis durch den Wasserdampfdruckkopplungseffekt bei. Wasserdampfschätzungen sind aufgrund der geringen Anzahl an Bodenstationen und Schwierigkeiten in der Satellitenfernerkundung unsicher. Daher überrascht es nicht, dass Wasserdampftrends ebenfalls in verschiedenen Reanalysen, die jene Beobachtungen mitverwenden, unsicher sind. Im Gegensatz zu niedrigeren Breitengraden zeichnen sich arktische Feuchteprofile durch Inversionen aus, bei denen die spezifische Feuchte mit der Höhe zunimmt. Die Repräsentation von Feuchteinversionen in aktuellen Modellen und Satellitenprodukten, sowie der Strahlungseffekt von Feuchteinversionen wurde bislang nur wenig untersucht. Des Weiteren wurde die Fähigkeit von bodengebundenen Mikrowellenradiometern (MWR) Feuchteinversionen zu erfassen noch nicht analysiert. Die ganzjährige Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) Expedition in der Arktis bietet exzellente Referenz-Wasserdampfbeobachtungen für die Evaluierung von Wasserdampfprodukten von Modellen und Satelliten. Beobachtungen von Radiosonden werden durch zwei MWR, die komplementäre Feuchtigkeitssensitivitäten haben, ergänzt.

Der erste Teil der Dissertation umfasst zwei Studien, die zur Quantifizierung des Vorteils der Synergie der zwei MWR für Wasserdampfprodukte gegenüber der Verwendung einzelner MWR beitragen. In der ersten Studie wurde die Qualität der Messungen jedes MWR kontrolliert und atmosphärische Parameter, darunter grobe Feuchteprofile und der integrierte Wasserdampfgehalt, mithilfe von Regression und neuronalen Netzwerken abgeleitet. Die individuellen MWR Produkte wurden mittels der MOSAiC Radiosonden evaluiert. In der zweiten Studie wurden die Messungen beider MWR mit neuronalen Netzwerken kombiniert, um deren komplementäre Feuchtigkeitssensitivitäten zu nutzen. Der Vorteil der Synergie wurde durch den Vergleich der durch die Evaluierung bestimmten Fehler der Synergie mit denen der individuellen MWR ermittelt. Die Synergie reduziert den Fehler der spezifischen Feuchte in der unteren Troposphäre um 50 % und die Wurzel der mittleren Fehlerquadratsumme des integrierten Wasserdampfgehalt um 15 % über einen weiten atmosphärischen Feuchtebereich. Zudem wurde die vertikale Auflösung der spezifischen Feuchteprofile in der unteren Troposphäre um einen Faktor 2 verbessert.

Im zweiten Teil der Dissertation wurden die Wasserdampfprodukte der MWR Synergie, sowie die Radiosondenbeobachtungen als Referenz genutzt, um Wasser-

dampf von vier aktuellen Modellen (zwei globale Reanalysen, ein regionales und ein globales Wettervorhersagemodell) und zwei Satellitenprodukten zu evaluieren. Ein Schwerpunkt liegt auf der Analyse der Repräsentation von Feuchteinversionen im Vergleich zu Radiosonden. Für die Satellitendaten wurde eine starke Tendenz zu negativen Abweichungen des integrierten Wasserdampfgehalts und der spezifischen Feuchte in feuchten Bedingungen gefunden. Die Modelle unterschätzen die spezifische Feuchte in der unteren Troposphäre, abgesehen von den untersten 100 m, in den kalten Jahreszeiten. Die Präsenz von bodengebundenen Inversionen wird von den Modellen und der MWR Synergie gut dargestellt, allerdings nicht von den Satellitenbeobachtungen. Höher liegende Inversionen werden sowohl von der MWR Synergie, als auch von den Satellitenbeobachtungen nicht erfasst. Zudem tendieren alle getesteten Datensätze dazu, die Anzahl an Inversionen in einem Profil und die Inversionsstärke zu unterschätzen, während die vertikale Erstreckung der Inversionen (Tiefe) überschätzt wird.

Anschließend wurden Strahlungstransportsimulationen in wolkenfreien Fällen verwendet, um die Sensitivität von nach unten gerichteter langwelliger Strahlung (ULS) zu Wasserdampf innerhalb von Feuchteinversionen zu untersuchen. Somit wurde ebenfalls der Strahlungseffekt von Feuchteinversionen quantifiziert. Die Sensitivitätstests mit künstlichen Veränderungen der Feuchteinversionsstärke ergaben, dass die Strahlung, die von Wasserdampf innerhalb der Inversion emittiert wurde, in den meisten Fällen nur einen merklichen Effekt auf die ULS hat, wenn sich die Inversion in der unteren Troposphäre befindet. In den meisten der wolkenfreien Fälle trägt eine bodengebundene Inversion 60–100 % zum gesamten Strahlungseffekt von Feuchteinversionen bei, der Werte bis zu 16 W m^{-2} erreichen kann. Des Weiteren wurden ULS Abweichungen quantifiziert, die sich durch die spezifischen Feuchteprofile der verschiedenen Modelle und Beobachtungen (bodengebunden und weltraumgestützt) ergeben, wenn diese als Eingangsparameter in den Strahlungstransportsimulationen verwendet werden. ULS Abweichungen, gegenüber der ULS basierend auf den Radiosondenprofilen, können in manchen Fällen mehr als 5 W m^{-2} betragen, liegen aber meist darunter. Die Feuchteprofile der MWR Synergie ergeben mit die geringsten ULS Abweichungen, was die hohe Qualität der MWR Feuchteprofile demonstriert. Die Abweichungen konnten gleichermaßen auf Unterschiede in der niedertroposphärischen spezifischen Feuchte und im integrierten Wasserdampfgehalt zurückgeführt werden. Den Ergebnissen entsprechend sind der integrierte Wasserdampfgehalt und die niedertroposphärische spezifische Feuchte, und damit die bodennahen Feuchteinversionen, gleichermaßen wichtig für akkurate ULS Berechnungen.

Contents

Abstract	i
Zusammenfassung	iii
Contents	v
List of Figures	viii
List of Tables	xi
List of Abbreviations	xii
1 Introduction	1
1.1 Motivation	1
1.2 Water vapour observations in the Arctic	7
1.3 MOSAiC expedition	12
1.4 Goals of this thesis	14
1.5 Overview of the studies	15
2 Microwave radiative transfer	18
2.1 Emission and absorption	18
2.2 Scattering	22
2.3 Radiative transfer equation	23
3 Retrieval theory	28
3.1 Inverse modelling problem	28
3.2 Regression	29
3.3 Neural Networks	31
4 Single microwave radiometer retrievals	35
4.1 Abstract	36
4.2 Background & Summary	36
4.3 Methods	38

4.3.1	Microwave radiometers on board Polarstern	38
4.3.2	HATPRO: Retrieval via regression	40
4.3.3	MiRAC-P: Retrieval via Neural Network	41
4.4	Data records	43
4.5	Technical Validation	45
4.5.1	Brightness temperatures	46
4.5.2	Derived products	47
4.6	Usage Notes	57
5	Synergistic microwave radiometer retrievals	61
5.1	Abstract	62
5.2	Introduction	63
5.3	Data sets	66
5.3.1	Retrieval development data	66
5.3.2	MOSAiC observations for retrieval application and evaluation	66
5.4	Methods	70
5.4.1	Retrieval preparation	71
5.4.2	Retrieval setup	72
5.4.3	Metrics for retrieval evaluation and vertical information content	73
5.5	Retrieval evaluation	74
5.5.1	IWV	75
5.5.2	Specific humidity profiles	75
5.5.3	Temperature profiles	77
5.6	Information benefit analysis	79
5.6.1	IWV	80
5.6.2	Specific humidity profiles	81
5.6.3	Relative humidity profiles	84
5.7	Conclusions	87
5.A	Neural Network retrieval details	90
5.B	Information benefit: Influence of different method	91
5.C	Comparison with smoothed radiosonde profiles	92
6	Evaluation of water vapour products	96
6.1	Introduction	96
6.2	Data	101
6.2.1	MOSAiC observations	101
6.2.2	Numerical models	104
6.2.3	Satellite products	105
6.3	Methods	106

6.3.1	Data processing	106
6.3.2	Humidity inversion detection	108
6.3.3	Contingency table for inversion detectability	110
6.3.4	Radiative transfer model	110
6.3.5	Specific humidity profile modifications	112
6.4	Water vapour comparison	114
6.4.1	Integrated water vapour comparison	114
6.4.2	Specific humidity profile comparison	116
6.5	Humidity inversions	120
6.5.1	Inversion detectability	120
6.5.2	Inversion characteristics	122
6.6	Downwelling longwave radiation sensitivity	127
6.6.1	Sensitivity to inversion height and strength	128
6.6.2	Radiative effect of humidity inversions	132
6.6.3	Radiative impact of specific humidity deviations of different observations and models	136
6.7	Conclusion	139
6.A	Weighted average interpolation	144
6.B	Integrated water vapour comparison with 7-day running mean . .	145
6.C	Vertical grid spacing of all data sets	146
6.D	Benefit of the microwave radiometer synergy for humidity inversion detection	147
6.E	Downwelling longwave radiation sensitivity: Influence of tempera- ture inversion	148
7	Conclusions and Outlook	150
7.1	Summary and conclusions	151
7.2	Outlook	156
	Bibliography	160
	Acknowledgements	194
	Eidesstattliche Versicherung	196

List of Figures

1.1	Monthly IWV statistics at Ny-Ålesund	3
1.2	Specific humidity latitude-height cross-section in the northern hemisphere	5
1.3	Drift track of RV <i>Polarstern</i> during MOSAiC	13
2.1	Transmissivity in the microwave spectrum	21
2.2	Weighting functions between 22 and 340 GHz from a ground-based perspective	26
3.1	Schematic of a multilayer perceptron	32
4.1	IWV time series during MOSAiC	48
4.2	Comparison of single-MWR IWV retrievals with radiosondes	50
4.3	Time series of daily mean LWP during MOSAiC	53
4.4	Errors of absolute humidity profiles derived from HATPRO	54
4.5	Errors of temperature profiles derived from HATPRO	56
4.6	Thermodynamic overview of a strong moist air intrusion showing the MWR products and radiosondes	59
5.1	Brightness temperatures in the microwave spectrum in Arctic conditions	69
5.2	ERA5 training data grid points and RV <i>Polarstern</i> drift track	72
5.3	Synergistic IWV retrieval evaluation with synthetic evaluation data from ERA5	76
5.4	Synergistic specific humidity profile retrieval evaluation with synthetic evaluation data from ERA5	77
5.5	Synergistic temperature profile retrieval evaluation with synthetic evaluation data from ERA5	79
5.6	IWV error comparison of the synergistic retrieval with single-MWR retrievals	81
5.7	Specific humidity error comparison of the synergistic retrieval with single-MWR retrievals	83

5.8	Degrees of freedom for specific humidity profiles for different frequency combinations	85
5.9	Vertical resolution of the retrieved specific humidity profiles . . .	85
5.10	Relative humidity error comparison of the synergistic retrieval with single-MWR retrievals	86
5.11	IWV error comparison of the single-MWR retrievals with the Neural Network HATPRO retrieval	92
5.12	Specific humidity error comparison of the HATPRO regression retrieval with the Neural Network HATPRO retrieval	92
5.13	Specific humidity errors for profiles smoothed with the Averaging Kernel	94
6.1	Humidity inversion detection demonstration	109
6.2	Comparison of simulated and measured DLR	112
6.3	Specific humidity profile with modifications of the inversion strength (a) and base height (b). In the example (radiosonde launched on 26 February 2020, at 10:55 UTC), only the strongest inversion with a base height above 1000 m is modified.	113
6.4	IWV evaluation of models and satellite products	117
6.5	Specific humidity profile evaluation of models and satellite products	119
6.6	Specific humidity profile example showing radiosonde and remote sensing observations and models	120
6.7	Statistics of humidity inversion characteristics of models and remote sensing products compared with radiosondes	123
6.8	Comparison of humidity inversion characteristics of models and remote sensing observations with radiosondes for seasons and height layers	126
6.9	Humidity inversion strength and base height modifications	129
6.10	DLR deviations due to inversion strength and base height modifications	130
6.11	Humidity inversion strength and base height modifications with IWV conservation	131
6.12	DLR deviations due to inversion strength and base height modifications with IWV conservation	132
6.13	DLR deviations due to strength modifications for all 120 clear sky cases	134
6.14	Radiative effect of humidity inversions	136

6.15 DLR deviations caused by using specific humidity profiles of different models and observations, with DLR from radiosonde profiles as reference	138
6.16 Influence of different interpolation strategies on the specific humidity profile	144
6.17 Visualization of interpolation using weighted averages	145
6.18 IWV evaluation of models and satellite products with a 7-day running mean	146
6.19 Vertical grid spacing of the models, radiosondes and remote sensing observations	147
6.20 Humidity inversion strength and base height modifications and temperature profile of a case study	148
6.21 DLR deviations due to inversion strength and base height modifications with IWV conservation and isothermal temperature profile	149

List of Tables

4.1	Published data sets	44
4.2	Single-MWR IWV errors with respect to radiosondes	51
4.3	Published codes	52
5.1	Settings of the synergistic Neural Network retrievals	91
6.1	Contingency table	111
6.2	Humidity inversion detectability statistics for all models, remote sensing observations and radiosondes computed from the contingency table	122
6.3	Sensitivity of DLR deviations to water vapour changes within inversions at different heights	130
6.4	Humidity inversion detectability statistics comparing HATPRO and the synergistic retrieval	148

List of Abbreviations

(AC) ³	Arctic amplification: Climate Relevant Atmospheric and Surface Processes and Feedback Mechanisms
AK	Averaging Kernel
AMSR	Advanced Microwave Scanning Radiometer
AMSU	Advanced Microwave Sounding Unit
ARM	Atmospheric Radiation Measurement
ASFS	Atmospheric Surface Flux Stations
AVHRR	Advanced Very High Resolution Radiometer
AWI	Alfred Wegener Institute
BL	Boundary layer
C3S	Copernicus Climate Change Service
CAFS	Coupled Arctic Forecast System
CFC	Chlorofluorocarbon gases
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DKRZ	Deutsches Klimarechenzentrum
DLR	Downwelling longwave radiation
DOF	Degrees of freedom
DOY	Day of the year
DWD	Deutscher Wetterdienst (German Weather Service)
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF reanalysis v5
ERA-I	ECMWF reanalysis interim
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GCOM-W1	Global Change Observation Mission 1st-Water
GCOS	Global Climate Observing System
GFS	Global Forecast System
GMAO	Global Modeling and Assimilation Office
GNSS	Global Navigation Satellite Systems
GRUAN	GCOS Reference Upper-Air Network
HALO	High Altitude and Long-range Research Aircraft
HATPRO	Humidity And Temperature PROFiler
HPC	High-Performance Computing
IASI	Infrared Atmospheric Sounding Interferometer
ICAO	International Civil Aviation Organization
ICON	ICOsahedral Non-hydrostatic
IDL	Interactive Data Language
IPCC	Intergovernmental Panel on Climate Change
IQR	Interquartile range
IWC	Ice water content
IWV	Integrated water vapour
JAXA	Japan Aerospace Exploration Agency
LHUMPRO	Low HUMidity PROFiler
LWC	Liquid water content
LWP	Liquid water path
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications, Version 2

MHS	Microwave Humidity Sounder
MiRAC-P	Microwave Radiometer for Arctic Clouds - Passive
MODIS	Moderate Resolution Imaging Spectrometer
MOSAiC	Multidisciplinary drifting Observatory for the Study of Arctic Climate
MPI-M	Max Planck Institute for Meteorology
MWR	Microwave radiometer
MWRRET	Microwave radiometer retrievals
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NN	Neural Network
NOAA-PSL	National Oceanic and Atmospheric Administration Physical Sciences Laboratory
OEM	Optimal estimation method
PAMTRA	Passive and Active Microwave radiative TRAnsfer model
PANGAEA	PAleoNetwork for GeologicAl and EnvironmentAl data - Data Publisher for Earth & Environmental Sciences
PPF	Product Processing Facility
Polly ^{XT}	Portable Lidar system, next generation
PREFIRE	Polar Radiant Energy in the Far-InfraRed Experiment
RASM	Regional Arctic System Model
RMSD	Root mean squared deviation
RMSE	Root mean squared error
RPG	Radiometer Physics GmbH
RQ	Research Question
RRTMG	Rapid radiative transfer model for general circulation models
RS	Radiosonde
RV	Research vessel
T-CARS	TROPOS Cloud and Aerosol Radiative effect Simulator
TB	Brightness temperature
TOA	Top of the atmosphere
TROPOS	Leibniz Institute for Tropospheric Research
WAI	Warm air intrusion

Chapter 1

Introduction

1.1 Motivation

The Earth is warming rapidly. Throughout the history of Earth, tropospheric temperatures have changed in response to several factors, including variations in the Earth's orbit around the Sun, the distribution of land masses, or changes in greenhouse gas concentrations. However, the current rate of warming is considerably higher than expected from natural causes. It is scientific consensus that the warming is attributed to human activities, such as the enhanced emissions of CO₂ (IPCC, 2023).

The hot spot of the current climate change is the Arctic (Screen et al., 2012; Wendisch et al., 2023). Here, the troposphere has warmed about three times faster than the global average, a rate that climate models struggle to simulate (Rantanen et al., 2022; Zhou et al., 2024). The warming is not uniformly distributed in space or time but especially pronounced in winter (Serreze et al., 2009; Maturilli et al., 2015; Maturilli and Kayser, 2017a) and in the Barents and Kara Seas (Rantanen et al., 2022). In these regions, near-surface air temperatures have increased up to 7 K since 1979 (Rantanen et al., 2022). However, in summer, the warming at the surface is less evident because the heat excess melts the sea ice, keeping the near-surface temperatures close to the freezing point (Tjernström et al., 2004; Graversen et al., 2008).

These drastic changes in the Arctic temperatures have implications for the local climate system and also affect lower latitudes through heat exchange driven by atmospheric circulation. The most striking effect of the strong warming in the Arctic is the sea ice decline, thinning and shift to younger sea ice (IPCC, 2022). The thinner ice is more susceptible to melting over the subsequent summers. In some Arctic regions, the ecosystem changes drastically as the climate shifts towards current sub-polar conditions (Francis et al., 2009). For example, fish species

from sub-polar latitudes are expected to migrate into Arctic seas and compete with local species for food (IPCC, 2022). The Arctic climate change can also trigger new emissions of greenhouse gases from thawing permafrost. Additionally, thawing permafrost enhances the risk of landslides in mountainous regions, posing a threat to the local wildlife and inhabitants (IPCC, 2022; IPCC, 2023). These mentioned impacts are merely the tip of the iceberg of the consequences of climate change.

A dedicated team of international scientists within the Transregional Collaborative Research Center TRR 172 on "Arctic amplification: Climate Relevant Atmospheric and Surface Processes and Feedback Mechanisms (\mathcal{AC})³" contributed significantly to the advancement in understanding the processes causing Arctic amplification through dedicated measurement campaigns and model experiments (Wendisch et al., 2023). Despite the scientific advancement in the past years, challenges remain regarding understanding the Arctic climate system and the correct representation of individual processes in climate models. Especially clouds and their radiative impact are not yet well represented in climate models (Pithan et al., 2014; Sedlar et al., 2020). A particular challenge is posed by mixed-phase clouds, which contain both ice and supercooled liquid particles and frequently occur in the Arctic (Nomokonova et al., 2019; Gierens et al., 2020).

Arctic amplification results from several climate responses that amplify an initial temperature perturbation, known as positive climate feedback loop. The strongest contributors are the ice albedo¹ (Serreze et al., 2009; Screen and Simmonds, 2010; Serreze and Barry, 2011; Screen et al., 2012) and the lapse rate² feedback (Pithan et al., 2014; Feldl et al., 2020; Hahn et al., 2021; Linke et al., 2023). In autumn, the additional open water regions due to the decline of the sea ice extent result in strong heat and moisture fluxes from the warm ocean to the cold atmosphere, providing an additional local moisture source (Brümmer and Pohlmann, 2000; Alekseev et al., 2019).

Water vapour, which also contributes significantly to the warming of the Arctic (Winton, 2006; Graversen and Wang, 2009; Serreze and Barry, 2011), is the focus of this thesis. It is the strongest greenhouse gas as it absorbs and re-emits radiation in a broad spectral range in the thermal infrared ('longwave' spectrum) (e.g., Kiehl and Trenberth, 1997; Trenberth, 1998). Ghatak and Miller (2013) found a large sensitivity of downwelling longwave radiation to the integrated wa-

¹Warming results in snow and ice melt, reducing the reflectivity (albedo) of the surface and allowing more solar radiation to be absorbed. The absorbed radiation heats the darker surface (open ocean or land), which promotes melting and warming of the near-surface air through enhanced radiative and turbulent heat fluxes.

²In contrast to the Tropics, where convection distributes heat over the entire troposphere, the Arctic boundary layer (lowermost part of the troposphere that interacts with the surface) is strongly stratified, trapping heat near the surface.

ter vapour (IWV) in the Arctic. The sensitivity is stronger in dry conditions (low IWV, up to 31 W kg^{-1}) than in humid conditions (high IWV, $\approx 18 \text{ W kg}^{-1}$) (Ghatak and Miller, 2013). The IWV is defined as the vertical integral of the specific humidity q , which is the relation of the mass of water vapour to the total mass of air, over the atmospheric column:

$$\text{IWV} = -\frac{1}{g} \int_{p_{\text{sfc}}}^{p_{\text{top}}} q \, dp, \quad (1.1)$$

where g is the standard gravitational acceleration, and p_{sfc} and p_{top} are the air pressure at the surface and the top of the atmosphere, respectively. Typical IWV values at the Arctic research site in Ny-Ålesund (Svalbard, Norway) range from 2 kg m^{-2} in winter to 17 kg m^{-2} in summer (interquartile ranges in Fig. 1.1). In extreme conditions, IWV can be as low as 0.4 kg m^{-2} in January and reach up to 34 kg m^{-2} in July (Fig. 1.1).

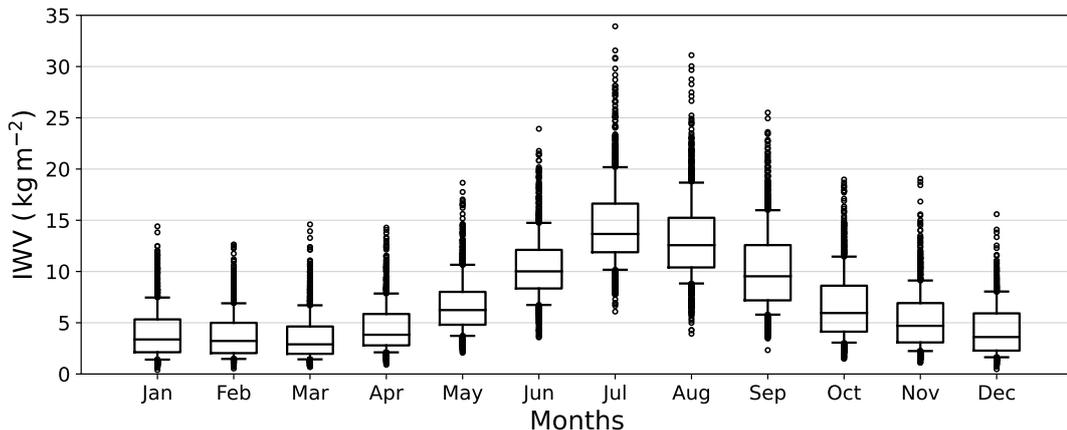


Figure 1.1: Monthly statistics of integrated water vapour (IWV) computed from radiosonde observations at Ny-Ålesund covering the years 1993–2023 (Maturilli and Kayser, 2016, 2017b; Maturilli, 2020). The box indicates the interquartile range and the horizontal line within the box shows the median. Whiskers indicate the 10th and 90th percentiles. IWV values outside the 10th or 90th percentile are illustrated as dots.

Water vapour contributes to the rapid warming of the Arctic via the water vapour feedback loop (Alexeev et al., 2005): According to the Clausius-Clapeyron equation, a warmer atmosphere can contain more water vapour before saturation is reached (15 % higher water vapour pressure per 1 K warming at 200 K, and only 6 % K^{-1} at 300 K, Held and Soden, 2000). Thus, higher temperatures and enhanced evaporation from larger open water regions intensify the longwave radiation at the surface, which radiatively heats the surface and near-surface air, triggering further ice melt and a temperature increase. According to Zhang et al. (2001), the IWV has a greater effect on snow melt than the mean atmospheric

temperature, but both effects are also linked by the Clausius-Clapeyron relation.

Furthermore, water vapour is a key part of the global water cycle (Trenberth, 1998; Bengtsson, 2010) and is strongly connected to the formation of clouds. In case the air is saturated with water vapour (over liquid water or ice), clouds form due to condensation, freezing or deposition (Lohmann et al., 2016, Chapter 6.1). Through clouds, water vapour also has an *indirect* radiative effect: Clouds scatter radiation in the visible spectrum of electromagnetic radiation (solar 'shortwave' radiation) but absorb and emit in the longwave (thermal infrared) (Shupe and Intrieri, 2004). The net radiative effect of the clouds at the surface compared to clear sky conditions is known as the cloud radiative effect. Liquid clouds dominate the longwave emission signal of clouds compared to pure ice clouds (Hong and Liu, 2015) unless the integrated cloud liquid water content (liquid water path, LWP) is very small ($< 5 \text{ g m}^{-2}$, Ebell et al., 2020). During polar night, the effect of solar radiation is absent, yet clouds continue to emit longwave radiation and trap the heat below the clouds, which results in higher surface and air temperatures. During polar day over sea ice, when both the solar and the infrared radiative effects play a role, the cloud radiative effect can become negative in mid-summer when the shading of the shortwave radiation dominates the longwave emission (Shupe and Intrieri, 2004; Nomokonova et al., 2020). In the warming Arctic, some observational studies suggest a regional increase in cloud occurrence (Devasthale et al., 2016; Nomokonova et al., 2020) but the overall trend is uncertain (Vihma et al., 2016; Wendisch et al., 2023). Because of the net positive cloud radiative effect (Shupe and Intrieri, 2004), an increase in cloud cover can contribute to Arctic amplification.

The vertical water vapour distribution is important for direct radiative effects (Colman, 2001; Devasthale et al., 2011) and the interaction of water vapour with clouds (Solomon et al., 2011; Sedlar et al., 2012; Brunke et al., 2015). Tjernström et al. (2019) found that when temperature and specific humidity increase rather than decrease with height, the surface received additional radiative heating of up to 25 W m^{-2} . However, the cloud occurrence was also higher. The increase of specific humidity with height is known as humidity inversion and is characteristic of the Arctic. Figure 1.2a shows that the specific humidity increases with height at latitudes north of 60°N in winter while it decreases with height at lower latitudes. In the Arctic Ocean, the humidity inversion is very distinct (Fig. 1.2b).

Humidity inversions occur throughout the entire year, with occurrence rates of 90 % in winter and more than 60 % in summer (Naakka et al., 2018). Humidity inversions can be formed by several processes: Strong radiative cooling at the surface during polar night (Curry, 1983) results in the formation of temperature inversions and reduces the saturation water vapour pressure according to the

Clausius-Clapeyron relation. Thus, as the water vapour pressure reaches saturation, the water vapour condenses to the liquid phase, reducing the water vapour concentration and resulting in the formation of fog (Curry, 1983). The fog can dissolve due to gravitational settling or precipitation of the condensates. Condensation driven by radiative cooling dries the atmosphere and forms a surface-based humidity inversion, often coinciding with a temperature inversion (Brunke et al., 2015; Naakka et al., 2018). Moisture advection from lower latitudes is another essential mechanism to form humidity (and temperature) inversions in the Arctic as the lower tropospheric part cools and water vapour condenses (Tjernström et al., 2004; Brunke et al., 2015). Such elevated temperature and humidity inversions are often accompanied by low-level mixed-phase clouds, whose cloud top is near the inversion top (Sedlar et al., 2012; Brunke et al., 2015). Entrainment at the cloud top, driven by radiative cooling, and a downward moisture gradient due to the humidity inversion draws the higher water vapour concentration downwards into the cloud (Solomon et al., 2011). Consequently, humidity inversions serve as a moisture source for low-level mixed-phase clouds, contributing to their persistence over time.

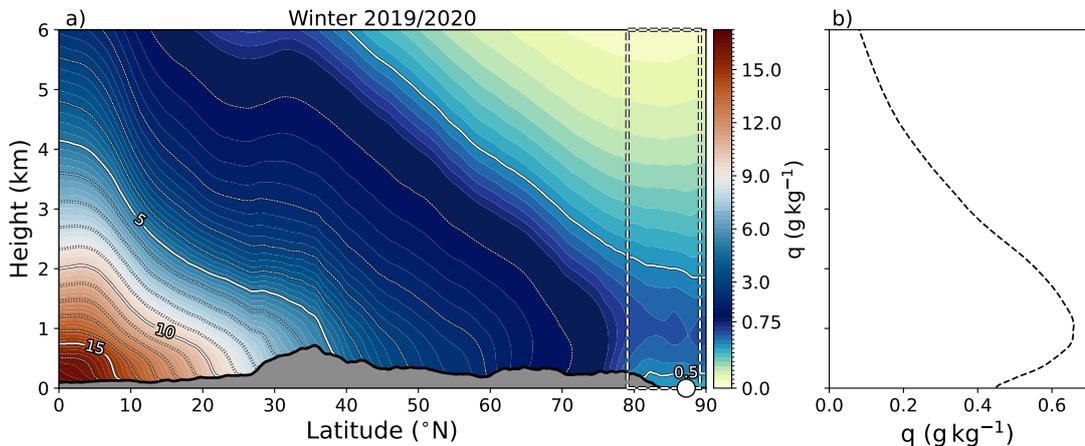


Figure 1.2: Specific humidity q latitude–height cross-section of the northern hemisphere for winter 2019/2020 (a). The dashed white box indicates the latitude range of the research vessel (RV) *Polarstern* during the MOSAiC expedition (Sect. 1.3). The white circle at 87°N indicates the average latitude of RV *Polarstern* during winter 2019/2020, which is the latitude of the mean specific humidity profile shown in (b). The data is based on the European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis v5 (ERA5).

Arctic amplification has also been related to changes in the global atmospheric circulation, which affects the transport of moisture into and out of the Arctic: Crasemann et al. (2017) found that regional changes in atmospheric circulation patterns could be attributed to sea ice loss using reanalyses and general circulation models. However, the relation between Arctic amplification and the long-term

changes of the atmospheric circulation is still under debate (Vavrus, 2018; Lee et al., 2019). For example, the results from Francis and Vavrus (2015) suggest that the jet stream, a strong wind band driven by the temperature contrast between the tropics and the polar regions, tends to meander more in the future because of the reduced meridional temperature gradient in the lower troposphere. However, several recent studies using reanalyses and climate models have indicated that the warming of the tropical upper troposphere and the simultaneous cooling of the polar lower stratosphere strengthens the jet stream as the meridional temperature gradient is increased in the upper troposphere (Lee et al., 2019; Stendel et al., 2021; Blackport and Fyfe, 2022; Nie et al., 2023). Climate simulations performed by Priestley and Catto (2022) also indicate that the number of cyclones reaching Svalbard and the Fram Strait in winter will decrease by the end of the 21st century under high anthropogenic emission scenarios. These results suggest a reduced waviness of the jet stream.

Past studies identified trends in the occurrence of warm air intrusions (WAIs) because of decreased meridional temperature gradients (Mewes and Jacobi, 2019) and increased occurrence of blocking situations (Rinke et al., 2017; You et al., 2022b). Such trends may be related to the enhanced waviness of the jet stream observed in recent decades. In the mid-latitudes, an increased occurrence of blocking situations results in an increase in the number and intensity of heat waves and droughts on the one hand, but also more cold air outbreaks and floods on the other hand (Stendel et al., 2021). Additionally, atmospheric blocking and meandering jets increase the atmospheric exchange of heat and moisture between the mid-latitudes and the Arctic through warm air intrusions (WAIs) and cold air outbreaks, which contribute to the large variability of IWV seen in Fig. 1.1. The findings from Alexeev et al. (2005) and Hahn et al. (2021) indicate that the increase of poleward heat and moisture transport is an important contributor to Arctic amplification. However, climate model simulations indicate that the observed (and simulated) recent positive jet waviness trend will reverse in the future (Cattiaux et al., 2016).

The strong poleward heat and moisture transport related to WAIs greatly affect the Arctic. The advected heat and the enhanced greenhouse effect due to clouds and water vapour result in higher temperatures, enhancing ice melt or hindering ice formation and growth (Kapsch et al., 2013; Tjernström et al., 2015; Johansson et al., 2017; Graham et al., 2019c; You et al., 2021, 2022a). WAIs that passed through the North Atlantic sector of the Arctic, which is a major pathway of WAIs (Serreze et al., 1995; Mewes and Jacobi, 2019), have contributed to the warming trend at Ny-Ålesund (especially in winter, Rinke et al., 2017; Dahlke and Maturilli, 2017). Also in the Barents Sea, trends in the occurrence of WAIs in

winter could explain half of the temperature trend (Woods and Caballero, 2016).

As the positive trend of WAIs is correlated with positive temperature trends, and the water vapour amount is related to temperature, it is not surprising that moistening of the Arctic has been observed (Chen and Liu, 2016; Maturilli and Kayser, 2017a; Rinke et al., 2019). The moistening trend is mainly confined to lower altitudes (Maturilli et al., 2013; Allan et al., 2022), where also temperatures increased the most, and has a strong regional and seasonal variability (Maturilli and Kayser, 2017a; Parracho et al., 2018). Based on reanalyses, the strongest relative trends have been identified around Svalbard in late autumn and early winter (up to 12 % decade⁻¹, Rinke et al., 2019), coinciding with the enhanced sea ice loss (Rantanen et al., 2022). Other regions and seasons also show drying trends (e.g., eastern Siberia in late winter or northern central Siberia in summer).

Findings from Nygård et al. (2020) indicate a correlation between increased moisture transport into the Arctic due to changes in the circulation and regional and seasonal IWV trends. Wang et al. (2024) estimated that enhanced moisture intrusion activity in summer is responsible for more than a third of the summertime moistening trend in the Arctic. However, different observations or reanalyses strongly disagree on the regional distribution and the magnitude of trends (Parracho et al., 2018; Rinke et al., 2019). In summer, the variation across data sets even exceeds the magnitude of the trend itself, so the sign of the trend is uncertain. The highest uncertainties occur in summer in the data sparse regions of the central Arctic, but also Siberia and the North Atlantic sector of the Arctic (Rinke et al., 2019). Trend estimates in the Arctic are challenging because of the low accuracy of water vapour observations (Schröder et al., 2018). For example, Crewell et al. (2021) studied the errors of satellite products and reanalyses during a field campaign carried out in the vicinity of Svalbard in May–June 2017. In the central Arctic, the monthly mean IWV varied up to 30 % between satellite products, and for single events the errors can be much higher (also due to the spatio-temporal variability of water vapour).

Uncertainties in water vapour estimates from reanalyses in the Arctic are related to the low density of ground observations and difficulties in satellite remote sensing. Thus, data assimilation is limited in the Arctic, and reanalyses have to rely more on the underlying model. The options and challenges of water vapour measurements in the Arctic are elaborated in the following section.

1.2 Water vapour observations in the Arctic

As for most other standard meteorological parameters, water vapour can be measured by a sensor directly interacting with the air (in-situ) or observing an air

volume from a distance (remote sensing). Remote sensing is not trivial as water vapour is not directly measured but only its interaction with electromagnetic radiation. Inverse modelling techniques, which are explained in Chapter 3, must be used to derive atmospheric parameters from the remote sensing observations.

In the Arctic, in-situ humidity measurements are rare because only a few ground stations (synoptic stations) exist at the coasts of the Arctic seas (e.g., Ny-Ålesund, Svalbard; Barrow, Alaska; Eureka and Alert, Canada, Uttal et al., 2016), and only a fraction of the stations launch radiosondes. Autonomous drifting buoys ([International Arctic Buoy Programme, accessed 16 May 2024](#)) are scattered in the central Arctic Ocean, but they usually only provide surface measurements (e.g., temperature, Rigor et al., 2000) and rarely measure humidity.

Nowadays, most ground stations, e.g., at Ny-Ålesund, and nearly all radiosondes measure humidity with capacitive sensors (capacitive hygrometer) that have a water vapour absorbent polymer between two electrodes (Maturilli et al., 2013; Foken, 2021, Chapters 8.3 and 8.4). Capacitive hygrometers have a high accuracy (usually 2–5 %) and low response time to allow for a high vertical resolution of soundings (measuring at 1 Hz with ascent rates of about 5 m s^{-1}). The radiosonde processing software corrects potential biases caused by long response times of the humidity sensor at low temperatures, radiative heating of the sensor, or wake effects of the ascending balloon (Foken, 2021, Chapter 46.3). Connell and Miller (1995) estimated the IWV error due to response times of the temperature and humidity sensors to be generally less than 0.1 kg m^{-2} . The Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN), which provides high-quality processing of radiosonde measurements, estimates the IWV uncertainty of the Vaisala RS41 radiosondes used in this thesis to be less than 0.5 kg m^{-2} in Arctic conditions (Sommer et al., 2023). It is important to respect the drift of the radiosonde due to wind during the ascent and descent when comparing the sounding to other measurements.

In-situ measurements from synoptic stations or radiosondes are often used as reference data to evaluate other humidity products because of their high accuracy (and vertical resolution). However, they only provide point measurements in space and time that might only be representative for a small area around the station. The poor spatial coverage and low temporal sampling rates of radiosondes impede an adequate representation of the water vapour variability and distribution in the Arctic.

Ground-based microwave radiometers (MWRs) also provide highly accurate IWV, as well as coarse humidity (and temperature) profiles (and LWP, Solheim et al., 1998; Crewell et al., 2001; Ware et al., 2003; Löhnert et al., 2004). MWRs have been deployed at different sites around the world, from polar regions to mid-

latitudes and subtropics (Macke et al., 2010; Steinke et al., 2015; Nomokonova et al., 2019; Bock et al., 2021). The exploration of the sensitivity of microwave radiance observations to water vapour dates back to 1946, when Dicke et al. (1946) investigated the absorption strength of water vapour around the 22.235 GHz absorption line. Combining radiance measurements at strongly absorbing frequencies with weakly absorbing frequencies allows the derivation of IWV, LWP and coarse humidity profiles (e.g., Solheim et al., 1998). Details on the physics behind microwave remote sensing follow in Chapter 2. As the 22.235 GHz absorption line is relatively weak, observations along the much stronger 183.31 GHz absorption line have been used to enhance the accuracy of water vapour observations in dry conditions (Cadeddu et al., 2007). By combining low- and high-frequency microwave radiometers, Cadeddu et al. (2022) achieved IWV uncertainties of only 0.2 kg m^{-2} and specific humidity profile standard deviations of $0.2\text{--}0.3 \text{ g kg}^{-1}$ at extremely dry Antarctic sites, where IWV did not exceed 9 kg m^{-2} . However, in the Arctic, the variability of IWV is much higher with values ranging from less than 1 to more than 30 kg m^{-2} (see Fig. 1.1). It is more challenging for retrievals to be accurate for both extremely dry and relatively humid conditions. A significant advantage of MWRs over radiosondes is the high temporal resolution of 1 second, allowing them to better capture water vapour variability during air mass transitions and storms (Crewell et al., 2021). Furthermore, modern MWRs are relatively compact, require little power supply and maintenance, and can be automated, making them suitable for continuous operational monitoring of the atmosphere (Rose et al., 2005).

Enhanced spatial coverage can be realized through remote sensing based on the Global Navigation Satellite System (GNSS). GNSS send signals to receiving stations on Earth at frequencies of about 1–1.5 GHz (Ning et al., 2016). Compared to vacuum conditions, the signal is delayed (known as slant total delay) due to the refractive index of the atmosphere (Foken, 2021, Chapters 36.1 and 36.3). From the slant total delay, the zenith total delay is derived, which mainly depends on the atmospheric pressure, temperature and humidity. To accurately derive IWV, pressure and temperature information must be added from additional observations, assumptions or numerical weather forecast. Compared to radiosondes, IWV errors are on the order of $1\text{--}2 \text{ kg m}^{-2}$ (Ning and Elgered, 2012; Van Malderen et al., 2014). The advantages of GNSS remote sensing include fully automated operation and the ability to derive three-dimensional water vapour distributions by combining multiple GNSS satellites and receiving stations at the ground (Van Malderen et al., 2014; Foken, 2021, Chapters 36.1 and 36.6). However, receiving stations are sparse in the Arctic, with no permanent installations on the sea ice.

For Arctic-wide coverage, satellite remote sensing observations are needed. Polar orbiting satellites are on sun-synchronous orbits that sample a region at roughly the same local time each day. They orbit around Earth at altitudes of 300–1000 km within approximately 90–103 min (Foken, 2021, Chapter 37.1). In the Arctic, the orbits are sufficiently close together to achieve full Arctic coverage at least once per day per satellite. Given the numerous polar orbiting satellites from various space agencies (National Aeronautics and Space Administration (NASA), European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), Japan Aerospace Exploration Agency (JAXA), etc.), the entire Arctic can be sampled several times per day with different instruments. However, all these instruments, operating at different wavelengths of the electromagnetic spectrum, face distinct challenges in the Arctic:

As an example of the near-infrared spectrum, the Moderate Resolution Imaging Spectrometer (MODIS) IWV retrieval is described. Solar light radiances reflected at the Earth’s surface are measured at wavelengths around the $0.94\ \mu\text{m}$ water vapour absorption line (Gao and Kaufman, 2003). The ratio between strongly and weakly absorbing channels is used to derive IWV and to cancel out surface and other atmospheric attenuation effects (e.g., aerosols). This method requires reflective surfaces (ocean areas with Sun glint, clouds, sea ice) and daylight, and can therefore only be used from spring to early autumn in the central Arctic. Additionally, clouds can obscure the lower atmosphere, leading to biases in the retrieved IWV, and the signal can be weak in very dry conditions (Alraddawi et al., 2018; Crewell et al., 2021). Errors are typically between 5 and 10 % under ideal conditions (Gao and Kaufman, 2003) but can be much greater in the Arctic.

In the thermal infrared and microwave spectrum, the emission of radiation from water vapour around resonant rotational or vibrational absorption lines is used to derive IWV and coarse humidity profiles. In contrast to the visible and near-infrared spectrum, water vapour observations are possible throughout the year because of the independence of solar radiation. For example, the Infrared Atmospheric Sounding Interferometer (IASI, Blumstein et al., 2004) measures infrared radiances at a high spectral resolution between 3.6 and $15.5\ \mu\text{m}$ and therefore includes the strong water vapour absorption line at around $6\ \mu\text{m}$. As mentioned for the ground-based MWRs, observations at frequencies with different absorption strengths can be used to derive water vapour profiles and IWV. Infrared remote sensing observations are limited to clear sky or partly cloudy scenes because clouds are strong absorbers in the infrared, attenuating signals below the clouds. At microwave frequencies, clouds are semi-transparent so that the whole atmosphere can be sensed in almost all sky conditions. The atmospheric opacity is generally low in the lower part of the microwave spectrum (1–100 GHz), except

at frequencies close to resonant absorption line centres, and increases with frequency due to the water vapour continuum absorption (see Chapter 2). At low frequencies and in dry conditions, the microwave radiation emitted by the surface strongly contributes to the total signal received by the satellite sensor. Over open ocean, the surface emission signal is well known and much smaller than over sea ice due to lower emissivities (Prigent et al., 2017). Over sea ice, the surface signal is stronger (generally higher emissivities) and substantially more uncertain because of the high spatio-temporal variability of the surface emissivity (Willmes et al., 2014; Lee et al., 2017; Risse et al., 2024). The sea ice emissivity depends on its dielectric and scattering properties, which may rapidly change (e.g., depending on sea ice temperature, salinity, and snow properties, Mathew et al., 2008; Wang et al., 2017). Consequently, water vapour products from satellite-based MWRs have high uncertainties over sea ice (Crewell et al., 2021) and are usually only used over open ocean. Over (polar) oceans, IWV errors are typically around 10 % (25 %) (Schröder et al., 2018).

The combination of microwave and infrared radiances can significantly reduce IWV errors compared to pure infrared or microwave retrievals in the Arctic and is currently one of the most accurate approaches (Roman et al., 2016; Crewell et al., 2021). For example, on the Metop satellites from EUMETSAT, infrared radiances (e.g., IASI) are combined with microwave radiances from the Advanced Microwave Sounding Unit (AMSU-A) and Microwave Humidity Sounder (MHS) (August et al., 2012; Roman et al., 2016). IWV errors of the combined product are on the order of 1–1.5 kg m⁻² with positive (negative) biases when IWV is below (above) 10 kg m⁻² (Roman et al., 2016; Crewell et al., 2021).

New satellites, for example, the recently launched Arctic Weather Satellite (launched in August 2024), extend the microwave observations to the sub-millimetre range (> 300 GHz). Due to the higher atmospheric opacity of the sub-millimetre spectrum compared to lower microwave frequencies, the measurements are less affected by the surface, which has been shown to be a source of error in short-term forecasts over sea ice (Lawrence et al., 2019). Furthermore, the Polar Radiant Energy in the Far-Infrared Experiment (PREFIRE) satellite was launched in May 2024 to close the gap of observations beyond 15 μm , exploring the emission of radiation in the far infrared up to 50 μm (L’Ecuyer et al., 2021).

Other methods, such as ground-based (or airborne) active remote sensing techniques in the visible and near-infrared spectrum, e.g., Raman lidar or differential absorption lidar (Wirth et al., 2009; Foken, 2021, Chapters 25 and 26), are also known to provide accurate IWV and water vapour profiles, but are limited by clouds. Currently, differential absorption radar techniques are also explored for water vapour profiles inside clouds (Schnitt et al., 2020; Roy et al., 2021).

However, these active remote sensing techniques are currently only used at a few ground stations or only during field campaigns.

Due to the low amount of ground stations available in the central Arctic Ocean, the quality of satellite products is mainly assessed with reanalyses, which assimilate satellite measurements and are therefore not independent of each other. Field campaigns provide great opportunities to evaluate satellite products and reanalyses through high-quality measurements. A particularly valuable field campaign, the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition (Shupe et al., 2022), will be the focus of this thesis and is introduced in the following section.

1.3 MOSAiC expedition

The MOSAiC expedition took place in the Arctic Ocean from October 2019 to September 2020 to study the Arctic climate system for a complete annual cycle (Shupe et al., 2022). It has been the most extensive Arctic field campaign and provides the most comprehensive set of atmospheric, sea ice, oceanic, ecosystem and biogeochemical measurements in the central Arctic to date. The German research vessel (RV) *Polarstern* from the Alfred Wegener Institute (AWI), Helmholtz Centre for Polar and Marine Research, was frozen into the ice and began to passively drift along with an ice floe on 04 October 2019. The drift started in the northern Laptev Sea and ended in the marginal ice zone in the Fram Strait on 31 July 2020, with a logistical interruption between 16 May and 19 June (Fig. 1.3).

From January to March 2020, the atmospheric circulation was dominated by a record-breaking positive phase of the Arctic Oscillation index, which means that the pressure over the central Arctic was lower while the pressure over the Atlantic and Pacific Oceans was higher (Lawrence et al., 2020). The atmospheric conditions and the unusually thin sea ice resulted in a very rapid sea ice drift (Kruppen et al., 2021) so that RV *Polarstern* reached the marginal ice zone in the Fram Strait, where the ice floe disintegrated, earlier than expected. Therefore, between 31 July and 21 August 2020, the ship transited through the ice towards the North Pole, where the autumn refreezing was captured. Overall, the atmospheric conditions were within the interquartile range of the 1979–2019 climatology (Rinke et al., 2021). Thus, the MOSAiC expedition represents the typical Arctic conditions well. However, also a few record-breaking events were observed (e.g., the WAI in mid-April 2020, Kirbus et al., 2023). The melting season along the MOSAiC track was approximately one month longer, related to changes in the sea ice concentration, and the summer months featured the highest temperatures and IWV compared to the climatological record (Rinke et

al. 2021).

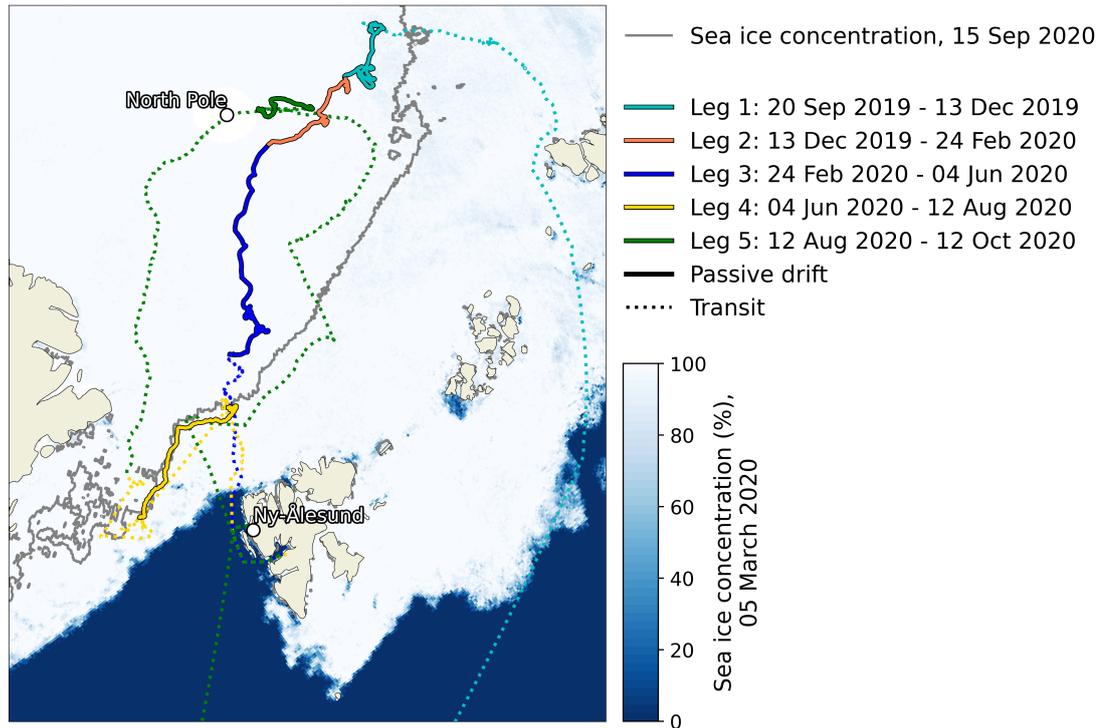


Figure 1.3: Track of RV *Polarstern* in transit (dotted lines) and when passively drifting with the sea ice (solid lines) during the MOSAiC expedition, separated into several 'legs' (colours). Shading shows the sea ice concentration from the Advanced Microwave Scanning Radiometer (AMSR, Spreen et al., 2008) for 05 March 2020 and the grey line indicates the 15% sea ice concentration on 15 September 2020.

During MOSAiC, extensive humidity measurements were performed using different approaches (Shupe et al., 2022). Near-surface measurements are available at two locations on the ship and on the sea ice in at least 300 m distance. While the instruments on board RV *Polarstern* can be influenced by turbulence around the ship's structure, the central measurement site on the sea ice (Met City) was located at a distance where the ship's influence was estimated to be negligible. Radiosondes, which were launched at the aft of RV *Polarstern* at 12 m height at least every 6 hours, measure water vapour (also temperature, pressure and wind) profiles, from which the IWV can be calculated. Remote sensing water vapour measurements are available from four MWRs and the Raman lidar Polly-XT. Two of the MWRs, a two-channel (23 and 31 GHz) and a three-channel (also includes 90 GHz) from the Atmospheric Radiation Measurement (ARM) research facility, only provide IWV (and liquid water path, LWP). The other two, the Humidity And Temperature PROfiler (HATPRO) from the Leibniz Institute for Tropospheric Research (Leipzig) and the Microwave Radiometer for Arctic Clouds - Passive (MiRAC-P) from the University of Cologne, can also be used for coarse

humidity profiling because they measure at multiple frequencies between 22 and 340 GHz. Finally, the Raman lidar Polly-XT provides humidity profiles during polar night and clear sky conditions or below cloud base height (Engelmann et al., 2016; Dai et al., 2018). The wealth of measurements from the MOSAiC expedition allows for detailed studies on inter-instrumental validations and consistency checks.

1.4 Goals of this thesis

As outlined in the previous sections, water vapour plays a crucial role in the Arctic climate system. Due to the sparsity of ground stations and challenges in satellite remote sensing, water vapour observations are highly uncertain in the central Arctic. It is especially challenging to correctly capture the vertical humidity structure, which typically features humidity inversions. In such data-sparse regions, highly accurate reference observations from field campaigns are crucial for improving operational ground-based and satellite remote sensing products and models by identifying current limitations and exploring the benefit of new methods. The importance of humidity inversions for the downwelling longwave radiation at the surface has not been sufficiently explored. Consequently, it is also unclear how misrepresentations of humidity inversions by remote sensing observations and models influence the downwelling longwave radiation at the surface. Fortunately, the MOSAiC expedition provides unique, high-quality measurements that are well-suited to address these issues and assess Arctic water vapour for the complete annual cycle.

In this thesis, new retrievals are developed using modern machine learning techniques to derive highly accurate IWV and humidity profiles (as well as LWP and temperature profiles) for the full annual cycle, covering a wide range of water vapour conditions. Radiosonde observations and synthetic evaluation data sets are used to explore the capabilities and limitations of low- and high-frequency MWRs for humidity profiling, especially regarding the ability to resolve humidity inversions. The high temporal resolution IWV product from the MWRs and the vertically detailed radiosonde observations of the MOSAiC expedition are used as reference for assessing the quality of the IWV and specific humidity profiles of state-of-the-art models and satellite products. Particular emphasis is placed on the detectability and representation of humidity inversions in these data sets, including the MWR humidity profiles. For the first time, the impact of humidity inversions on downwelling longwave radiation is quantified without the influence of clouds or temperature changes. Specifically, the following research questions (RQs) are answered:

- RQ1: What is the benefit of combining low- and high-frequency MWRs with different moisture sensitivity for humidity profiling compared to more common single-MWR measurements?
- RQ2: How well can ground-based MWRs capture the main features of the Arctic humidity profile, especially concerning humidity inversions? Can MWRs be used to correctly detect and monitor Arctic humidity inversions over long times in regions where radiosonde measurements are missing?
- RQ3: Can the reference measurements from the MOSAiC expedition be used to identify biases in the humidity profile representation in satellite products and reanalyses? What are the capabilities of space-based remote sensing for humidity profiling as assessed by MOSAiC?
- RQ4: How well do humidity inversions have to be represented to avoid biases in downwelling longwave radiation?

An overview of the studies addressing these RQs is given below. Afterwards, the microwave radiative transfer for ground-based MWRs is introduced in Chapter 2. Chapter 3 explains how atmospheric parameters are derived from measured microwave radiances. These two theory chapters are followed by the three studies (Chapters 4, 5, and 6), two of which are already published or accepted for publication. Finally, Chapter 7 answers the research questions by concluding the findings and provides an outlook for future research.

1.5 Overview of the studies

Study 1: Single microwave radiometer retrievals

The first study (Chapter 4, Walbröl et al., 2022) is the first step towards a high-accuracy reference IWV data set for the central Arctic, using the measurements of the low-frequency Humidity and Temperature Profiler (HATPRO) and the high-frequency Microwave Radiometer for the study of Arctic Clouds - Passive (MiRAC-P) during the MOSAiC expedition. The measurements of the MWRs were processed and quality controlled in preparation to derive atmospheric parameters. Standardized regression retrievals were used to derive coarse temperature and humidity profiles, as well as IWV and LWP from the HATPRO radiance measurements. For MiRAC-P, the standardized regression retrievals did not sufficiently capture the nonlinear relationship between the radiance measurements at high frequencies and IWV. To better capture the nonlinear relationship, we developed a new IWV retrieval using Neural Networks. The IWV, as well as the

humidity and temperature profile retrievals were evaluated using the radiosonde observations from the MOSAiC expedition. The errors of the retrieved humidity and temperature profiles are generally similar to findings from earlier studies but face challenges in the lower troposphere due to inversions. Regarding IWV, the evaluation showed that HATPRO and MiRAC-P perform well in a complementary sense. However, none of the IWV retrievals were accurate over the entire range of IWV conditions. This study revealed the potential for improved water vapour products by combining the low- and high-frequency measurements from HATPRO and MiRAC-P.

Study 2: Synergistic microwave radiometer retrievals

The findings of the first study and Cadeddu et al. (2022) motivated us to combine HATPRO and MiRAC-P measurements in a synergistic approach for enhanced IWV and humidity profiles (Chapter 5, Walbröl et al., 2024). With the experience from the first study, we developed new retrievals of IWV, specific humidity profiles and temperature profiles using Neural Networks. The synergistic retrievals were evaluated with a synthetic data set and MOSAiC radiosondes. To quantify the benefit of the synergy compared to single-MWR retrievals from the first study (addressing RQ1), the errors of the retrievals were compared. Additionally, the vertical information content and vertical resolution were computed for the humidity profiles of the synergy and the single-MWR retrievals.

Study 3: Evaluation of water vapour products and assessing the importance of humidity inversions

The third study addresses RQ2 to RQ4 (Chapter 6). The IWV from the MWR synergy from the second study and radiosonde observations were used as reference to evaluate state-of-the-art models and satellite products. Root mean squared deviations and biases were computed to quantify the performance of the water vapour products of all data sets to answer parts of RQ3. One part of the evaluation focused on the detectability of humidity inversions and the representation of their characteristics in the models, satellite observations and the ground-based MWRs. Comparing the inversion characteristics of the radiosondes with those of all other data sets allowed us to assess how well humidity inversions are represented, addressing RQ2 and RQ3. RQ4 is answered in three steps: Firstly, radiative transfer simulations were used to estimate the effect of humidity inversions at different altitudes on the downwelling longwave radiation. The findings help to get a sense of the heights at which humidity inversions are relevant for downwelling longwave radiation. Secondly, the radiative effect of humidity

inversions was quantified without the influence of clouds or temperature changes. Thirdly, the impact of different representations of humidity inversions on downwelling longwave radiation was analyzed by using the humidity profiles of the different data sets in the radiative transfer simulations. With the results from the simulations, we estimated to what extent the vertical water vapour distribution, characterized by inversions, affects the downwelling longwave radiation compared to the IWV.

Chapter 2

Microwave radiative transfer

The electromagnetic radiation emitted, absorbed and scattered by the Earth's atmosphere and surface can be used for remote sensing of meteorological parameters. In this thesis, the focus lies on the microwave spectrum (wavelengths λ : 1–100 mm, frequencies ν : 3–300 GHz) and parts of the sub-millimetre spectrum ($\lambda < 1$ mm, $\nu > 300$ GHz). In contrast to the visible, ultraviolet or (near and thermal) infrared spectrum, clouds are semitransparent in the microwave spectrum, allowing instruments to sense the entire atmosphere even in cloudy conditions. In this chapter, an overview of the interaction of microwave radiation with the atmosphere from a ground-based perspective is provided, following the works of Liou (2002) with some additional remarks from Janssen (1993), Petty (2006) and Foken (2021), to understand what a microwave radiometer (MWR) measures and which meteorological parameters can be derived. The MWR only measures atmospheric radiation and does not emit a beam itself (as a radar does). For this reason, this type of observation is often called passive microwave remote sensing.

2.1 Emission and absorption

Electromagnetic radiation is emitted when an atom or molecule changes from one of its higher energy states (quantized energy states) to a lower state given the energy difference $\Delta E = \Delta n h \nu$, where Δn is the change of the quantum number, h the Planck constant and ν the frequency of the emitted radiation. The idealized perfect absorber and emitter of electromagnetic radiation is known as black body. When a black body is in thermodynamic equilibrium with its surroundings at temperature T , it emits spectral radiance (or radiant energy intensity) at a certain frequency according to Planck's law (Liou, 2002, Chapter 1.2)

$$B_\nu(T) = \frac{2h\nu^3}{c^2 \left(e^{\frac{h\nu}{k_B T}} - 1 \right)}, \quad (2.1)$$

where c and k_B are the velocity of light in vacuum and the Boltzmann constant, respectively. The spectral radiance is given in units of energy per time, area, solid angle and frequency ($\text{W m}^{-2} \text{sr}^{-1} \text{s}$). In thermodynamic equilibrium, an emitting body absorbs the same amount of radiation for a given frequency according to Kirchhoff's law (emissivity ϵ_ν equals absorptivity α_ν , Kirchhoff, 1860). If the emissivity or absorptivity is less than one, the absorbing or emitting body is referred to as a grey body, which emits less spectral radiance ($\epsilon_\nu B_\nu(T)$) than a black body at the same temperature and frequency.

Similar to emission, absorption occurs when incident radiation, where each photon has a certain energy $E = h\nu$, changes the quantized energy state of an atom or a molecule to a higher level (from a ground state to an excited state). In the microwave spectrum, the main absorbing gases are oxygen (O_2) and water vapour (H_2O), and energy state changes are realized by rotational transitions. Due to the structure of H_2O (isosceles triangle between the hydrogen and oxygen atoms), the molecule has three different moments of inertia I_i ($i \in \{0, 1, 2\}$) and a permanent electric dipole moment (Liou, 2002, Chapter 3.2). The diatomic oxygen molecule O_2 is linear and therefore has two equal moments of inertia and negligible inertia along the axis connecting the oxygen atoms. Due to symmetric charge distributions, the oxygen molecule has no permanent electric dipole moment but a weak magnetic dipole moment because of two unpaired electrons. Either electric or magnetic dipole moments are needed for rotational transitions (Petty, 2006, Chapter 9.2). Otherwise, an incident electromagnetic wave cannot apply torque to change the angular momentum of the molecule.

In the following, the energy levels of rotational transitions are outlined for a linear rigid rotor, which can be used as an approximation for O_2 : During rotational transitions, the rotational energy state, governed by the angular momentum L of the molecule, is changed. The angular momentum and moments of inertia are related by $L = I\omega$, where ω is the angular velocity. For the rotating molecule, the angular momentum takes discrete values according to the quantized energy states

$$L = \frac{h}{2\pi} \sqrt{J(J+1)}, \quad (2.2)$$

where J is the rotational quantum number. Using Eq. 2.2 and $\omega = L/I$, the energy $E = L\omega/2$ of the rotating dipole for a certain rotational quantum number is then given by

$$E_J = BhJ(J+1), \quad (2.3)$$

with the rotational constant $B = h/(8\pi^2 I)$ (Janssen, 1993, Chapter 2.2). The molecule mostly transitions from one energy state to the next higher or lower

state ($\Delta J = \pm 1$). For example, to excite the molecule from the lower energy state E_J to a higher energy state E_{J+1} , the frequency of the incident photon must be

$$\nu = \frac{E_{J+1} - E_J}{h} = \frac{Bh(J+1)(J+2) - BhJ(J+1)}{h} = 2B(J+1). \quad (2.4)$$

Thus, resonant absorption lines for this rigid rotor example are separated by $\Delta\nu = 2B$. The absorption lines are generally not monochromatic because of natural, pressure (or collisional) and Doppler broadening (Petty, 2006, Chapter 9.3). Natural broadening occurs due to Heisenberg's uncertainty principle and the finite time a molecule stays in an excited state before returning to a lower energy state and emitting radiation (Janssen, 1993, Chapter 2.1). The energy of the emitted photon is uncertain according to Heisenberg's uncertainty principle, broadening the frequency spectrum of the emitted photon. Pressure broadening is caused by collisions between molecules, which can modify the natural time the molecule is in the excited state before returning to its ground state. Doppler broadening is observed when molecules have a velocity component in the direction of observation, causing Doppler frequency shifts. The direction of motion of a molecule is arbitrary, and its velocity depends on its temperature. Due to the high abundance of molecules in the troposphere, pressure (or collisional) broadening dominates. All three effects together are often referred to as line broadening.

For this thesis, the essential resonant absorption lines are those for water vapour at 22.235 GHz and 183.31 GHz, and those for oxygen around 60 GHz (Liou, 2002, Chapter 7.5). Figure 2.1 shows the transmissivity of the atmosphere in the microwave spectrum from a ground-based perspective. The transmissivity of the atmosphere indicates the fraction of radiation intensity reaching the sensor (here, at the surface) to the radiation intensity at the top of the atmosphere. Section 2.3 provides a formal definition of the transmissivity. Due to line broadening, numerous oxygen absorption lines at 60 GHz overlap, forming a strong absorption complex with high atmospheric opacity and, therefore, low transmissivity (Fig. 2.1). Frequency ranges with low opacity have a high transmissivity and are called atmospheric windows. Atmospheric windows are found at low frequencies between resonant absorption lines (e.g., at about 30 GHz). The transmissivity decreases with frequency as the opacity increases due to the water vapour continuum absorption, whose origin is not fully explored. The water vapour continuum absorption has been related to the formation of H₂O molecule clusters and to broadening due to collisions of water molecules with other water molecules or other gases (nitrogen, oxygen, etc.) (Rosenkranz, 1998; Turner et al., 2009). Due to the dependence of the water vapour continuum absorption

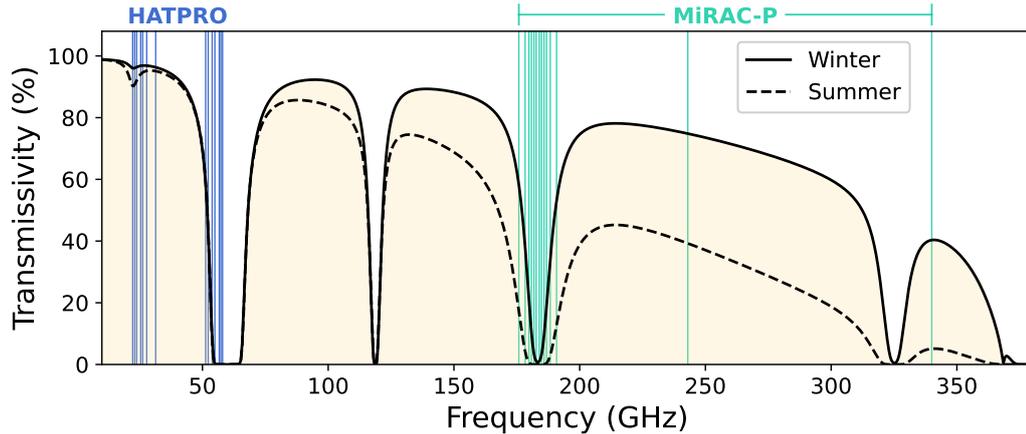


Figure 2.1: Calculated transmissivity of microwave radiation between 10 and 380 GHz simulated with MOSAiC radiosondes averaged over summer (01 May–31 August 2020, integrated water vapour: 13.3 kg m^{-2}) and winter (01 October 2019–30 April 2020 and 01–30 September 2020, integrated water vapour: 3.9 kg m^{-2}). Blue (green) lines indicate the frequencies of the microwave radiometer HATPRO (MiRAC-P).

on the amount of water vapour, the atmospheric transmissivity is low at frequencies above 300 GHz in humid conditions (Petty, 2006, Chapter 7.4). Figure 2.1 illustrates that the transmissivity at high frequencies is much lower in the Arctic summer than in winter because of the higher water vapour amount. However, as the transmissivity is not 0, remote sensing of the entire atmosphere is still possible.

As oxygen is a well-mixed gas whose concentration is known, temperature profiles can be derived by combining observations at frequencies close to the centre of the oxygen absorption complex (low transmissivity) with frequencies at the wings (high transmissivity). Similarly, observations along the water vapour absorption lines can be used to derive humidity profiles and the integrated water vapour (IWW). As the 183.31 GHz absorption line is much stronger than the 22.235 GHz line (see also transmissivity in Fig. 2.1), the former is favoured for ground-based microwave remote sensing in dry conditions (e.g., Kerber et al., 2014). At window frequencies, where the resonant absorption by water vapour and oxygen is weak, the emission from liquid hydrometeors (e.g., cloud droplets) is pronounced. The emission strength also increases with frequency. By combining observations at frequencies where liquid emission dominates (e.g., 30 GHz) with those where water vapour absorption dominates (e.g., the 22.235 GHz resonant absorption line), the total amount of cloud liquid water (liquid water path, LWP) can be derived (Crewell et al., 2001). The observation at a frequency with strong water vapour absorption is needed to account for the water vapour emission signal when deriving LWP.

2.2 Scattering

Scattering is another form of interaction of electromagnetic radiation with matter, where the radiation is redirected. The scattering properties of a particle strongly depend on its relative size to the wavelength of the radiation. This relationship is expressed by the size parameter $X = 2\pi a/\lambda$, where a is the radius of the particle (Liou, 2002, Chapter 1.1). Typically, three forms of scattering are distinguished:

- Rayleigh scattering: $a \ll \lambda \implies X \ll 1$,
- Mie scattering: $a \simeq \lambda \implies X \gtrsim 1$,
- Geometric optics: $a \gg \lambda \implies X \gg 1$.

In the frequency range considered in this thesis (20–350 GHz), only Rayleigh and Mie scattering are relevant and have been respected in the radiative transfer simulations to prepare the retrievals of atmospheric parameters (see Chapters 4 and 5). At low frequencies (< 100 GHz) and in the absence of large precipitating hydrometeors (rain, graupel, snow or hail), scattering is negligible compared to absorption by gases and hydrometeors (Janssen, 1993, Chapter 1.2 and 3.2). In the Mie scattering regime (higher frequencies and/or larger hydrometeors), scattering dominates the total extinction of radiation (extinction includes absorption and scattering) (Petty, 2006, Chapter 12.3).

In the following, Rayleigh scattering for a spherical particle will be briefly outlined to understand the underlying processes of scattering. As the wavelength of the incident radiation is significantly larger than the scattering particle, it can be assumed that the electric field of the incident radiation is homogeneous throughout the particle, which may have an electric field by itself. Due to the incident radiation, all dipoles within the scattering particle experience torque, realigning the dipoles along the electric field. Thus, as the electric field of the radiation propagates over time, the dipoles in the particle oscillate at the frequency of the incident radiation. The acceleration of the particle's oscillating dipole moments produces the electric field of the scattered radiation. The squared absolute value of the electric field then yields the intensity of the scattered radiation. If the incident radiation is unpolarized, the intensity of the Rayleigh-scattered radiation is relatively uniform in all directions around the scattering particle with slight maxima in the forward (propagation direction of the incident radiation) and backward direction. Thus, an upward-looking ground-based sensor (e.g., MWR) can also detect radiation from particles and molecules not directly in the line of sight. However, in the Rayleigh scattering regime, the contribution of scattering is small compared to the emission signal (Petty, 2006, Chapter 12.3).

When approaching the Mie scattering regime, scattering can no longer be neglected as the extinction efficiency of a particle gets more and more dominated by scattering compared to absorption. As the wavelength approaches the size of the scattering particles, the electric field inside the particle cannot be assumed to be homogeneous. Therefore, the Rayleigh theory no longer applies, and the intensity of the scattered radiation is computed using Mie theory instead. In contrast to Rayleigh scattering, radiation is scattered much more in the forward direction (Petty, 2006, Chapter 12.3). Thus, for an upward (zenith) looking ground-based sensor, the downwelling microwave radiation emitted by gases and hydrometeors is rather scattered in the downward direction than in the horizontal or upward direction. The scattering intensity highly depends on the frequency of the radiation, as well as the size, shape, density, orientation and complex index of refraction of the particle (e.g., Hong et al., 2009; Foken, 2021, Chapter 41.3).

2.3 Radiative transfer equation

The radiative transfer equation explains how (microwave) radiation propagates through the atmosphere. Here, the radiative transfer equation will be derived for an upward-looking ground-based sensor (MWR), assuming a plane-parallel atmosphere (horizontally homogeneous). However, we start from the perspective of the downwelling radiation, looking down towards the surface from a certain height in the atmosphere. Absorbing (and therefore emitting) and scattering particles in the atmosphere change the intensity I_ν of radiation at each distance step ds according to

$$\frac{dI_\nu}{ds} = -S + O. \quad (2.5)$$

While absorption is a clear sink S and emission O is a source of radiation, scattering can be both, depending on whether radiation is scattered towards (source) or away (sink) from the direction of the observer. For simplification, low microwave frequencies are considered here so that scattering can be disregarded. Equation 2.5 can therefore be written as

$$\frac{dI_\nu}{ds} = \beta_{a,\nu} (-I_\nu + B_\nu(T)), \quad (2.6)$$

forming Schwarzschild's equation, where B_ν is the emitted radiation according to Planck's law and $\beta_{a,\nu}$ is the volume absorption coefficient (Petty, 2006, Chapter 8.1). The volume absorption coefficient can be rewritten in terms of the mass absorption coefficient $k_{a,\nu}$ and the density of the absorbing medium ρ : $\beta_{a,\nu} = k_{a,\nu}\rho$. Note that the density $\rho(z)$ of the absorbing medium is only known for well-mixed

gases such as oxygen or carbon dioxide. In contrast, the water vapour density is unknown and has a high spatio-temporal variability. Here, the absorbing medium is air (dry air combined with water vapour), where water vapour and oxygen contribute most to the absorption. The mass absorption coefficient is obtained by summing over the product of all absorption line strength and shape functions, which depend on the pressure and temperature (Liou, 2002, Chapter 4.2). As the shape function depends on the pressure (line broadening), which depends on the height according to the hydrostatic equation, the absorption coefficient also varies with height.

As radiation propagates through the atmosphere, it is continuously attenuated through absorption (scattering is neglected). Beer-Bouguer-Lambert's law describes that an initial radiation intensity $I_\nu(0)$ decreases exponentially with distance, given by

$$I_\nu(s_1) = I_\nu(0)e^{-\int_0^{s_1} k_{a,\nu}(s)\rho(s) ds}, \quad (2.7)$$

where $s_1 > 0$. The term $\tau(s) = \int_s^{s_o} k_{a,\nu}(s)\rho(s) ds$ is often referred to as the optical depth or optical thickness and is a measure of the absorption strength between any point s and the observer s_o (Chandrasekhar, 1960, Chapter 1.7). Therefore, the optical depth is 0 at the position of the observer. Further, $t(s) = e^{-\tau(s)}$ is known as the transmissivity, which describes to what fraction the radiation intensity is reduced (by absorption) between s_o and s . Over an infinitesimal distance ds , the absorption coefficient and density can be considered constant. Using the differential form of the optical depth $d\tau = -k_{a,\nu}\rho ds$ and multiplying Eq. 2.6 with $e^{-\tau}$ yields

$$-k_{a,\nu}\rho \frac{dI_\nu}{d\tau} e^{-\tau} = k_{a,\nu}\rho (-I_\nu + B_\nu(T)) e^{-\tau}. \quad (2.8)$$

Note that the mass absorption coefficient and the density have replaced the volume absorption coefficient. Rearrangement and integration over $d\tau$ with integration limits 0 and τ' yields

$$I_\nu(0) = I_\nu(\tau')e^{-\tau'} + \int_0^{\tau'} B_\nu(T)e^{-\tau} d\tau. \quad (2.9)$$

For the zenith-looking ground-based MWR, the distance ds of the initially considered perspective (looking down from a given height) corresponds to the height difference $-dz$. Thus, the optical depth is now given by $\tau(z) = \int_{z_o}^z k_{a,\nu}(z)\rho(z) dz$, where z_o is the altitude of the MWR. The upper integration limit τ' is now chosen to be the top of the atmosphere z_{TOA} so that $\tau' = \tau(z_{\text{TOA}})$. In the microwave spectrum, measured radiation intensities are often expressed in terms of the brightness

temperature T_b , which would be the physical temperature of a black body if it emitted the same radiation that was measured. For low frequencies, Planck's law can be simplified to a term that linearly depends on the temperature using the Rayleigh-Jeans approximation ($h\nu \ll k_B T$), and the radiation intensity I_ν can be similarly scaled (Liou, 2002, Chapter 7.5):

$$B_\nu(T) \approx \frac{2k_B\nu^2}{c^2}T, \quad I_\nu \equiv \frac{2k_B\nu^2}{c^2}T_{b,\nu} \quad (2.10)$$

At frequencies above 200 GHz, the Rayleigh-Jeans approximation produces errors of a few percent (Janssen, 1993, Chapter 1.2) and the inverse of Planck's law should be used instead (Petty, 2006, Chapter 6.2). Using $d\tau = k_{a,\nu}\rho dz$ and the Rayleigh-Jeans approximation, Eq. 2.9 can be written as

$$T_{b,\nu}(z_o) = T_{b,\nu}(z_{\text{TOA}})e^{-\tau(z_{\text{TOA}})} + \int_{z_o}^{z_{\text{TOA}}} T(z)k_{a,\nu}(z)\rho(z)e^{-\tau(z)} dz. \quad (2.11)$$

This equation describes what a ground-based microwave radiometer observes when scattering is neglected. The first term on the right-hand side of Eq. 2.11 is the cosmic background radiation $T_{b,\text{cos}} = 2.73$ K, which is attenuated while propagating through the entire atmosphere according to Beer-Bouguer-Lambert's law (Foken, 2021, Chapter 29.3). At frequencies and in atmospheric conditions where the transmissivity is high (e.g., around 30 GHz), the contribution of the cosmic background radiation is significant. The second term describes the emission of radiation from an atmospheric layer at any height between the surface (z_o) and the top of the atmosphere, which is also attenuated by absorption before reaching the MWR at z_o . The information to derive, for example, the water vapour profile, is hidden in the second term, which can be rewritten as

$$k_{a,\nu}(z)\rho(z)e^{-\tau(z)} = k_{a,\nu}(z)\rho(z) \overbrace{e^{-\int_0^z k_{a,\nu}(z')\rho(z') dz'}}^{t(z)} = -\frac{dt(z)}{dz} \equiv W(z). \quad (2.12)$$

$W(z)$ is the so-called weighting function for temperature, which depends on the density of the absorbing medium (e.g., air) and the absorption coefficient. The weighting function indicates how the transmissivity t changes with each height layer dz . If the transmissivity, which is always 1 at the sensor, strongly decreases in a height layer, radiation is strongly absorbed and emitted. Conversely, if the transmissivity is constant over that height layer ($W(z) = 0$), radiation is neither absorbed nor emitted.

The weighting function thus indicates the absorption rate within a certain atmospheric height layer at a given frequency and provides information about

the height from which an emission signal originates. Figure 2.2 illustrates the weighting functions for temperature for the frequencies of the two ground-based microwave radiometers HATPRO and MiRAC-P. From a ground-based perspective, the weighting functions are highest near the surface. For a frequency near the centre of an absorption line (e.g., around 58 GHz for oxygen or 183.31 GHz for water vapour), the absorption coefficient is high, resulting in a strong peak of the weighting function near the surface (e.g., Fig. 2.2b, c). Away from the centre of an absorption line, the absorption coefficient is smaller and the transmissivity higher, resulting in higher weighting function values at higher altitudes. For example, at altitudes above 1.5 km, the weighting function at 190.81 GHz is stronger than the weighting function at 183.81 GHz (Fig. 2.2c). Thus, the frequency away from the absorption line is more sensitive to water vapour absorption above 1.5 km than the frequency in the centre of the absorption line.

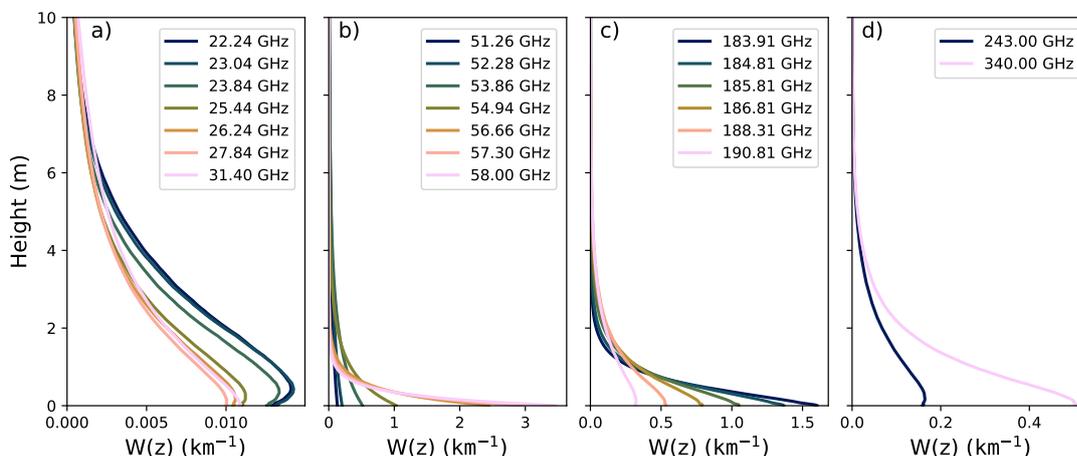


Figure 2.2: Calculated weighting functions for temperature $W(z)$ for the frequencies of the microwave radiometers HATPRO (a–b) and MiRAC-P (c–d) based on the mean of the MOSAiC radiosondes between 01 October 2019 and 30 September 2020.

If weighting functions were narrow peaks and did not overlap with height for adjacent frequencies (e.g., along an absorption line wing), the average density of the absorbing medium over the height range of the weighting function could be directly estimated (Petty, 2006, Chapter 8.3). However, in a real atmosphere, weighting functions are broad, which means that measurements at adjacent frequencies are highly correlated and contain similar information. Consequently, the information content of microwave observations is smaller than the number of frequency channels of an instrument. For example, the microwave observations at the seven frequency channels between 22.24 and 31.4 GHz show broad weighting functions (Fig. 2.2a), which provide only 1–3 independent pieces of information for humidity profiles (Löhnert et al., 2009). Extracting an absorber density pro-

file from correlated observations is not straightforward. The challenges of this problem and the theory of the methods to overcome them are presented in the following chapter.

Chapter 3

Retrieval theory

Chapter 2 explained that passive microwave observations contain information on atmospheric parameters, such as temperature and humidity profiles, liquid water path, and integrated water vapour. This chapter focuses on the methods to infer atmospheric parameters from the information hidden in the highly correlated microwave observations. The derivation of atmospheric parameters from (microwave) remote sensing observations is often referred to as retrieval.

3.1 Inverse modelling problem

Deriving a state vector \mathbf{x} (e.g., humidity profile) from measurements \mathbf{y} (e.g., brightness temperatures) is known as the inverse problem. The inverse problem is the opposite of the forward problem where a state vector is used to simulate measurements. The forward problem is formally defined as $\mathbf{y} = F(\mathbf{x}, \mathbf{b}) + \epsilon$, where F is the forward model (e.g., radiative transfer equation) based on the physical understanding and ϵ is the measurement error vector (Rodgers, 2000, Chapter 2.1). \mathbf{b} contains additional parameters that affect the measurements but are not derived in the inverse problem.

In the forward problem, one set of measurements is simulated for a given atmospheric state. The simulated measurements agree with the true measurements within the range of the measurement and model errors. Therefore, within an uncertainty range, it is known what measurements to expect for a given atmospheric state. However, the inverse case is not as trivial in passive microwave remote sensing because one set of highly correlated brightness temperature measurements (e.g., at 7 frequency channels) can generally be caused by an infinite amount of atmospheric states. As the atmospheric state is generally a continuous function, the inverse problem is under-determined and possesses either no or non-unique solutions (Rodgers, 2000, Chapter 2.1). Usually, the state vector is discretized to a height grid with N levels. If the number of unknown state vector

elements N was smaller than the number of measurements M the problem would no longer be under-determined (except for mixed-determined conditions, which are not discussed here). However, due to the weighting function overlap discussed in the previous chapter, the M measurements provide less than M independent pieces of information (Petty, 2006, Chapter 8.3). It is therefore more accurate to state that the number of independent pieces of information in the measurements must be $\geq N$ to avoid the problem being under-determined. One exact solution for the inverse problem only exists if $M = N$ (well-determined) and the measurements also provide M independent pieces of information (Rodgers, 2000, Chapter 2.2).

In real passive microwave remote sensing applications, the inverse problem is typically under-determined and additionally ill-conditioned, meaning that measurement errors can be greatly amplified in the derived state vector (Rodgers, 2000, Chapter 1.3). In this thesis, statistical retrievals are used to overcome the inverse problem. Statistical retrievals require large data bases to derive coefficients that relate measurements (brightness temperatures) to the atmospheric state (e.g., humidity profile). To perform well, the data base should be representative of the conditions of the region where the retrieval is applied (here, Arctic Ocean). Here, the coefficients are obtained via the least squares approach as shown for the regression in Sect. 3.2 and by training Neural Networks as presented in Sect. 3.3.

Physical approaches, such as Optimal Estimation (Rodgers, 2000, Chapters 4 and 5), yield accurate and physically consistent solutions to the inverse problem and provide error estimates. This thesis does not consider such retrievals because modern statistical approaches provide similarly accurate results when trained carefully while being computationally much cheaper (Solheim et al., 1998). Due to the low computational costs, statistical retrievals are also easier to apply to other sites (especially for continuous operation) with the same instrumentation and similar atmospheric conditions. Additionally, this thesis includes microwave observations at frequencies where scattering from ice particles is relevant. Thus, the retrieval would require assumptions on the ice crystal properties (shape, size, concentration, ...) or further observations. Assumptions about the ice crystal properties could increase retrieval errors due to the ill-conditioned nature of the inverse problem.

3.2 Regression

When the inverse problem is linear or nearly linear and a large data base (n_s data samples) is available, regression coefficients between the observations and

each state vector element can be derived, and yield accurate results. Slight non-linearities between observations and the state can be handled by adding non-linear terms to the regression (Rose et al., 2005). In this thesis, multiple regression with linear and quadratic terms is used. The i -th element (e.g., i -th height level) and k -th sample of the state vector $x_{i,k}$ ($i = 1, 2, \dots, N$, $k = 1, 2, \dots, n_s$) can be related to the observations $y_{l,k}$ ($l = 1, 2, \dots, M$) as follows:

$$x_{i,k} = a_i + \sum_{l=1}^M b_{i,l} y_{l,k} + \sum_{l=1}^M c_{i,l} y_{l,k}^2 + \epsilon \quad (3.1)$$

ϵ is the error between the true state (left side of Eq. 3.1) and the predicted state (right side of Eq. 3.1 without the error term). Thus, we have n_s equations to solve $2M + 1$ unknown coefficients ($a_i, b_{i,1}, \dots, b_{i,M}, c_{i,1}, \dots, c_{i,M}$) for the i -th element of the state vector. If the data base is sufficiently large $n_s \geq 2M + 1$, the inverse problem is formally over-determined. Usually, $n_s \gg 2M + 1$ to improve the accuracy over a wider range of (atmospheric) conditions and to be less susceptible to noise (Rodgers, 2000, Chapter 6.2). The solution would only be exact if neither the observations nor the state vectors of the data base had any errors or uncertainties (Rodgers, 2000, Chapter 2.2). For example, two identical atmospheric states can result in two slightly different sets of observations due to measurement errors.

Equation 3.1 must be solved for each data sample k and element i of the state vector, resulting in a total of $n_s \times N$ equations to obtain the $N(2M + 1)$ regression coefficients. For convenience, the equations can be summarized over all data samples using matrix notation

$$\mathbf{x}_i = \mathbf{G}\mathbf{m}_i + \epsilon, \quad (3.2)$$

where $\mathbf{m}_i = (a_i, b_{i,1}, \dots, b_{i,M}, c_{i,1}, \dots, c_{i,M})^T$ contains the regression coefficients and \mathbf{G} is the sensitivity of the state vector to the coefficients for element i ($\partial \mathbf{x}_i / \partial \mathbf{m}_i$) (Oliver et al., 2008, Chapter 3.1). More specifically, \mathbf{G} is a $n_s \times (2M + 1)$ matrix that contains the measurements

$$\mathbf{G} = \begin{pmatrix} 1 & y_{1,1} & \cdots & y_{M,1} & y_{1,1}^2 & \cdots & y_{M,1}^2 \\ \vdots & & & & & & \\ 1 & y_{1,k} & \cdots & y_{M,k} & y_{1,k}^2 & \cdots & y_{M,k}^2 \\ \vdots & & & & & & \\ 1 & y_{1,n_s} & \cdots & y_{M,n_s} & y_{1,n_s}^2 & \cdots & y_{M,n_s}^2 \end{pmatrix}. \quad (3.3)$$

In general, the error term $\epsilon = \mathbf{x}_i - \mathbf{G}\mathbf{m}_i$ in Eq. 3.2 is large without the trained

coefficients (e.g., using random values for each coefficient). Thus, to obtain useful regression coefficients, the squared error between the true state vector \mathbf{x}_i and the predicted one $\mathbf{G}\mathbf{m}_i$ is minimized using

$$\nabla_m \left((\mathbf{x}_i - \mathbf{G}\mathbf{m}_i)^T (\mathbf{x}_i - \mathbf{G}\mathbf{m}_i) \right) = 0, \quad (3.4)$$

where $\nabla_m = (\partial/\partial m_{i,1}, \partial/\partial m_{i,2}, \dots, \partial/\partial m_{i,2M+1})^T$ is the gradient with respect to the regression coefficient vector \mathbf{m}_i (Oliver et al., 2008, Chapter 3.2). This approach is known as the least squares approach as it minimizes the squared errors. Solving Eq. 3.4 yields

$$\mathbf{m}_i = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{x}_i \quad (3.5)$$

for the i -th element of the state vector if $\mathbf{G}^T \mathbf{G}$ is non-singular (Oliver et al., 2008, Chapters 3.1 and 3.2), which is usually the case. Once all $N(2M+1)$ coefficients have been determined, Eq. 3.2 can be applied to unseen data by inserting the new observations in \mathbf{G} to obtain the new predicted state vector \mathbf{x} .

3.3 Neural Networks

With increasing computational power and simpler implementation methods (e.g., via the Python libraries *Keras* or *tensorflow*, Chollet et al., 2015; Abadi et al., 2015), Neural Networks (NNs) have become a popular tool in science for a multitude of problems (e.g., pattern recognition, classification, clustering, regression, etc., Aires et al., 2001; Chatterjee et al., 2023; Jozef et al., 2024). The idea of a NN is to imitate the neurons of the human brain, which are connected with a certain strength to exchange information (Haykin, 1998; Mas and Flores, 2008). In a NN, weights w represent the strength of the neuron connection. NNs can capture non-linearities much better than linear or quadratic regression (see Sect. 3.2). In fact, Tang and Yang (2021) showed that NNs can approximate any continuous function to a given accuracy if the network has a sufficient number of neurons.

This thesis uses multilayer perceptron NNs (Haykin, 1998) to solve the regression problems of deriving atmospheric parameters from brightness temperature measurements. A multilayer perceptron consists of an input layer, at least one hidden layer and an output layer. An example with two hidden layers is illustrated in Fig. 3.1. If the NN has more than one hidden layer, it is usually called a deep NN. Each layer consists of a certain number of neurons (e.g., a_3^2 in Fig. 3.1), which are connected to all neurons of the next and previous layer by the weights w , if the NN is fully connected (dense). The number of neurons in the input

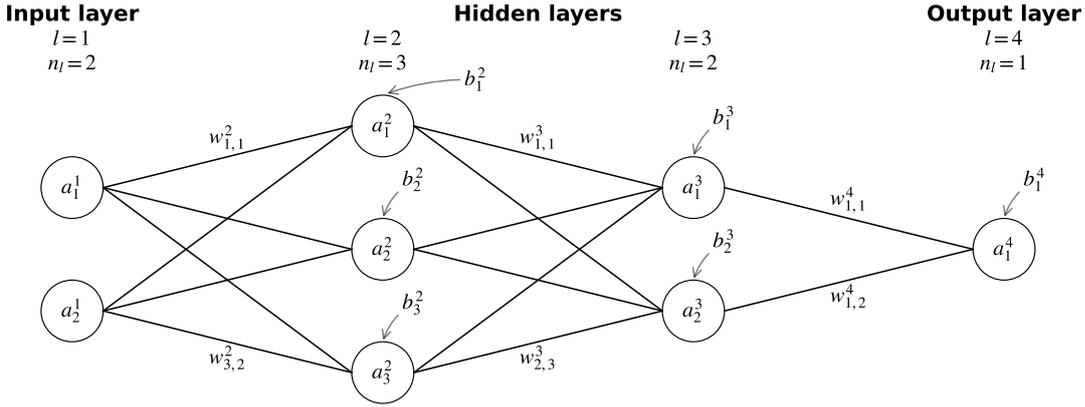


Figure 3.1: Schematic of a multilayer perceptron with an input layer consisting of 2 neurons, two hidden layers with 3 and 2 neurons, respectively, and an output layer with 1 neuron. The weight $w_{j,k}^{l+1}$ connects neuron a_k^l of layer l with neuron a_j^{l+1} of layer $l+1$, which is also influenced by the bias value b_j^{l+1} (see text). n_l is the number of neurons of layer l .

layer corresponds to the size of the input vector (e.g., brightness temperatures at different frequencies), while the output layer has the size of the desired state vector.

The NN learns the relationship between the input (measurement vector \mathbf{y}) and output (state vector \mathbf{x}) by iterating through the training data set (e.g., a radiosonde climatology with simulated brightness temperatures) and updating the network parameters w and b . The network parameters are initialized with random values. Each training data sample (e.g., one sounding with corresponding brightness temperatures) is then processed by the NN in the forward direction (input \rightarrow output) to estimate how well the computed output $f(\mathbf{y})$ (predicted state vector) compares to the true output \mathbf{x} (true state vector). This process is referred to as the feedforward pass.

Consider a multilayer perceptron with L layers, including the input and output layers. The k -th element of the input vector (measurement) y_k is provided to the neuron a_k^1 of the input layer and is forwarded to the next layer using a weighted sum of the input and an additional bias term b (Fig. 3.1). a_k^l represents the information stored in neuron k for layer l . The propagation from layer l to the next layer $l+1$ is given by

$$z_j^{l+1} = \sum_{k=1}^{n_l} w_{j,k}^{l+1} a_k^l + b_j^{l+1} \quad (3.6)$$

$$a_j^{l+1} = \sigma(z_j^{l+1}), \quad (3.7)$$

where $w_{j,k}^{l+1}$ connects neuron k of layer l with neuron j of layer $l+1$ and n_l is the

number of neurons of layer l (using the notation style of Tang and Yang, 2021). b_j^{l+1} is the bias term of neuron a_j^{l+1} (Fig. 3.1). The result of the weighted sum and the addition of the bias term is forwarded to an activation function σ , which is used to increase the non-linearity of the network or limit the value of a neuron to a certain value. Once the feedforward pass reaches the output layer $l = L$ for the training sample i , the error \mathcal{L} (or loss) between the truth x_j and prediction $f(\mathbf{y})_j = a_j^L$ is computed, for example, using

$$\mathcal{L}(i) = \sum_{j=1}^N (x_j(i) - a_j^L(i))^2 \equiv \sum_{j=1}^N (\epsilon_j(i))^2, \quad (3.8)$$

where N is the number of state vector elements (Haykin, 1998, Chapter 4.3). The error information of the feedforward pass is used to update the weights and biases of the NN in the backpropagation pass (input \leftarrow output). The sensitivity of the weights and biases to the loss is given by

$$\frac{\partial \mathcal{L}}{\partial w_{j,k}^l} = \delta_j^l \frac{\partial \sigma(z_j^l)}{\partial z_j^l} a_k^{l-1} \quad \text{and} \quad (3.9)$$

$$\frac{\partial \mathcal{L}}{\partial b_j^l} = \delta_j^l \frac{\partial \sigma(z_j^l)}{\partial z_j^l} \quad \text{with} \quad (3.10)$$

$$\delta_j^l = \begin{cases} 2\epsilon_j & \text{if } l = L \\ \sum_{k=1}^{n_{l+1}} \delta_k^{l+1} \frac{\partial \sigma(z_k^{l+1})}{\partial z_k^{l+1}} w_{k,j}^{l+1} & \text{else} \end{cases} \quad (3.11)$$

with $l = 2, 3, \dots, L$ (Haykin, 1998, Chapter 4.3). Note that the training data sample index i (learning step) has been omitted for clarity. The factor of 2 in Eq. 3.11 if $l = L$ originates from the derivative of the loss function.

The goal of the NN training procedure is to minimize the loss function \mathcal{L} . The minimization is usually realized by gradient descent methods (and optimizations thereof) so that the step to update the weights and biases after processing training sample i is given by

$$\Delta w_{j,k}^l(i) = -\eta \frac{\partial \mathcal{L}(i)}{\partial w_{j,k}^l(i)} + \alpha \Delta w_{j,k}^l(i-1) \quad \text{and} \quad (3.12)$$

$$\Delta b_j^l(i) = -\eta \frac{\partial \mathcal{L}(i)}{\partial b_j^l(i)} + \alpha \Delta b_j^l(i-1), \quad (3.13)$$

where η is the learning rate to control the update rate of the weights and biases, and α is the momentum constant (Haykin, 1998, Chapter 4.3). High learning rates may cause oscillations around the minimum while low learning rates generally result in smoother weight and bias update paths towards the minimum. However,

learning rates that are too small are more likely to find a local instead of the global minimum. The terms associated with the momentum constant improve the learning process by respecting the weight or bias update of the previous learning step $i - 1$. For example, if the weight update step $\Delta w_{j,k}^l$ had the same sign over consecutive training samples (steps during learning), the weight update would increase, and a minimum would be approached faster. The learning process is therefore less likely to be stuck in a local minimum when including the momentum term (e.g., $\alpha > 0$). Conversely, if the weight update step changes sign over consecutive samples, the weight update steps would become smaller, dampening the oscillation around a minimum.

The iteration through the training data continues when the weights and biases are updated according to

$$w_{j,k}^l(i+1) = w_{j,k}^l(i) + \Delta w_{j,k}^l(i) \quad \text{and} \quad (3.14)$$

$$b_j^l(i+1) = b_j^l(i) + \Delta b_j^l(i). \quad (3.15)$$

Once the network has processed the last training data sample to adapt the weights and biases, one epoch of the training process is completed. Usually, the training data set is cycled through multiple times, given by the number of epochs, until some stopping criterion is fulfilled. Common stopping criteria include a maximum number of epochs or a sufficiently small change in the loss function, indicating that the network is converging to a minimum. Above, the weights and biases were updated after processing each training data sample i . However, it is more common to evaluate the loss and update the weights and biases after a certain number of training data samples (batch of samples) have been processed by the network rather than after each sample (Aggarwal, 2023, Chapter 2.6). This training approach is known as mini-batch stochastic gradient descent and improves the accuracy of the loss function gradients with respect to the weights and biases but uses more memory. The number of training data samples processed before the weights and biases are updated is known as the batch size. Similar to regression, the NN can be applied to new data once the training of the NN has been completed according to the stopping criteria.

In modern NN tools, such as Python's Keras library, all the above-mentioned parameters and many more can be easily adapted to optimize the learning process. The learning rate, batch size, epoch number, activation function, loss function, but also the architecture (number of neurons and hidden layers) are common tuning parameters. The retrieval setups for the applications in this thesis are explained in the appropriate sections of Chapters 4 and 5.

Chapter 4

Single microwave radiometer retrievals

In this study, the microwave radiometer measurements from the MOSAiC expedition are quality controlled and atmospheric parameters are derived using statistical retrievals. This study has been published in:

Walbröl, A. et al., 2022: Atmospheric temperature, water vapour and liquid water path from two microwave radiometers during MOSAiC, *Scientific Data*, 9, 534, [10.1038/s41597-022-01504-1](https://doi.org/10.1038/s41597-022-01504-1).

Detailed author contributions: Susanne Crewell, Kerstin Ebell and Andreas Walbröl conceptualized the study. Ronny Engelmann, Hannes Griesche, Martin Radenz, Julian Hofer and Dietrich Althausen collected the microwave radiometer data and maintained the instruments onboard RV *Polarstern* during the MOSAiC expedition. Marion Maturilli provided the radiosonde data. Kerstin Ebell and Andreas Walbröl quality controlled the HATPRO brightness temperature measurements. Kerstin Ebell derived atmospheric parameters (integrated water vapour, liquid water path, temperature profiles and absolute humidity profiles) using the regression coefficients trained with Ny-Ålesund radiosondes and synthetic brightness temperatures computed with radiative transfer simulations. Andreas Walbröl quality controlled the MiRAC-P brightness temperature measurements. Emiliano Orlandi provided the MiRAC-P training data. Andreas Walbröl developed the integrated water vapour retrieval from MiRAC-P observation, supported by discussions with Emiliano Orlandi. Furthermore, Andreas Walbröl evaluated all derived products (with scientific input from Susanne Crewell, Kerstin Ebell and Emiliano Orlandi) and wrote the majority of the manuscript. Parts of Sect. 4.3.1 and 4.3.2 were written by Kerstin Ebell and Andreas Walbröl together. All authors reviewed the manuscript.

Atmospheric temperature, water vapour and liquid water path from two microwave radiometers during MOSAiC

Andreas Walbröl¹, Susanne Crewell¹, Ronny Engelmann², Emiliano Orlandi³, Hannes Griesche², Martin Radenz², Julian Hofer², Dietrich Althausen², Marion Maturilli⁴, Kerstin Ebell¹

¹Institute for Geophysics and Meteorology, University of Cologne, Cologne, 50969, Germany

²Leibniz Institute of Tropospheric Research (TROPOS), Leipzig, 04318, Germany

³RPG Radiometer Physics GmbH, Meckenheim, 53340, Germany

⁴Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, 14473 Potsdam, Germany

Corresponding author(s): Andreas Walbröl (a.walbroel@uni-koeln.de)

4.1 Abstract

The microwave radiometers HATPRO (Humidity and Temperature Profiler) and MiRAC-P (Microwave Radiometer for Arctic Clouds - Passive) continuously measured radiation emitted from the atmosphere throughout the Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) expedition on board the research vessel Polarstern. From the measured brightness temperatures, we have retrieved atmospheric variables using statistical methods in a temporal resolution of 1 s covering October 2019 to October 2020. The integrated water vapour (IWV) is derived individually from both radiometers. In addition, we present the liquid water path (LWP), temperature and absolute humidity profiles from HATPRO. To prove the quality and to estimate uncertainty, the data sets are compared to radiosonde measurements from Polarstern. The comparison shows an extremely good agreement for IWV, with standard deviations of 0.08 – 0.19 kg m⁻² (0.39 – 1.47 kg m⁻²) in dry (moist) situations. The derived profiles of temperature and humidity denote uncertainties of 0.7 – 1.8 K and 0.6 – 0.45 g m⁻³ in 0 – 2 km altitude.

4.2 Background & Summary

Observations show that the Arctic is warming at a greater rate than the global average, a feature known as Arctic Amplification (Serreze et al., 2009; Screen and Simmonds, 2010; Serreze and Barry, 2011). Complex mechanisms that are not yet fully understood contribute to the enhanced warming. Water vapour is the strongest greenhouse gas and plays a major role in several processes related

to the amplification, but yet to an uncertain degree (i.e., downward longwave radiation flux, clouds, precipitation) (Graversen and Wang, 2009; Serreze and Barry, 2011; Ghatak and Miller, 2013). The enhanced sea ice loss in the Arctic in summer and autumn causes a greater evaporation, increasing the water vapour load of the warmed atmosphere, which can contain more moisture (Graversen and Wang, 2009; Screen and Simmonds, 2010). Moist air frequently intrudes into the Arctic, impeding sea ice formation and driving the retreat of the sea ice edge by increased net radiative warming and mechanical forcing (Woods and Caballero, 2016; Graham et al., 2019c; Komatsu et al., 2018). Strong moisture transports with a filamentary geometry are called Atmospheric Rivers (Newell et al., 1992), where information with a high temporal resolution is needed to capture the water vapour variability.

Within the past decades, a robust increase of moisture has been detected in the Arctic for certain regions and seasons (Maturilli and Kayser, 2017a; Parracho et al., 2018; Rinke et al., 2019). The increase of moisture content enhances the downward longwave radiation flux and therefore contributes to warming. Especially the autumn and winter months in the Barents Sea and Arctic Ocean are affected by positive moisture trends (Rinke et al., 2019). However, inconsistencies in the moistening trend among reanalyses call for reliable reference data to evaluate them in the data sparse region of the central Arctic. Radiosonde and satellite data are assimilated in reanalyses and therefore not suitable for independent evaluation. Additionally, water vapour estimations from different satellites disagree among each other, partly due to different measurement principles (Crewell et al., 2021). Despite the accuracy and high vertical resolution of water vapour and temperature profiles from radiosondes, low sampling rates (one to four sondes per day) and the poor spatial coverage of launch sites in the Arctic impede an adequate representation of the water vapour variability. Remote sensing in the microwave spectrum (satellite- or ground-based) is generally less accurate (lower vertical resolution) and faces several difficulties but has the potential to fill the gaps: Microwave radiometers (MWRs) on board polar orbiting satellites can sample the entire Arctic more than once per day even in cloudy conditions but suffer from uncertainties, for example, due to the lack of knowledge of the highly variably sea ice emissivity (Scarlat et al., 2017) and coarse vertical resolution.

Robust reference water vapour data sets are required for process studies and to evaluate reanalyses and satellite products in the Arctic. The Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAIC) expedition (Shupe et al., 2020, 2022) from September 2019 to October 2020 offers a unique set of detailed measurements in the central Arctic. During the expedition the research vessel (RV) *Polarstern* (Knust, 2017) from the Alfred Wegener Insti-

tute, Helmholtz Centre for Polar and Marine Research (AWI) drifted with the sea ice to investigate coupled atmosphere-ice-ocean processes in the central Arctic to ultimately improve climate models. In this data descriptor, we focus on measurements from the two MWRs MiRAC-P (Microwave Radiometer for Arctic Clouds - Passive), a high frequency MWR especially tailored for low water vapour conditions, and HATPRO (Humidity and Temperature Profiler), a standard MWR commonly used for monitoring of integrated water vapour (IWV). The multi-frequency HATPRO also allows for thermodynamic profiling. MWRs are the the only measurement systems to derive the total cloud liquid (liquid water path (LWP)) in all cloud conditions. From the MiRAC-P observations, we only present the IWV but humidity profiling and LWP derivation will be explored in the future.

The data introduced in this descriptor will be the base of upcoming studies within the Transregional Collaborative Research Centre TR 172 “Arctic Amplification: Climate Relevant Atmospheric and Surface Processes, and Feedback Mechanisms (AC)3” (Wendisch et al., 2017) to study the influence of water vapour and its variability on Arctic Amplification. They can support process studies with high quality IWV and LWP, as well as examinations of boundary layer developments with temperature and humidity profiles with a temporal resolution of one second. Furthermore, the data sets can be used as reference for the evaluation of satellite water vapour products and reanalyses.

4.3 Methods

In this section, we describe the two MWRs HATPRO and MiRAC-P and their measuring principles. Both radiometers were manufactured by RPG-Radiometer Physics GmbH (RPG). In the following, the regression for HATPRO and Neural Network for MiRAC-P to derive meteorological quantities from the raw sensor data are elaborated.

4.3.1 Microwave radiometers on board Polarstern

The RPG HATPRO G5 (Rose et al., 2005) from the Leibniz Institute of Tropospheric Research (TROPOS) was mounted on the OCEANET-Atmosphere container, which is routinely operated aboard RV Polarstern since 2009 (e.g., Kanitz et al., 2011, 2013; Engelmann et al., 2021). Its two receivers measure radiation emitted from atmospheric gases and liquid water in the microwave spectrum as brightness temperatures (TBs) in 14 channels with an absolute accuracy of 0.5 K. The half-power beam-widths of the receivers are in the range $2 - 4^\circ$. Seven of the

channels detect radiation at frequencies between 22.24 and 31.4 GHz (K-band, first receiver) and the remaining ones between 51.26 and 58.0 GHz (V-band, second receiver). The lower frequency band lies along the wing of a weak rotational water vapour absorption line at 22.24 GHz. Channels further away from the absorption line feature lower opacities and are therefore associated with the atmospheric window (e.g., at 31.4 GHz). Henceforth, these channels will be referred to as window channels. Despite the proximity of the 22.24 GHz channel to the water vapour absorption line, the opacity is still quite low so that radiance from all tropospheric layers contribute to the recorded signal (TBs are in the range of 10 – 40 K). Coarse water vapour profiles can be derived from the shape of the pressure broadened water vapour absorption line. Löhnert et al. (2009) found that 1 to 3 independent pieces of information (degrees of freedom) could be resolved in a central European and humid tropical climate. The emission of liquid water is more prominent in window channels and increases with frequency in the microwave spectrum (Janssen, 1993; Foken, 2021). Signals from ice clouds can be neglected because they are transparent in the range of HATPRO frequencies. Apart from humidity profiles, we use the K-band TBs to derive the integrated water vapour (IWV) and liquid water path (LWP). The higher frequency band covers a wing of the oxygen absorption complex at 60 GHz, allowing for temperature profile retrievals because the vertical distribution of the well-mixed oxygen is known (Foken, 2021). Channels close to the absorption line (58 GHz) feature a high opacity, sensing the radiation emitted from oxygen in the vicinity of the instrument. About 1 to 4 independent pieces of information can be resolved for temperature profiling (Löhnert et al., 2009), depending on the climate and scanning strategy. Most of the time the instrument operated in zenith mode with the elevation angle remaining at 90.0°. The zenith measurements were carried out with a temporal resolution and integration time of 1 s and were interrupted every 30 minutes for 110 s to perform so-called boundary layer (or elevation) scan, sensing the atmosphere at elevation angles of 5.4, 6.6, 8.4, 11.4, 14.4, 19.2, 30.0, and 90.0°. This elongates the instrument’s line of sight through the atmosphere and therefore increases the sensitivity in the atmospheric boundary layer. With this scanning method, we can derive temperature profiles with improved vertical resolution in the lower troposphere (adding about 2 independent pieces of information) resulting in more distinctly resolved height levels when combined with the zenith mode (Löhnert et al., 2009; Löhnert and Maier, 2012). HATPRO can operate in nearly all weather conditions, except during heavy precipitation. A dry blower keeps the radome dry even during slight precipitation, which is recorded by a simple yes/no sensor. Since the measurements are not reliable when the radome is wet, a rain flag has been applied to the data when necessary. Absolute

calibrations, where the receivers point at a built-in target at ambient temperature and a target cooled with liquid nitrogen, need to be performed about every 3 months to ensure the TB accuracy. Additionally, a gain calibration is performed automatically to avoid TB drifts (during MOSAiC, the interval of gain calibrations was 315 seconds).

The University of Cologne’s MiRAC-P (Mech et al., 2019a) (RPG-LHUMPRO-243-340 G5) is a passive MWR that measures atmospheric radiances as TBs at a temporal resolution of 1 s with six channels along the 183.31 GHz (G-band) water vapour absorption line and two window channels centered at 243 and 340 GHz. It was mounted next to HATPRO during the MOSAiC expedition. The six double-sided G-band channels are located at 183.31 ± 0.6 , ± 1.5 , ± 2.5 , ± 3.5 , ± 5.0 , and ± 7.5 GHz and, together with the window channels, can be used to derive IWV, LWP, and humidity profiles. The window channels at 243 and 340 GHz feature much higher opacities than the HATPRO window channels because the water vapour continuum absorption strength increases with frequency in the microwave spectrum (Janssen, 1993). At these frequencies, ice particles in clouds scatter atmospheric radiation causing uncertainties in radiative transfer modelling. The G-band water vapour absorption line is significantly stronger than the 22.24 GHz line and can get saturated if the water vapour load is sufficiently high, making the atmosphere opaque. Then, the TBs in the inner G-band (close to the 183.31 GHz line) are in the range of about 240 – 280 K, depending on the low-tropospheric temperature and moisture distribution. As it will be pointed out later, we can exploit the different absorption line strengths for a complementary usage of HATPRO and MiRAC-P. All MiRAC-P channels use a double side band heterodyne receiver design and have a half-power beam-width ranging from 1 to 1.3°. The off-axis parabolic mirror allows to point the radiometer to 0 – 180° elevation for sky view or to the internal ambient temperature calibration target (accuracy 0.2 K). During MOSAiC, MiRAC-P operated in zenith mode only. The measurement noise is below 0.5 K for all channels at one second integration time.

4.3.2 HATPRO: Retrieval via regression

In order to apply the regression with linear or quadratic terms, an example of the latter is given in equation (4.1), coefficients that map TBs to the desired meteorological quantities (IWV, LWP, absolute humidity and temperature profiles) need to be derived by training (c_0 , c_1 , and c_2). The IWV of the k -th sample in the training data set (N samples in total) is computed by

$$IWV_k = c_0 + \sum_{i=1}^m (c_{1,i} TB_{k,i} + c_{2,i} TB_{k,i}^2) \quad \text{with } k = 1, \dots, N, \quad (4.1)$$

where m is the number of MWR channels considered for this retrieval (7 K-band channels in this case). Here we use coefficients (Ebell, 2022) determined by Nomokonova et al. (2019) who applied them to HATPRO data at an Arctic site (Ny-Ålesund, Svalbard). The climatology behind the regression consists of $N = 2744$ radiosondes launched daily at 12 UTC in Ny-Ålesund covering the period 2006-05-21 to 2017-03-31 (Ebell and Walbröl, 2021). The radiosondes have been processed with the GRUAN version 2 algorithm (Sommer et al., 2012; Maturilli and Kayser, 2017a). For the regression, simulated TBs from the atmospheric state given by the radiosonde data were obtained with a one-dimensional radiative transfer model that only respects absorption and emission. Since radiosondes cannot measure the liquid water content of clouds, a simple cloud model was applied. Following Karstens et al. (1994), a liquid cloud is detected when the relative humidity in a height layer is greater than 95 % for temperatures above 253.15 K and the liquid water content is computed with a modified adiabatic approach. Ice clouds are transparent at HATPRO frequencies and therefore not taken into account. The radiative transfer model follows Rosenkranz (1998) for oxygen absorption, Ellison (2006) for liquid cloud absorption, Turner et al. (2009) for water vapour continuum absorption, Rüeiger (2002) for air mass corrections, and Liljegren et al. (2005) for the water vapour line width modelling. Random numbers with a normal distribution multiplied by 0.5 have been added to the TBs to imitate instrument noise with a strength of 0.5 K to match the instrument specifications given by RPG. For the retrievals of IWV, LWP, absolute humidity and temperature profiles from HATPRO’s zenith operation mode, the regression includes both linear and quadratic terms (Löhnert and Crewell, 2003) and only linear terms for the temperature profile based on the boundary layer mode. The evaluation with the test data (identical to training data as in Nomokonova et al., 2019) yields overall negligible biases, and standard deviations of 0.37 kg m^{-2} and 14.3 g m^{-2} for IWV and LWP, respectively. The humidity profile standard deviation over the entire test data set is 0.65 g m^{-3} at the surface and decreases to 0.17 g m^{-3} at 5 km altitude. Temperature profiles retrieved from zenith and elevation mode feature the lowest standard deviation at low altitudes, e.g. 250 m (1.5 K) and 150 m (1.0 K), respectively, increasing with altitude.

4.3.3 MiRAC-P: Retrieval via Neural Network

Given the saturated 183 GHz line, the retrieval problem is strongly non-linear for MiRAC-P. Therefore, we developed a Neural Network (NN), which is described following the published script (Walbröl, 2022), based on Python’s tensorflow and keras modules to retrieve IWV. The idea of a NN is to process a given input

(e.g., TBs) through one or more hidden layers, connected by so-called activation functions, to generate an output (e.g., IWV). We have refrained from using the Ny-Ålesund radiosonde data for the training of the MiRAC-P IWV retrieval because the dry conditions, where the sensitivity of this instrument is best, were not sufficiently represented. Instead, the training and test data consist of the ERA-Interim (ERA-I) reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011) and simulated TBs. The total number of samples is 24835, distributed over 8 virtual stations (certain grid points) north of 84.5° N and a period from 2001-01-01 to 2017-12-31 with data samples at 00, 06, 12, and 18 UTC (Orlandi and Walbröl, 2022). A subset of 12 years from the entire data set, which has been provided by the instrument manufacturer RPG, has been randomly selected as training and the remaining 5 years as test data. Each double side band frequency of the simulated G-band TBs has been averaged to be comparable to the measurements of MiRAC-P.

To obtain a more robust result from the training, we performed the training and evaluation (with the test data) 20 times with different random number seeds. The 20 random seeds were obtained by producing a set of 20 random numbers that lie between 0 and 1000 (boundaries have been chosen arbitrarily). At the beginning of the loop, the seeds of numpy's and tensorflow's random number generator were set to the random value. This random value affects the choice of training and test years because a permutation of an index ranging over all years (0 – 16) defines which ones are selected for training and testing. Of the 17 permuted indices, the first 12 (last 5) mark the training (test) years, respectively. For example, the test data can be 2002, 2006, 2007, 2010, 2016 with the remaining years being used for training. Furthermore, the initialization of the weights in the NN is affected by the seed of tensorflow's random number generator. As for HATPRO, we also added a random Gaussian noise to the synthetic TBs with a strength of 0.75 K for the G-band channels, 4.2 K for the 243 GHz, and 4.5 K for the 340 GHz channel, as recommended by the manufacturer. The higher noise for the two window channels reduced their weights in the retrieval and therefore diminishes the impact of signals from sources other than water vapour, such as cloud liquid emission or radiation scattered at ice particles. In correspondence with RPG, we chose the input vector of the NN to consist of all MiRAC-P TBs of a time step and the cosine and sine of the day of the year as additional information. The input was scaled to a feature range of -3 to 1 using the MinMaxScaler of the sklearn.preprocessing module. The input layer is connected to the only hidden layer, which has 32 nodes, with an exponential activation function. A linear activation function then links the hidden with the output layer, which only consists of the retrieved IWV. All layers are fully connected. The kernels of the

layers are initialized with the default Glorot uniform distribution. Similar to a regression approach, the goal of the training procedure is to adapt the weights of the NN to minimize a loss function that evaluates the predicted with the target IWV. In our case, the mean squared error is used as loss function, minimized with the Adam optimizer (Kingma and Ba, 2017). The maximum number of training epochs (number of times the entire training data is cycled through) is 100 with a batch size (number of samples to estimate error gradient before weights are updated) of 64.

During the optimization process, the mean squared error of the test data is monitored to avoid overfitting. Once the test loss did not improve for at least 20 epochs, the training was stopped and the weights that resulted in the lowest test loss were saved. As mentioned before, the training procedure was performed with 20 different randomly chosen seeds to assess the robustness of the NN. Hence, we get the mean and spread of the retrieval performance, quantified by the standard deviation (see equation (4.2)), from the test data ($0.55 \pm 0.03 \text{ kg m}^{-2}$). This value is also noted as a comment in each published retrieval file of the MiRAC-P (Walbröl et al., 2022b) and computed as the square root of the bias corrected test loss

$$\tilde{\sigma} = \sqrt{\frac{1}{N} \sum_{k=1}^N \left(IWV_{\text{pred},k} - Bias - \widehat{IWV}_k \right)^2}, \quad (4.2)$$

with \widehat{IWV}_k being the test data and $IWV_{\text{pred},k}$ the predicted IWV of the k -th sample. The bias is the mean difference between the target (in this case, ERA-I) and predicted IWV. After training, the model is applied to the observed TBs from MiRAC-P with the random seed that produced the lowest overall test loss (seed value: 558).

4.4 Data records

In this section, the data for the retrieval developments, the measured TBs, and retrieved products are presented for both HATPRO and MiRAC-P. The files for the retrieval development (Ebell and Walbröl, 2021; Ebell, 2022; Orlandi and Walbröl, 2022) have been uploaded to Zenodo, while the remaining files (Engelmann et al., 2022; Ebell et al., 2022; Walbröl et al., 2022a,b) have been published on PANGAEA. All data files are in netCDF format and summarized in Table 4.1.

The retrieval training data for HATPRO (Ebell and Walbröl, 2021) consists of one file that contains the entire training and test data for the retrieval of temperature (variable name in the file: ta) and humidity (hua) profiles, IWV (prw), and LWP (clwvi) from TBs (tb) measured by HATPRO. The data set is

Filename	Purpose	Published on	DOI
_nya_rt00.nc	Retrieval coefficients for HATPRO	Zenodo (Ebell, 2022)	10.5281/zenodo.6673886
MOSAIC_hatpro_retrieval_nya_v00.nc	Retrieval development HATPRO	Zenodo (Ebell and Walbröl, 2021)	10.5281/zenodo.5741350
MOSAIC_mirac-p_retrieval_pol_v00.nc	Retrieval development MIRAC-P	Zenodo (Orlandi and Walbröl, 2022)	10.5281/zenodo.5846394
Prefix	Content	Published on	DOI
ioppol_tro_mwr00_11_tb_v01_	HATPRO TBs (zenith mode)	PANGAEA (Engelmann et al., 2022)	10.1594/PANGAEA.941356
ioppol_tro_mwrBL00_11_tb_v01_	HATPRO TBs (boundary layer mode)	PANGAEA (Engelmann et al., 2022)	10.1594/PANGAEA.941356
MOSAIC_uoc_lhumpo-243-340_11_tb_v01_	MIRAC-P TBs	PANGAEA (Walbröl et al., 2022a)	10.1594/PANGAEA.941407
ioppol_tro_mwr00_12_clwvi_v01_	HATPRO IWP	PANGAEA (Ebell et al., 2022)	10.1594/PANGAEA.941389
ioppol_tro_mwr00_12_hua_v01_	HATPRO absolute humidity profile	PANGAEA (Ebell et al., 2022)	10.1594/PANGAEA.941389
ioppol_tro_mwr00_12_prw_v01_	HATPRO IWV	PANGAEA (Ebell et al., 2022)	10.1594/PANGAEA.941389
ioppol_tro_mwr00_12_ta_v01_	HATPRO temperature profiles (zenith mode)	PANGAEA (Ebell et al., 2022)	10.1594/PANGAEA.941389
ioppol_tro_mwrBL00_12_ta_v01_	HATPRO temperature profiles (boundary layer mode)	PANGAEA (Ebell et al., 2022)	10.1594/PANGAEA.941389
MOSAIC_uoc_lhumpo-243-340_12_prw_v01_	MIRAC-P IWV	PANGAEA (Walbröl et al., 2022b)	10.1594/PANGAEA.941470

Table 4.1: Filenames (prefixes of daily files), purpose (content), publication platform, and DOI of the published retrieval development data sets (measured TB and retrieved product data sets). Behind the prefixes of the daily files the year (yyyy), month (mm), day (dd), hour (HH), minute (MM), and second (SS) are noted as follows: yyyyymmddHHMMSS.nc

composed of meteorological observations from radiosondes, and simulated TBs. Elevation angles (*ele*) lower than 90° are only needed for the boundary layer temperature profile. Nomokonova et al. (2019) created the regression coefficients (Ebell, 2022) for zenith temperature (*tze*) and humidity (*hze*) profiles, boundary layer temperature (*tel*) profiles, and for IWV (*iwv*) and LWP (*lwp*) with this training data set.

The retrieval training data for MiRAC-P (Orlandi and Walbröl, 2022) has been provided by the manufacturer RPG and consists of one file that contains the entire training and test data for the retrieval of IWV (*prw*) from TBs (*tb*) measured by the MiRAC-P. The sine and cosine of the day of the year, computed from the time variable, are also included. The outline of the data set has been given in the previous section.

The HATPRO TB data set (Engelmann et al., 2022) contains daily files of atmospheric radiance measured as TBs (*tb*) during zenith (file name contains *mwr00*) and elevation (file name contains *mwrBL00*) mode. The retrieved products from HATPRO TBs include daily files of IWV (*prw*), LWP (*clwvi*), temperature (*ta*) and humidity (*hua*) profiles (Ebell et al., 2022). Temperature profiles have been retrieved from both zenith (filename contains *mwr00*) and elevation (filename contains *mwrBL00*) modes. The uncertainties of the variables are denoted by the expected standard error (*prw_err*, *clwvi_err*, *hua_err*, *ta_err*). The measured and retrieved data cover the period 2019-10-19 to 2020-10-02. Flag values indicate the quality of the data. The latitude and longitude coordinates of both instruments have been taken from RV Polarstern track data (Rex, 2020; Haas, 2020; Kanzow, 2020; Rex, 2021a,b).

The MiRAC-P TB data set (Walbröl et al., 2022a) is likewise structured as daily files of atmospheric radiation measured as TBs (*tb*). The TBs of the double side band frequencies (G-band) are averaged and labeled with the upper part of the band (e.g., 190.81 GHz instead of 183.31 ± 7.5 GHz). Similar to the training data (Orlandi and Walbröl, 2022), the sine and cosine of the day of the year are included for the NN retrieval. The retrieved IWV (*prw*) (Walbröl et al., 2022b) from MiRAC-P TBs is also compiled into daily files. The IWV uncertainty computed from the retrieval test data is noted as a comment to the retrieved variable and is also given in three categories (dry: $[0, 5)$, intermediate: $[5, 10)$, moist: $[10, 100)$ kg m^{-2}).

4.5 Technical Validation

In this section, we first discuss the accuracy of TBs and subsequently demonstrate the quality of the derived products — IWV, LWP, absolute humidity and

temperature profiles — by comparing them, where possible, to radiosonde observations. Since there is no direct measurement for LWP, we refer to past studies that show the quality of LWP derived from HATPRO (Rose et al., 2005; Toporov and Löhnert, 2020). Additionally, we compare our LWP with that from the Atmospheric Radiation Measurement (ARM) research facility MWR. The codes to analyze the derived products and generate Figures 4.1–4.5 are openly available (Walbröl, 2022) (see also Table 4.3).

The retrieved temperature and humidity profiles, as well as the IWV, will be compared with radiosonde measurements that have been gathered during the MOSAiC expedition (Maturilli et al., 2021). The radiosondes have been launched from RV Polarstern at least four times per day. We have converted the relative humidity to specific and absolute humidity by using the saturation water vapour pressure method suggested by Hyland and Wexler (1983). Then, we integrated the specific humidity over the pressure levels and divided by the standard gravitational acceleration to obtain IWV. For the comparison of the temperature and humidity profiles with HATPRO, we interpolated each radiosonde onto the height grid of HATPRO profiles. Radiosondes that did not reach at least 10 km altitude and that contained missing values have been rejected in the analysis (23 out of 1522). Drifts of the radiosondes with wind and uncertainties of the temperature, relative humidity, and pressure sensors, which are 0.2 – 0.4 K, 3 – 4 %, and 0.6 – 1.0 hPa, respectively (Maturilli et al., 2021), are error sources for the comparison with MWR data.

4.5.1 Brightness temperatures

Before the retrievals are applied, the quality of the measured TBs was checked following the procedure suggested by Löhnert et al. (2009). This involved the flagging of time steps when the rain flag is set, when the sun is within $\pm 7^\circ$ (elevation and azimuth) of the line of sight of the instrument, when TBs exceed the range 2.7 – 330.0 K, and when a receiver sanity check fails. The receiver sanity check is based on status flags of an internal procedure implemented by the manufacturer RPG in the housekeeping files of the MWRs, respecting also the receiver stability. Besides automated checks, a manual inspection of the TB data was performed to flag those time steps that show obvious artifacts not related to atmospheric signals (i.e., the crane at the bow of RV Polarstern causing sudden leaps in the TBs). In the following examinations, only time steps with good quality (flag = 0 or *nan*) have been used. The dates when the MWRs were calibrated with liquid nitrogen (Rose et al., 2005), to ensure the absolute accuracy of the TBs, are given in Figure 4.1. On 2019-10-19, 06:30 UTC, the first calibration of MiRAC-P was

carried out but yielded values that differed significantly from previous tests or expectations because the calibration integration time exceeded the maximum value supported by the software. Therefore, the calibration was repeated on 2019-10-22, 05:40 UTC. The MiRAC-P did not require as many calibrations as HATPRO because it showed a negligible drift of TBs over time whereas HATPRO is a standardized instrument recommended to be calibrated every 3 months (Rose et al., 2005). Slight jumps in the retrieved data can be found around calibration times. For example, the most noticeable and concurrently the highest absolute jump in IWV is 0.3 kg m^{-2} on 2020-03-01, 11:00 UTC in the HATPRO data.

4.5.2 Derived products

After applying the retrieval algorithms, the meteorological quantities were inspected whether or not they lie within a reasonable range. LWP must be within $[-200, 3000] \text{ g m}^{-2}$, IWV in $[0.0, 100.0] \text{ kg m}^{-2}$, temperature in $[180.0, 330.0] \text{ K}$, and absolute humidity in $[-0.5, 30] \text{ g m}^{-3}$, otherwise a flag value was set. The lower end of the thresholds for LWP and absolute humidity are chosen to respect slightly negative values that might result from the regression. For LWP, a further processing step is done. Potential offsets in LWP can be partly corrected using a clear-sky offset correction. Under clear-sky, i.e. here liquid-free, conditions, the LWP should be zero. To determine if a scene is liquid-free, the standard deviation of LWP within a 2-min time interval was analyzed. If this value is below a certain threshold, we assume that no liquid occurs. The threshold depends on the instrument and climate of the location. Based on visual inspection of the derived LWP and also cloud radar reflectivity, the best offset correction was achieved with a LWP standard deviation threshold of 1.5 g m^{-2} for almost the entire MOSAiC period. Only on 2020-07-10, -11, and -12 we used 0.9 g m^{-2} because the other value resulted in highly negative LWP. If all 2-min intervals within a 20-min time window indicate liquid-free conditions, the mean value of the retrieved LWP is calculated and subtracted from the original values. For cloudy periods, the estimated offset values during clear-sky periods are linearly interpolated and subtracted from the retrieved LWP.

Integrated water vapour

The MOSAiC expedition gave the opportunity for high quality water vapour measurements in the central Arctic for an entire year. This allows to capture the vast contrasts between winter (polar night) and summer (polar day). The contrast is nicely reflected in the IWV time series over all five MOSAiC expedition legs (measurement periods with a certain scientific crew) from both MWRs and the

radiosonde data (see Figure 4.1). In winter, the net outgoing longwave radiation and missing energy input from the sun can cause temperatures to drop to values below $-35\text{ }^{\circ}\text{C}$ (Rinke et al., 2021) making the air extremely dry due to the Clausius-Clapeyron relation. IWV is frequently below 4 kg m^{-2} from December 2019 to mid-April 2020 and can even be as low as 0.8 kg m^{-2} (February and March 2020). Only during occasional storms the IWV peaks above 5 kg m^{-2} (i.e., mid February 2020). As soon as the melt season commences in late spring (May 2020), the IWV shows much higher values (up to 30 kg m^{-2}) and a greater variability on synoptic scales (few days). In general, when merely considering the time series, all three data sets capture the extreme differences between winter and summer very well, proving the capability of the MWRs to capture the full range of IWV conditions. During synoptic events, such as cold air outbreaks or moist air intrusions, the benefit of the MWRs compared to the radiosondes is obvious. The MWRs capture the temporal evolution of IWV much better with their resolution of 1 s than the radiosondes, which were mainly launched four times a day during the expedition. IWV variabilities, gradients and extreme values, of which the latter might be missed by radiosondes, can be resolved at time scales of minutes or even seconds (Steinke et al., 2015). The extraordinarily strong moist air intrusion that occurred in mid-April 2020 is shown in greater detail in the Usage Notes as an example of the retrieved products.

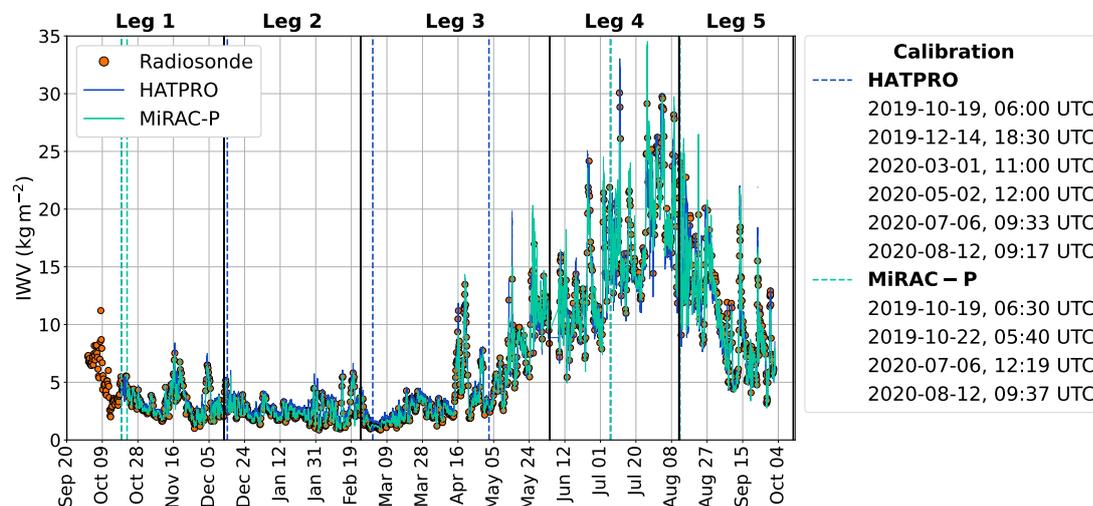


Figure 4.1: IWV time series from radiosondes (orange circles), HATPRO (blue), and MiRAC-P (cyan) covering the entire MOSAiC expedition (2019-09-20 – 2020-10-12). A 5-minute running mean has been applied to HATPRO and MiRAC-P data for smoothing. The calibration times of the MWRs are indicated as dashed vertical lines in their respective colours with the exact times noted in the legend. The MOSAiC legs are marked as black vertical lines.

To analyze the differences between radiosondes and the MWRs, the data sets

are displayed against each other in Figure 4.2. For the comparison, the MWR data has been averaged over 15 minutes starting from the radiosonde launch times. The standard deviation of these 15-minute periods are shown as error bars and indicate the noise but also the variability of the retrieved products. When we omit radiosondes that failed the quality check (as noted above) and MWR data where the flag value does not indicate good quality, a total of 1327 (1404) radiosonde launches are left to compare to HATPRO (MiRAC-P) data. From Figure 4.2, the complementary nature of HATPRO and MiRAC-P is visible. The MiRAC-P agrees better with radiosondes in dry conditions compared to HATPRO, which indicates the superior sensitivity of the strong G-band water vapour absorption line. To point out the complementary precision of MiRAC-P and HATPRO, Table 4.2 summarizes the standard deviations (computed as in equation (4.2), but with \widehat{IWV}_k representing the radiosonde and $IWV_{\text{pred},k}$ the MWR), biases, and root mean squared errors with respect to the radiosonde IWV for three IWV classes (dry: $[0, 5)$, intermediate: $[5, 10)$, moist: $[10, 100)$ kg m^{-2}). On average, HATPRO shows a bias of 0.35 kg m^{-2} for IWV smaller than 5 kg m^{-2} (see Table 4.2). Below 3.5 kg m^{-2} , the bias ranges from 0.25 to 0.75 kg m^{-2} . Here, higher biases occur in the drier conditions (lower IWV). Due to the superior sensitivity of MiRAC-P in dry conditions, a bias nearly three times lower (0.12 kg m^{-2} instead of 0.35 kg m^{-2}) can be seen for IWV smaller than 5 kg m^{-2} . In the dry regime, the MiRAC-P features a considerably lower standard deviation (0.08 kg m^{-2}) than HATPRO, which shows 0.19 kg m^{-2} . Even in the range $5 - 10 \text{ kg m}^{-2}$, the majority of the MiRAC-P data denotes differences to the radiosondes within $[-0.25, +0.25) \text{ kg m}^{-2}$ resulting in a bias of 0.0 kg m^{-2} , while the standard deviations of both MWR retrievals are similar ($\approx 0.3 \text{ kg m}^{-2}$) in that IWV range. When the IWV is greater than 10 kg m^{-2} , the retrieved IWV from MiRAC-P starts to scatter because the atmosphere becomes opaque to the G-band channels close to the absorption line. In other words, these channels become saturated and an increase in IWV does not change the TB any longer (e.g. Cadeddu et al., 2007, 2009). The higher the IWV, the more channels further away from the absorption line are affected by this saturation effect. The radiative transfer simulations of the training data have shown that the 183.31 ± 7.5 and 243 GHz channels are the only frequencies that can still detect IWV increases through TB changes for IWV above 15 kg m^{-2} . But in these frequencies and moist conditions, many TBs map to the same IWV so that no clear relation between the TBs and IWV can be inferred. This could explain the strong scattering of IWV from MiRAC-P when compared to the radiosonde measurements in moist conditions as seen in Figure 4.2, resulting in a standard deviation of 1.47 kg m^{-2} (see Table 4.2). HATPRO shows the opposite behaviour for high IWV, having an

uncertainty of 0.39 kg m^{-2} , which is almost a factor of 4 lower than the uncertainty of MiRAC-P.

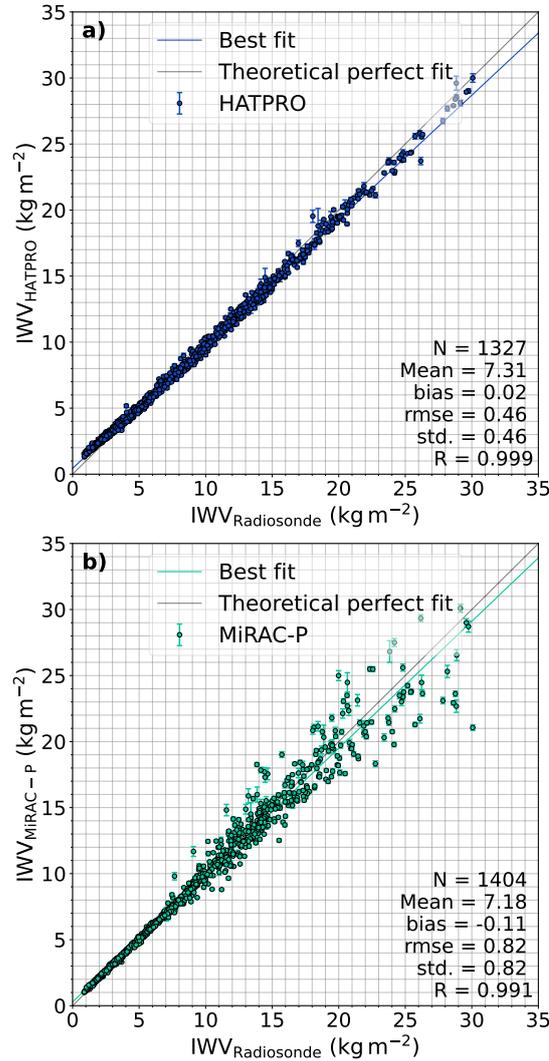


Figure 4.2: Comparison of the I WV from MOSAiC radiosondes (see text for details) with retrieved I WV from HATPRO (a) and MiRAC-P (b). The MWR data has been averaged over 15 minutes starting from each radiosonde launch time. The error bars denote the standard deviations of the 15-minute periods. A linear fit has been determined for both radiometers (coloured solid line) and a perfect fit is provided for orientation. Additionally, the number of samples (N), mean, bias, root mean squared error (RMSE), standard deviation (std.), and Pearson correlation coefficient (R) are given.

When considering the entire I WV range, the bias of the MiRAC-P (HATPRO) product is -0.11 kg m^{-2} (0.02 kg m^{-2}), with a standard deviation of 0.82 kg m^{-2} (0.46 kg m^{-2}). Compared to the Global Navigation Satellite System (GNSS) I WV retrieval performed by Männel et al. (2021), who found a bias of $0.08 \pm 0.04 \text{ kg m}^{-2}$ and a root mean squared error of 1.47 kg m^{-2} , the two MWRs yield more precise estimates of I WV, and HATPRO also a higher accuracy, when considering the

IWV range (kg m^{-2})	Instrument	N	RMSE (kg m^{-2})	bias (kg m^{-2})	σ (kg m^{-2})
[0, 5)	HATPRO	651	0.40	0.35	0.19
	MiRAC-P	730	0.15	0.12	0.08
[5, 10)	HATPRO	279	0.33	-0.14	0.29
	MiRAC-P	276	0.35	0.00	0.35
[10, 100)	HATPRO	397	0.61	-0.47	0.39
	MiRAC-P	398	1.49	-0.23	1.47

Table 4.2: Root mean squared error (RMSE), bias, standard deviation (σ) of IWV between HATPRO or MiRAC-P and radiosondes, divided into three categories of low, intermediate and high moisture load. Additionally, the number of samples (N) of the respective subclass and instrument is given.

entire range. The ARM research facility also derived IWV from their two-channel MWR, which was also located onboard RV Polarstern (Atmospheric Radiation Measurement (ARM) user facility, 2019). Their retrieval (MWRRET) combines a statistical and physical approach (Optimal Estimation), that also takes surface observations and radiosonde IWV into account, to generate a best estimate IWV data set (for a detailed description, please see Gaustad et al., 2011). ARM’s MWR provides a IWV record with a lower temporal resolution (26 seconds on average) and a roughly 20-day long data gap in August 2020. Both the lower resolution and the gap have to be taken into account when comparing their product with ours from HATPRO and MiRAC-P. Reducing the radiosonde and the three MWR data sets to a common time grid, where all quality flags indicate good quality, leaves us with 813 radiosondes to compare. The MWRRET best estimate yields a bias (standard deviation) of -0.21 kg m^{-2} (0.44 kg m^{-2}), while our products show -0.01 kg m^{-2} (0.44 kg m^{-2}) and -0.08 kg m^{-2} (0.75 kg m^{-2}) for HATPRO and MiRAC-P, respectively. Below 5 kg m^{-2} (10 kg m^{-2}), the performance of our products is especially good, having a standard deviation of 0.19 kg m^{-2} (0.29 kg m^{-2}) and 0.07 kg m^{-2} (0.28 kg m^{-2}) for HATPRO and MiRAC-P, respectively, while it is 0.40 kg m^{-2} (0.48 kg m^{-2}) for the MWRRET best estimate.

Liquid water path

The LWP is an important quantity for the evaluation of reanalyses and radiation balance. HATPRO, MiRAC-P and the two MWRs from the Atmospheric Radiation Measurement research facility (of which the three-channel MWR did not operate during most of the time Shupe et al., 2022) are the only instruments onboard RV Polarstern capable of retrieving LWP in all cloud conditions. Throughout the MOSAiC expedition, the LWP features a distinct seasonal variability (see Figure 4.3) with seasonally averaged daily mean LWP of 8, 25, 91,

Filename	Purpose
NN_retrieval_miracp.py	Neural Network retrieval training and application
data_tools.py	Module containing data analysis routines called by other scripts
import_data.py	Module containing various importer routines called by other scripts
met_tools.py	Module containing meteorological computations (humidity conversion, ...) called by other scripts
my_classes.py	Classes called by other scripts
case_study_overview_mwr_radiosonde.py	Script to generate Figure 4.6
mwr_pro_output_add_geoinfo.py	Adding Polarstern track data (Rex, 2020; Haas, 2020; Kanzow, 2020; Rex, 2021a,b) to HATPRO files
PANGAEA_tab_to_nc.py	Script to convert PANGAEA radiosonde (Maturilli et al., 2021) and Polarstern track data (Rex, 2020; Haas, 2020; Kanzow, 2020; Rex, 2021a,b) to netCDF format
plot_mwr_level_2a_radiosonde.py	Script to generate Figures 4.1, 4.2, and 4.3
plot_mwr_level_2bc_radiosonde.py	Script to generate Figures 4.4, and 4.5

Table 4.3: First block: Neural network retrieval development and application on MiRAC-P TB data (Walbröl et al., 2022a), generating the derived product (Walbröl et al., 2022b). Second block: Auxiliary modules called by other scripts. Third block: Visualization and data processing scripts. The codes are freely available (Walbröl, 2022).

and 40 g m^{-2} for winter (December – February), spring (March – May), summer (June – August), and autumn (September – November). Also the variability of the daily mean within a season, computed as seasonal standard deviations of the daily mean LWP, shows an annual cycle with 15, 38, 67, and 49 g m^{-2} for winter, spring, summer, and autumn, respectively. In summer, daily average LWP can exceed 250 g m^{-2} . This seasonality was also seen at Ny-Ålesund by Nomokonova et al. (2019). Higher values of LWP frequently occur in conjunction with high IWV because the moister air masses tend to generate more or deeper clouds. Former studies have proven the quality of the retrieved LWP, having an uncertainty of merely $14 - 23 \text{ g m}^{-2}$ (Rose et al., 2005; Toporov and Löhnert, 2020). In winter, when LWP is frequently within the uncertainty range (see Figure 4.3), the LWP estimates must be considered with care. Although retrieval noise might still result in slightly negative LWP, the clear-sky offset correction improved LWP biases. Comparing LWP derived from HATPRO with the best estimate from ARM’s two-channel MWR (MWRRET) (Atmospheric Radiation Measurement (ARM) user facility, 2019), we find that more than 81 % of the data values agree within $\pm 17.5 \text{ g m}^{-2}$ and 93 % within $\pm 27.5 \text{ g m}^{-2}$. For the comparison both data sets have been merged onto the same time grid due to differences in temporal resolution and data availability (as for IWV, see above).

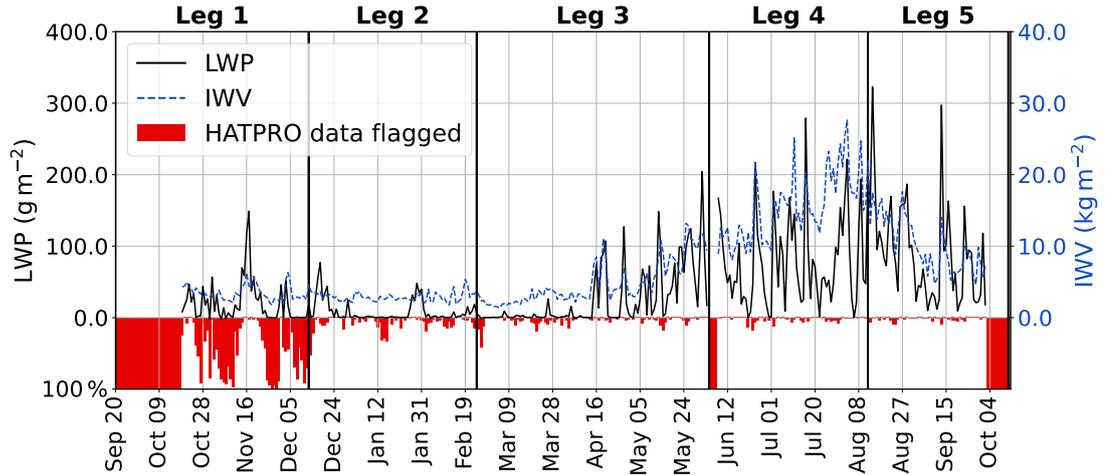


Figure 4.3: Time series of daily mean LWP (black) and IWV (blue dashed) values retrieved from HATPRO TBs for the entire MOSAiC period. The HATPRO data availability (red bars) shows the fraction of good quality (flag = 0) to the total number of data points on that day. A full red bar reaching from 0 to 100 % means that no data without a set flag is available on that day. The absence of red bars implies 100 % data availability with high quality.

The data availability in Figure 4.3 shows the fraction of non-flagged (flag = 0 or *nan*) values to the total number of data points of a day. During MOSAiC leg 1 (2019-09-20 – 2020-12-13), the internal sanity check of HATPRO frequently indicated a problem with the receiver of the V-band channels. The problem did not persist beyond the calibration on 2019-12-14, 18:30 UTC from where on the fraction of flagged values decreased significantly.

Humidity profiles

Humidity profiling from HATPRO data is more challenging than estimating the integrated amount because of the low information content (usually 1 to 3 independent pieces of information Löhnert et al., 2009). The dry conditions of the Arctic and the frequent occurrence of strong vertical gradients and moisture inversions (Devasthale et al., 2011; Nygård et al., 2014; Devasthale et al., 2016; Naakka et al., 2018) impede it further. The retrieved absolute humidity profile may still contain slightly negative values in high altitudes because of retrieval noise but flags are set for values below -0.5 g m^{-3} .

As for the comparison of IWV from HATPRO and radiosondes, the HATPRO data has been averaged over 15 minutes, starting from each radiosonde launch time, to evaluate the retrieved absolute humidity profiles. Systematic differences (bias) are expressed as the mean difference of absolute humidity over time on each height level ($\Delta\rho_v = \rho_{v,\text{HATPRO}} - \rho_{v,\text{radiosonde}}$) in absolute and relative terms (Figure 4.4a). The latter has been normalized by the mean absolute humidity

from radiosondes after averaging. The standard deviation of absolute humidity with radiosonde data as reference is also given in relative and absolute terms. As above, the relative term of absolute humidity standard deviation shown in Figure 4.4b has been computed by normalization with the mean radiosonde absolute humidity after determining the absolute term. While this procedure (normalizing after averaging) may not capture the individual relative differences for each radiosonde (normalizing before averaging), it is sufficient to give an idea of the relative uncertainty of the retrieved humidity profiles. We computed the bias and standard deviation for each MOSAiC leg so that Figure 4.4 displays the mean (standard deviation) of these quantities over the legs as black lines (shading).

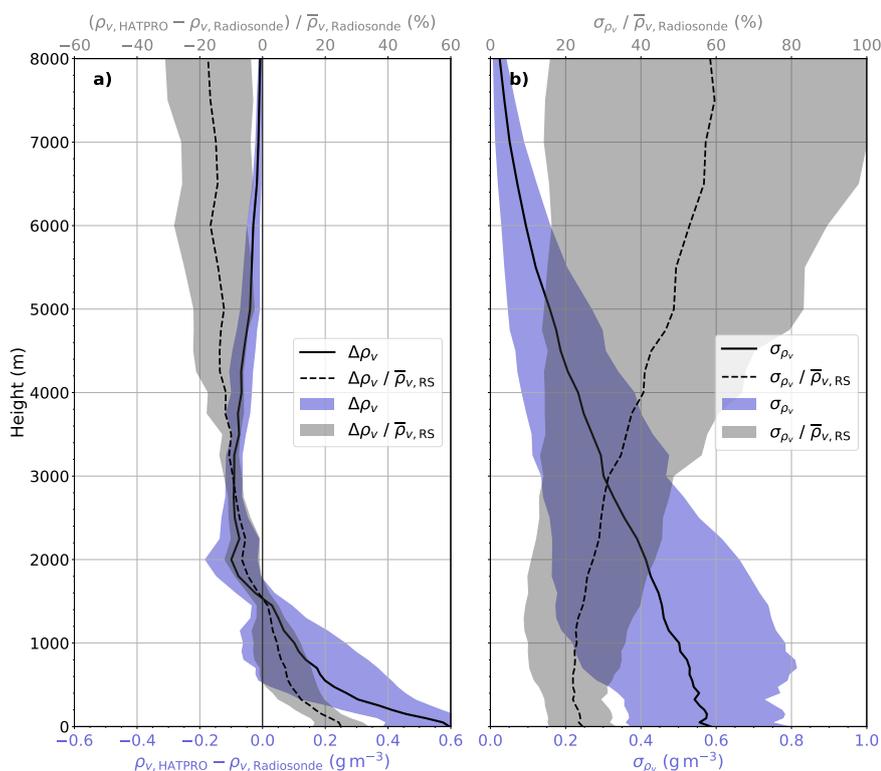


Figure 4.4: Bias (a) and standard deviation (σ) (b) of absolute humidity (ρ_v) profiles between radiosondes (RS) and HATPRO in absolute (solid) and relative (dashed) terms. Shading in blue (grey) indicates the variability over the MOSAiC legs as standard deviation in absolute (relative) terms (see text for details). The relative terms have been normed with the mean radiosonde absolute humidity. HATPRO data has been averaged over 15 minutes starting from radiosonde launch times.

In the lowest 1.5 km, HATPRO overestimates the absolute humidity with the highest bias (0.6 g m^{-3} or 25 %) at the surface. Further above, the bias becomes negative, up to about -0.1 g m^{-3} at 2 – 3 km height, and approaches zero in the remaining atmospheric column (up to 10 km). This is a typical behaviour when humidity inversions or strong moisture gradients are smoothed out in the

retrieved profile. The integrated humidity content (IWV) stays free of bias when a positive bias at the surface is balanced by a negative one in greater heights. In winter, when the humidity is low, the relative bias is usually higher than in summer.

At the surface, the standard deviation is 0.59 g m^{-3} or 25 % in absolute or relative terms, respectively (see Figure 4.4b). Because of the general decrease of absolute humidity with height, the standard deviation in absolute values also approaches zero with values of 0.41 g m^{-3} at 2 km height and 0.02 g m^{-3} at 8 km height. However, the relative standard deviation increases to 27 % and 58 % at those heights, respectively. Therefore, above 5 km altitude, when the standard deviation is near 50 %, the retrieved profile from HATPRO must be considered with care.

Past studies have found standard deviations of HATPRO-retrieved to radiosonde profiles of $0.9 - 0.6 \text{ g m}^{-3}$ in the lowest 2 km (Ebell et al., 2013, 2017), which are slightly higher than those found here ($0.59 - 0.41 \text{ g m}^{-3}$). However, their studies were carried out with data in the mid-latitudes, where the water vapour load is much higher. Ebell et al. (2013) found a relative uncertainty of 12 % in the lowest 2 km, while our analysis shows 22 – 27 %. To reduce humidity profile uncertainties and improve the information content, we are thus working on a synergetic retrieval of IWV and humidity profiles combining the measurements of both HATPRO and MiRAC-P.

Temperature profiles

HATPRO temperature profiles from the zenith mode have been averaged the same way as absolute humidity profiles for the comparison with radiosonde data. The measurements in boundary layer mode, performed only once every 30 minutes, were averaged over ± 30 minutes around radiosonde launch times due to the lower sampling rate. In Figure 4.5, the bias and standard deviation profiles can be seen for both measurement modes. Shading indicates the variability over the MOSAiC legs as described in the previous section and also shown in Figure 4.4. In the lowest 800 m, both modes show biases that quickly change with height (zenith: 1.2 to 2.0 to -2.8 K , elevation: -1.4 to 0.2 to -1.2 K). Radiative cooling over sea ice causes strong surface temperature inversions persisting almost throughout the entire winter (Tjernström and Graversen, 2009; Devasthale et al., 2010; Maturilli and Kayser, 2017a). The presence of low clouds may also generate inversions due to cloud top cooling (Sedlar et al., 2012; Devasthale et al., 2016). In summer, when the solar energy is used to melt the snow and sea ice (Graversen et al., 2008), temperatures remain close to the freezing point at the surface despite the possible presence of warmer air masses aloft. Therefore temperature inversions

are weaker and located at greater heights in summer (Tjernström and Graversen, 2009). Although the rapid changes of temperature over a few hundred metres below 1 km altitude cannot be resolved by HATPRO, the boundary layer mode denotes lower biases (standard deviations) than the zenith mode by up to 2 K (0.5 K). The standard deviation in 0 – 2 km height is 0.7 – 1.7 K (1.2 – 2.7 K) for the boundary layer (zenith) mode and therefore higher than the values found by Löhnert and Maier (2012) (0.5 – 1.4 K). This is likely due to the nearly permanent presence of inversions in the low Arctic troposphere. The switching signs of the biases up to 2 km height suggest that the inversions are smoothed out with a warm bias at their lower and a cold bias at their upper end for the zenith mode. Averaged over height and time, the boundary layer mode features a consistent cold bias.

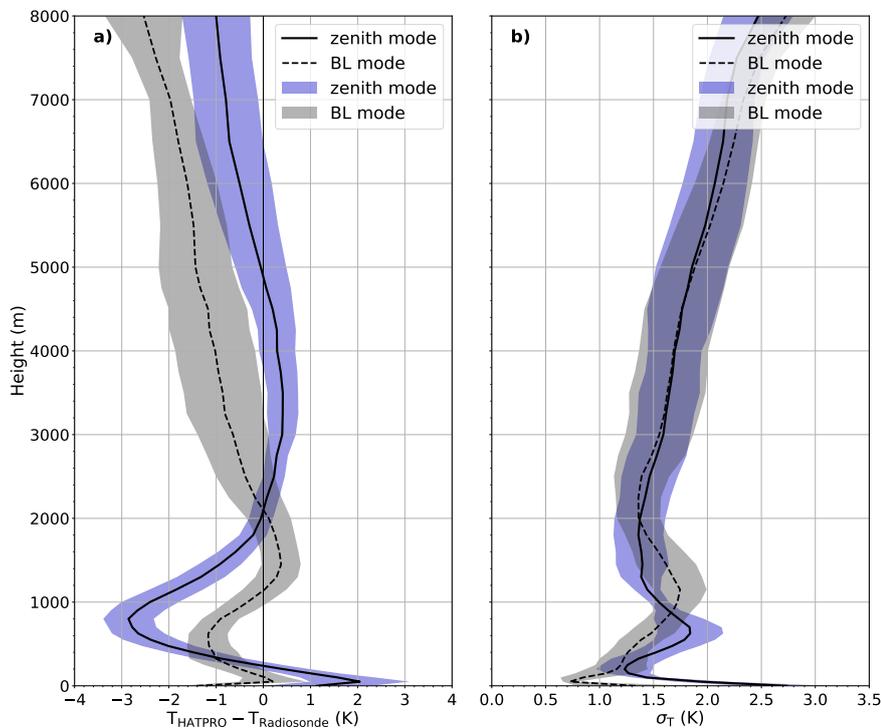


Figure 4.5: Bias (a) and standard deviation (σ) (b) of temperature (T) profiles between radiosondes (RS) and HATPRO zenith (solid) and boundary layer (BL, dashed) modes. Shading in blue (grey) indicates the variability over the MO-SAiC legs as standard deviation for the zenith (boundary layer) mode (as in Figure 4.4). HATPRO zenith measurements have been averaged over 15 minutes starting from radiosonde launch times while the boundary layer scans range ± 30 minutes around radiosonde launch times due to the lower sampling rate.

Above 2 km altitude, the bias of the zenith mode is smaller compared to the boundary layer scan (-1.0 to 0.4 K vs. -2.6 to 0.0 K). Up to about 4.5 km altitude, the standard deviations are similar, increasing from 1.4 to 1.8 K (see Figure 4.5). In greater heights, the standard deviation of the zenith tempera-

ture profile increases (up to 2.5 K in 8 km) because information comes from more height levels at once (broader weighting functions) providing less distinctly resolved levels in these altitudes. At heights between 2 and 4 km, we found similar uncertainties as in Löhnert and Maier (2012), who identified standard deviations of 1.4 – 1.7 K.

4.6 Usage Notes

In this section, we give some recommendations on handling the MWR data and provide an example of their capabilities for using them in a case study. When importing either the TB data or the retrieved products, data where the flag is not 0 or *nan* must be considered with care (remark: Python library *xarray* converts the good quality indicator to *nan* while *netCDF4* leaves it at the fill value 0). Importing the zenith temperature and humidity profiles from HATPRO over a long time period might result in large memory usage when using a library like *numpy* (one month of zenith temperature profiles results in roughly 100 million data points of type *float32*). Therefore, downsampling or the usage of a library like *xarray*, which compresses the data, is highly recommended. The comparisons of the HATPRO temperature and humidity profiles with radiosonde measurements is just one example of downsampling. Regarding the IWV, MiRAC-P should be used for values lower than 5 kg m^{-2} and HATPRO for values greater than 10 kg m^{-2} to optimally exploit the data sets. A transition zone from MiRAC-P to HATPRO IWV could be established in the range $5 - 10 \text{ kg m}^{-2}$ where both instruments work similarly well. For temperature profiles, a combination of zenith and boundary layer modes (0 – 2 km: boundary layer mode, 2 – 10 km: zenith mode) yields the best estimate.

The MOSAiC data policy requires a moratorium for the TB and retrieved data products until 2023-01-01. Only researchers that are a part of the MOSAiC community will have access before that date.

Moist air intrusion case

During the MOSAiC expedition a record breaking moist air intrusion was captured in April 2020 (Rinke et al., 2021). The codes used for importing, processing, and visualizing the data as seen in Figure 4.6 are published on Zenodo (Walbröl, 2022). Time stamps where the flag indicates bad quality have been filtered out and the radiosonde data was interpolated to a height grid with 5 m resolution, ranging from 0 to 15 km altitude. We have resampled HATPRO humidity and zenith temperature profiles to one-minute averages to reduce the number of data

points. In Figure 4.6, the time series of IWV, temperature and humidity profiles of the moist air intrusion case are shown. Since this was considered an intensive observation period due to the anomalous conditions, radiosondes were launched up to 7 times per day, which translates to a temporal resolution of 3 to 4 hours. However, despite the unusually high number of radiosonde launches, the full variability of IWV cannot be caught as well as with HATPRO and MiRAC-P. It is likely that the radiosonde data missed the maximum of 14.3 kg m^{-2} (HATPRO) IWV on 2020-04-19 by 0.8 kg m^{-2} . On that day, the radiosondes also experienced a strong horizontal drift due to high wind speeds, which could explain the discrepancy between HATPRO and radiosondes as well. Additionally, the MWRs detect the steep temporal gradient of the IWV much clearer than the radiosondes (i.e., between 2020-04-19 00 and 06 UTC). In the humidity profiles, the limited vertical resolution and biases at the surface of the HATPRO product are obvious. Even the strong humidity inversions that were detected by the radiosondes from 2020-04-19 00 UTC until 2020-04-21 12 UTC are not resolved. Also lifted layers of dry air with strong humidity inversions at their upper end (e.g., 2020-04-16 08 UTC) cannot be identified in the HATPRO humidity profiles. Below 2 km, small temperature inversions like the one seen in the radiosonde profile on 2020-04-16 are, at least, seen as isothermal layers in the HATPRO temperature profiles. The benefit of the boundary layer over the zenith mode is more distinct on 2020-04-18 00 UTC and 2020-04-19 00 – 12 UTC when the strong temperature inversion in 0 – 2 km height is clearly resolved.

Code availability

Almost all parts of this study have been coded with Python (version 3.8.10) using the following libraries: tensorflow (2.5.0), keras (2.5.0), numpy (1.17.4 and 1.19.5 (latter for NN retrieval)), sklearn (0.24.2), netCDF4 (1.5.3 and 1.5.7 (latter for NN retrieval)), matplotlib (3.4.3), and xarray (0.18.2). The codes of the NN retrieval, the visualization scripts for the Technical Validation and Usage Notes are openly accessible (Walbröl, 2022) and listed in Table 4.3. The scripts for HATPRO retrievals and processing of TB data of both instruments, written in the programming language IDL, are also available on Github and Zenodo (Walbröl, 2022).

Acknowledgements

We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Project 268020496 — TRR 172, within the Transregional Collaborative Research Center "Arctic Amplification: Climate

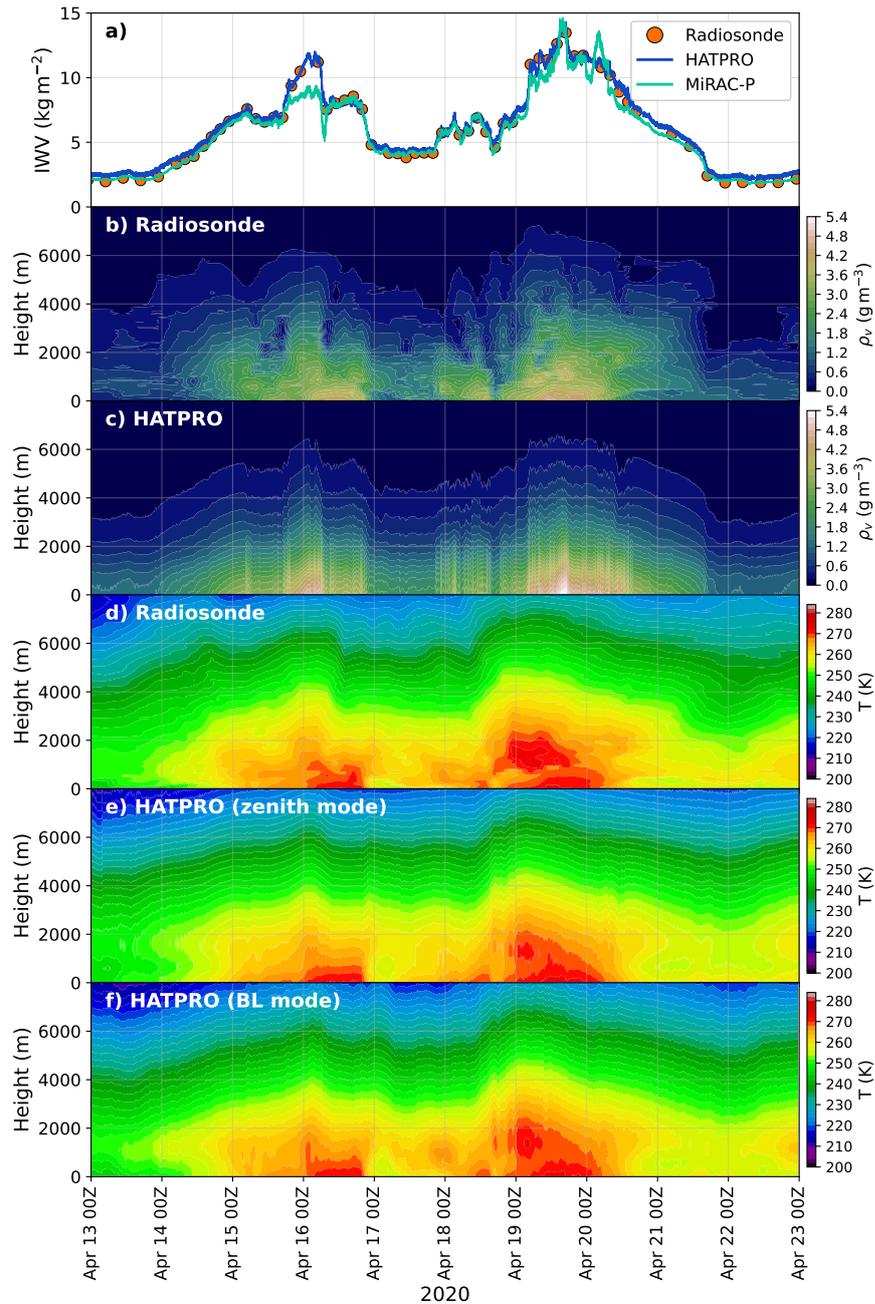


Figure 4.6: Overview of a moist air intrusion case during the MOSAiC expedition from 13th to 23rd April 2020 showing the IWV as in Figure 4.1 (a), absolute humidity (ρ_v) profiles from radiosondes (b) and HATPRO (c), as well as temperature (T) profiles from radiosondes (d), HATPRO zenith (e) and boundary layer (BL) (f) modes. HATPRO humidity and zenith temperature profiles have been resampled to one-minute averages.

Relevant Atmospheric and Surface Processes, and Feedback Mechanisms (AC)3". Data used in this manuscript was produced as part of the international Multi-disciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAiC20192020 and the Polarstern expedition AWI_PS122_00. We thank all those who contributed to MOSAiC and made this endeavour pos-

sible (Nixdorf et al., 2021). The microwave radiometer HATPRO was funded by Federal Ministry of Education and Research (BMBF) under FKZ: 01LKL1603A. Radiosonde data were obtained through a partnership between the leading Alfred Wegener Institute (AWI), the atmospheric radiation measurement (ARM) user facility, a US Department of Energy facility managed by the Biological and Environmental Research Program, and the German Weather Service (DWD).

Author contributions

R. E., H. G., M. R., J. H., and D. A. collected the data and maintained the instruments onboard RV Polarstern. K. E. applied the coefficients derived through training to HATPRO TB data. E. O. provided the MiRAC-P training data. M. M. provided the radiosonde data. A. W. wrote the majority of the manuscript, developed the NN retrieval (supported by E. O.), and analyzed the data (with scientific input from S. C., K. E., and E. O.). All authors reviewed the manuscript.

Chapter 5

Synergistic microwave radiometer retrievals

In this study, new retrievals of advanced water vapour products from the synergy of two microwave radiometers with complementary moisture sensitivity are developed, evaluated and compared to the single instrument retrievals. This study has been published in *Atmospheric Measurement Techniques* (see reference below). Here, the *accepted version* of the manuscript (03 September 2024) is presented.

Walbröl, A. et al. 2024: Combining low- and high-frequency microwave radiometer measurements from the MOSAiC expedition for enhanced water vapour products, *Atmospheric Measurement Techniques*, 17, 6223-6245, [10.5194/amt-17-6223-2024](https://doi.org/10.5194/amt-17-6223-2024).

Detailed author contributions: Susanne Crewell, Kerstin Ebell and Andreas Walbröl conceptualized this study. Mario Mech prepared the radiative transfer simulations on the Levante HPC System of the Deutsches Klimarechenzentrum (DKRZ). Andreas Walbröl modified the prepared scripts for his use to generate the training, test and evaluation data for this study. Andreas Walbröl also developed the retrievals, created all visualizations, analyzed the results and wrote all parts of the manuscript with scientific input from Susanne Crewell, Kerstin Ebell, Hannes Griesche and Mario Mech. Susanne Crewell, Kerstin Ebell, Hannes Griesche and Andreas Walbröl discussed the results. All authors reviewed the manuscript.

Combining low and high frequency microwave radiometer measurements from the MOSAiC expedition for enhanced water vapour products

Andreas Walbröl¹, Hannes J. Griesche², Mario Mech¹, Susanne Crewell¹, Kerstin Ebell¹

¹Institute for Geophysics and Meteorology, University of Cologne, 50969 Cologne, Germany

²Leibniz Institute for Tropospheric Research, Remote Sensing of Atmospheric Processes, 04318 Leipzig, Germany

Corresponding author(s): Andreas Walbröl (a.walbroel@uni-koeln.de)

5.1 Abstract

In the central Arctic, high quality water vapour observations are sparse due to the low density of meteorological stations and uncertainties in satellite remote sensing. Different reanalyses also disagree on the amount of water vapour in the central Arctic. The Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) expedition provides comprehensive observations that are suitable for evaluating satellite products and reanalyses. Radiosonde observations provide high quality water vapour estimates with a high vertical but a low temporal resolution. Observations from the microwave radiometers (MWRs) onboard the research vessel *Polarstern* complement these observations through high temporal resolution. In this study, we demonstrate the high accuracy of the combination of the two MWRs HATPRO (Humidity and Temperature Profiler) and MiRAC-P (Microwave Radiometer for Arctic Clouds - Passive). For this purpose, we developed new retrievals of integrated water vapour (IWV) and profiles of specific humidity and temperature using a Neural Network approach, including observations from both HATPRO and MiRAC-P to utilize their different water vapour sensitivity. The retrievals were trained with the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis version 5 (ERA5) and synthetic MWR observations simulated with the Passive and Active Microwave radiative TRANSfer tool (PAMTRA). We applied the retrievals on the synthetic and real observations and evaluated them with ERA5 and radiosondes launched during MOSAiC, respectively. To assess the benefit of the combination of HATPRO and MiRAC-P compared to single MWR retrievals, we compared the errors with respect to MOSAiC radiosondes and computed the vertical information content of the specific humidity profiles. The root mean squared error (RMSE) of IWV was reduced by up to 15 %. Specific humidity biases and RMSE were reduced by up to 75 and 50 %, respectively. The vertical information content of

specific humidity could be increased from 1.7 to 2.4 degrees of freedom. We also computed relative humidity from the retrieved temperature and specific humidity profiles and found that RMSE was reduced from 45 to 15 %. Finally, we show a case study demonstrating the enhanced humidity profiling capabilities compared to the standard HATPRO based retrievals. The vertical resolution of the retrieved specific humidity profiles is still low compared to radiosondes but the case study revealed the potential to resolve major humidity inversions. To which degree the MWR combination detects humidity inversions, also compared to satellites and reanalyses, will be part of future work.

5.2 Introduction

The amplified warming of the Arctic, known as Arctic amplification, is a well established phenomenon and has been discussed in several studies (e.g., Screen et al., 2012; Screen and Simmonds, 2010; Rantanen et al., 2022; Wendisch et al., 2023). Arctic amplification is caused by several positive climate feedback mechanisms, such as the ice albedo and the lapse rate feedback (Serreze and Barry, 2011; Wendisch et al., 2023). Following the Clausius Clapeyron relation, a warmer atmosphere can contain more water vapour before condensation occurs. Higher water vapour loads enhance the greenhouse effect (stronger emission in the thermal infrared) and thus increase temperatures at the surface (Held and Soden, 2000; Graverson and Wang, 2009; Ghatak and Miller, 2013). This positive feedback loop is known as the water vapour feedback and its role in Arctic amplification is still under investigation.

In the past decades, a moistening trend has been observed on a global scale (Chen and Liu, 2016; Allan et al., 2022) and also regionally in the Arctic (Ghatak and Miller, 2013; Maturilli and Kayser, 2017a; Parracho et al., 2018; Rinke et al., 2019; Serreze et al., 2012). The relative increase of the vertically integrated water vapour (IWV) is strongest in the Arctic (Chen and Liu, 2016). However, IWV trends have a high spatial heterogeneity and depend on the season (Parracho et al., 2018; Rinke et al., 2019). Many studies relied on atmospheric reanalyses, which assimilate measurements from synoptic stations, particularly radiosondes, satellites, etc. However, ground-based observations are sparse and satellite observations have different challenges in the Arctic (Crewell et al., 2021): The derivation of water vapour products from visible and infrared observations is hindered by darkness or clouds, and satellite products from microwave observations are uncertain due to the high and variable sea ice emissivity (Mathew et al., 2008; Wang et al., 2017; Scarlat et al., 2017). The lack of ground-based observations and difficulties in satellite remote sensing in the Arctic lead to high uncertainties

in water vapour products in reanalyses (Crewell et al., 2021; Parracho et al., 2018; Chen and Liu, 2016; Graham et al., 2019b). Therefore, it is not surprising to find a large spread of the IWV trend among reanalyses, often larger than the median trend itself for certain seasons and regions (Rinke et al., 2019).

A special feature of the Arctic is the high occurrence of humidity inversions, which are height layers where the water vapour concentration increases with height (Devasthale et al., 2011; Vihma et al., 2011; Nygård et al., 2014; Maturilli and Kayser, 2017a; Naakka et al., 2018). Humidity inversions are strongly coupled with temperature inversions (Tjernström et al., 2004), which form due to radiative cooling in clear sky conditions in winter, or due to sea ice melt or advection of warm and moist air above the boundary layer in summer (Graversen et al., 2008; Devasthale et al., 2010; Tjernström et al., 2019). Humidity inversions are a moisture source for the formation and maintenance of clouds through entrainment at the cloud top (Nygård et al., 2014). It is therefore important to have humidity observations with a sufficiently high vertical resolution that allows to capture this characteristic feature of the Arctic humidity profile. Additionally, the vertical water vapour distribution affects the downward thermal infrared radiation. Tjernström et al. (2019) showed that in cases when humidity inversions were present, the downward thermal infrared radiation was higher fostered by fog or low cloud formation.

Current reanalyses have difficulties in correctly representing the stable stratification of Arctic winter conditions (Wang et al., 2019; Yu et al., 2021; Graham et al., 2019a). For example, the widely used European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis version 5 (ERA5) (Hersbach et al., 2020), which is among the best performing global reanalyses in the Arctic, still shows positive near-surface air temperature and humidity biases (Graham et al., 2019a; Avila-Diaz et al., 2021; Loeb et al., 2022; Yu et al., 2021). The biases are highest in cold stable conditions found over sea ice in winter and smaller in summer or over the open Arctic Ocean (e.g., Fram Strait, Wang et al., 2019; Graham et al., 2019b). Herrmannsdörfer et al. (2023) suggested that ERA5 does not sufficiently represent sea ice thickness and snow depth. Difficulties in the representation of the stable conditions and positive biases of temperature and humidity at the surface result in errors in the temperature and humidity profiles of ERA5 (and other reanalyses).

It follows that reanalyses and satellite products struggle with the representation of water vapour in the Arctic. To evaluate the accuracy of water vapour in current reanalyses and satellite products, we need reference measurements. However, reliable and high quality water vapour measurements in the central Arctic are currently only available through field campaigns. The Multidisciplinary drift-

ing Observatory for the Study of the Arctic Climate (MOSAIC, Shupe et al., 2022) expedition, where the research vessel (RV) *Polarstern* (Knust, 2017) was frozen into the ice to observe the Arctic climate for a full annual cycle, provides unique observations for this purpose. Radiosonde measurements (Maturilli et al., 2021) yield IWV and humidity profiles with a high vertical but low temporal resolution (3–6-hourly). Additionally, water vapour products have been derived from upward looking microwave radiometers (MWRs) that were mounted on the OCEANET container (Macke et al., 2010; Engelmann et al., 2021) at the bow of RV *Polarstern*: Walbröl et al. (2022) created retrievals of IWV and profiles of absolute humidity and temperature from the low frequency Humidity and Temperature Profiler (HATPRO, Rose et al., 2005) and an IWV product specifically designed for dry conditions from the high frequency Microwave Radiometer for Arctic Clouds - Passive (MiRAC-P, Mech et al., 2019a). The MWR products have a high temporal resolution (almost every second) but the humidity profile from HATPRO is coarse with less than 2 degrees of freedom (Löhnert et al., 2009).

The high frequency observations from MiRAC-P have a high sensitivity to atmospheric water vapour in dry conditions ($\text{IWV} < 10 \text{ kg m}^{-2}$) but get saturated in humid conditions ($\text{IWV} \geq 10 \text{ kg m}^{-2}$, Cadeddu et al., 2007, 2022; Fionda et al., 2019). In contrast, the low frequency observations from HATPRO have a high sensitivity in humid conditions but a weak signal in the dry conditions of the Arctic in winter. The complementary moisture sensitivity of HATPRO and MiRAC-P motivates the synergy of both instruments, as it has been done for IWV in e.g., Cadeddu et al. (2009).

In this study, we develop retrievals of water vapour products combining observations from HATPRO and MiRAC-P to improve the vertical resolution of specific humidity profiles and reduce errors compared to single MWR retrievals. We retrieved specific humidity instead of absolute humidity because it is a more commonly used humidity measure in atmospheric reanalyses and satellite products. Specifically, we answer the following questions:

1. How much are IWV and humidity profile errors reduced compared to single instrument retrievals and what is the influence of using different retrieval setups?
2. What is the vertical information content benefit for humidity retrievals when combining two MWRs with different moisture sensitivity?
3. Is the vertical information content sensitive to cloud presence, temperature or water vapour amount?

The manuscript is structured as follows: In Sect. 5.3, we start with a descrip-

tion of the data used for the retrieval development and the measurements from the MOSAiC expedition, which will be used for the application and evaluation of the retrieval. In Sect. 5.4, we elaborate on the preparation of the retrieval development data before giving details on the retrieval setup and vertical information content estimation. Afterwards, we evaluate the retrieval in Sect. 5.5 and estimate the information benefit in Sect. 5.6 before concluding the manuscript in Sect. 5.7 by answering the questions raised above.

5.3 Data sets

5.3.1 Retrieval development data

Radiosondes are commonly used for the evaluation of temperature and humidity profile retrievals because of the high vertical resolution and accuracy (e.g., Cimini et al., 2010; Löhnert and Maier, 2012). Due to the lack of radiosonde stations and uncertain water vapour observations from satellites, we selected the ERA5 reanalysis (Hersbach et al., 2020) as a data source for the retrieval development. With a horizontal resolution of 31 km and 137 vertical levels, it has the highest horizontal and vertical resolution of all current global reanalyses. The high vertical resolution might be beneficial for developing humidity profile retrievals because a low vertical resolution could constrain the retrieval from reaching its true potential. ERA5 data is available for 1940–present with an hourly resolution. Despite having slightly higher biases in near-surface air temperatures and humidity in cold stable conditions over sea ice than other reanalyses, ERA5 overall performs best in the Arctic, especially concerning the representation of clouds and precipitation (Graham et al., 2019a). The better representation of clouds and precipitation is beneficial for the simulation of microwave radiances for the retrieval development (described in Sect. 5.4.1). Also, extreme precipitation and temperature events are better captured by ERA5 than other reanalyses (Avila-Diaz et al., 2021; Wang et al., 2019; Loeb et al., 2022).

5.3.2 MOSAiC observations for retrieval application and evaluation

RV *Polarstern* drifted with an ice floe from 04 October 2019 in the Laptev Sea across the central Arctic Ocean until it approached the marginal ice zone in the Fram Strait on 31 July 2020. Between mid-May and mid-June 2020, RV *Polarstern* had to leave the floe for logistical reasons. To capture the refreezing period of the ice, RV *Polarstern* drifted with a second ice floe close to the North

Pole from 21 August to 20 September 2020. In early October 2020, RV *Polarstern* left the sea ice.

Radiosondes

Throughout MOSAiC, Vaisala RS41 radiosondes have been launched from RV *Polarstern* at the standard synoptic times (00, 06, 12, and 18 UTC). The actual launch time is usually around 1 hour before the respective synoptic time due to the relatively slow ascent rates of about 5 m s^{-1} . During intense observation periods, additional radiosondes were launched at 03, 09, 15 and 21 UTC. Here, we use all radiosonde level 2 data from 01 October 2019 to 01 October 2020 (Maturilli et al., 2021). Radiosondes provide temperature, pressure, and relative humidity with accuracies of 0.2–0.4 K, 0.04–1.0 hPa, and 3–4 %, respectively. With a measurement frequency of 1 Hz, the vertical resolution is about 5 m. For the comparison with the retrievals, the radiosonde data has been interpolated onto the retrieval height grid (see Sect. 5.4.1).

Cloudnet and surface meteorology measurements

To evaluate the presented retrievals in different atmospheric conditions, we included additional data sets from the MOSAiC expedition: To distinguish between freezing and non-freezing conditions at the surface (temperatures below and above 273.15 K), the 2 m temperature measurements from the tower at the Met City site (Cox et al., 2023a) were used. The Met City site was located within the central observatory, only a few hundred metres away from RV *Polarstern*. Additionally, we identified cloudy scenes using the Cloudnet retrieval products (Griesche et al., 2024b). Cloudnet uses a synergy of passive and active atmospheric remote sensing to provide profiles of cloud macro- and microphysical properties (liquid and ice water content, effective radii of liquid droplets and ice crystals) with a time and height resolution of 30 s and 30 m, respectively (Illingworth et al., 2007; Tukiainen et al., 2020).

Cloudnet delivers, e.g., a classification of the atmospheric conditions, distinguishing between clear sky, different cloud types (ice, liquid, mixed-phase), and the presence of aerosols and insects, for each time-height pixel. Because of technical limitations, the Cloudnet product starts at a height of 182 m and can therefore miss the presence of low-level stratus clouds, which are common in the Arctic (Gierens et al., 2020; Griesche et al., 2020). The additional low-level stratus detection developed by Griesche et al. (2020) was used to mask these cases.

In this study, clear sky conditions were identified using Cloudnet target clas-

sification data (Engelmann et al., 2023) and the low-level stratus mask (Griesche et al., 2023) where quality flags indicated good quality (including also the Cloudnet issue dataset, Griesche and Seifert, 2023). As we compare our retrievals with radiosonde measurements, we selected Cloudnet data at times from the radiosonde launch to 15 minutes after the launch. A radiosonde launch is considered clear sky when no low-level stratus were present and the Cloudnet target classification indicated either clear sky, aerosols, or insects.

Microwave radiometers

The two upward-looking microwave radiometers HATPRO and MiRAC-P measure radiation emitted from water vapour, oxygen and hydrometeors. Measured radiances are typically expressed as brightness temperatures (TB). HATPRO detects radiances in seven channels between 22.24 and 31.4 GHz (K-band) and in seven channels between 51.26 and 58 GHz (V-band). MiRAC-P has a double-sideband receiver that measures radiances at six frequencies from 183.31 ± 0.6 to 183.31 ± 7.5 GHz (G-band) and a two-channel receiver for 243 and 340 GHz. At MiRAC-P frequencies, the scattering of radiation by hydrometeors is relevant, and the contribution of the continuum water vapour absorption is stronger (Rosenkranz, 1998).

Figure 5.1 shows TBs simulated with the Passive and Active Microwave radiative TRANSfer tool (PAMTRA, Mech et al., 2020), using two clear sky radiosondes from MOSAiC (winter: 05 March 2020, 06 UTC, summer: 06 August 2020, 00 UTC). A higher atmospheric opacity generally results in higher TBs in the zenith. In the K-band channels of HATPRO and the G-band channels of MiRAC-P, which are located around resonant water vapour absorption lines, the different water vapour loads of winter and summer can be well distinguished by their large TB differences of up to 40 K in the K-band and more than 100 K in the G-band. Also in MiRAC-P's high frequency channels at 243 and 340 GHz, TB differences are larger than at K-band frequencies (up to 200 K) due to continuum water vapour absorption. At the K-band frequencies, the relation between TBs and IWV is rather linear and becomes more nonlinear for the higher frequencies (G-band and above).

Observations along resonant water vapour absorption lines are well suited to derive IWV and humidity profiles (Crewell et al., 2001; Cadeddu et al., 2007; Cimini et al., 2010; Perro et al., 2016, e.g.). Because of the high water vapour sensitivity, most of the G-band channels are saturated in the summer case, meaning they do not observe radiances from the entire atmospheric column. In contrast, the K-band channels show almost no water vapour signal in the extremely dry winter case (IWV of 0.9 kg m^{-2}) while there is still a strong signal in the G-band.

Furthermore, higher TBs in summer compared to winter are caused by higher temperatures of the emitting gases and hydrometeors (Fig. 5.1). The V-band channels of HATPRO lie around the oxygen absorption complex and can be used for temperature profiling (Rose et al., 2005; Löhnert and Maier, 2012). As explained in Walbröl et al. (2022), HATPRO also measured atmospheric radiances at different elevation angles every 30 minutes during MOSAiC, allowing for more detailed temperature profile retrievals in the lower troposphere (boundary layer temperature profiles).

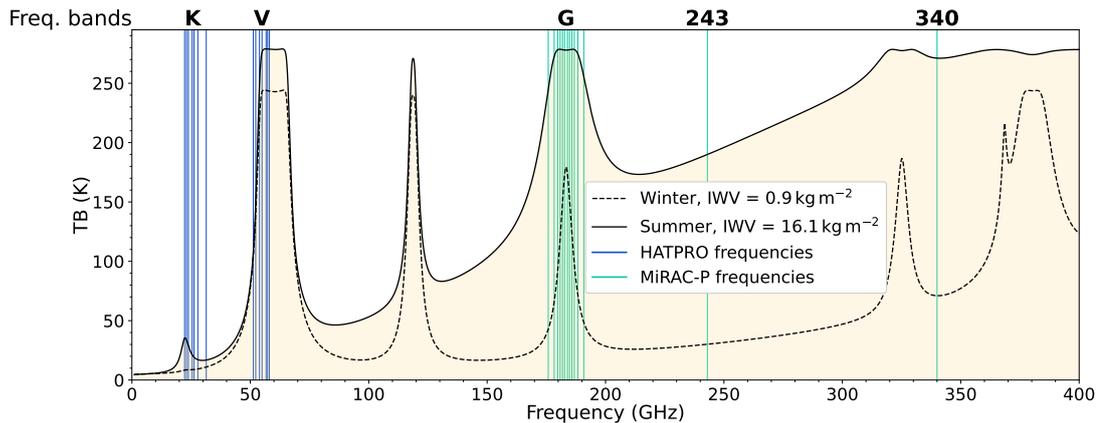


Figure 5.1: Brightness temperatures (TBs) from 1 to 400 GHz simulated with PAMTRA for two radiosondes launched from RV *Polarstern* during MOSAiC (winter: 05 March 2020, 06 UTC, summer: 06 August 2020, 00 UTC). The dashed (solid) black line shows the TBs simulated with meteorological data from the winter (summer) radiosonde. The blue (cyan) lines indicate the frequencies at which HATPRO (MiRAC-P) measures. The labels K, V, G, 243, and 340 represent abbreviations for sets of frequency channels (bands) of HATPRO and MiRAC-P.

In this study, we generally used TB measurements where flags indicate good quality (Walbröl et al., 2022). We identified a few rain events between late-May and late-June 2020 that were not flagged by visual inspection. The quality flags have been updated. Additionally, we checked whether other flag values could be accepted and found that a receiver sanity flag was often set although the data looked reasonable. Therefore, we also included that data in our analysis. Times before the first successful calibration of both MWRs (22 October 2019, 05:40 UTC) have been excluded.

For the information benefit analysis, we compared the new synergistic retrievals to the single instrument retrievals developed by Walbröl et al. (2022), i.e., the two IWV products (HATPRO and MiRAC-P), and profiles of temperature and absolute humidity from HATPRO (Ebell et al., 2022; Walbröl et al., 2022b). We converted the retrieved absolute to specific humidity using the retrieved tem-

perature profiles and air pressure from radiosondes. All retrieved quantities were averaged over 15 minutes, starting at the launch time of each radiosonde, for the comparison with MOSAiC radiosondes. For boundary layer temperature profiles, we extended the averaging window to 30 minutes before to 30 minutes after each radiosonde launch due to the lower sampling rate.

5.4 Methods

The retrieval of an atmospheric state vector \mathbf{x} (e.g., specific humidity profile) from an observation vector \mathbf{y} (e.g., TBs at different frequencies) is an inverse problem. In its simplest form, the inverse problem can be formulated as $\mathbf{x} = F^{-1}(\mathbf{y})$ where F is the forward operator (e.g., radiative transfer model, here, PAMTRA). In atmospheric remote sensing, inverse problems are often ill-conditioned because small changes in observations can lead to large changes in the retrieved state vector and many different atmospheric states can lead to the same observations. Furthermore, the inverse problem is ill-posed because the radiative transfer equation cannot be inverted in a direct way.

The challenge is to find the most probable and realistic state of the atmosphere that fits the observations. In physical retrievals (e.g., Optimal Estimation, Rodgers, 2000; Ebell et al., 2017), the state vector \mathbf{x} is adapted as long as the forward simulated observations $F(\mathbf{x})$ do not agree with the actual observations \mathbf{y} within a given uncertainty range. Physical retrievals are computationally expensive but provide physically consistent state vectors and uncertainty estimation. However, at the high frequencies of MiRAC-P, the scattering of radiation by frozen hydrometeors cannot be neglected and may therefore introduce uncertainties in the radiative transfer calculations needed for the forward simulation $F(\mathbf{x})$. The retrieval would require assumptions on hydrometeor properties (concentration, size, shape, orientation) or further hydrometeor observations, making it dependent on the availability of such observations.

Statistical retrievals are computationally cheap approaches that are also well established and provide similarly good results as physical retrievals (Solheim et al., 1998). In statistical retrievals, empirical relations are used to map observations to the state vector. The statistical relationship between observations and state vector must be trained with large data sets covering the conditions of the area of interest. Regression or deep learning algorithms are examples of statistical retrievals. In this study, we use Neural Networks (NNs) because they can deal better with the nonlinear relationship between IWV and TB measurements in the G-band compared to regression. During the development of the MiRAC-P-only retrieval (Walbröl et al., 2022), tests showed that the IWV retrieved with

a multiple nonlinear regression had a significantly higher spread than when retrieved with NNs.

5.4.1 Retrieval preparation

For the NN retrievals of IWV, specific humidity and temperature profiles during MOSAiC, a training data set is needed that covers the variability of the environmental conditions in the central Arctic over an annual cycle. We selected ERA5 data for 2001–2018 with 6-hourly temporal resolution at 12 grid points, of which 9 are located in the central Arctic and 3 in the Fram Strait (see Fig. 5.2). The grid points in the Fram Strait cover more humid conditions as this is a typical pathway for warm and moist air intrusions (Mewes and Jacobi, 2019).

Simulated HATPRO and MiRAC-P observations are needed in conjunction with the ERA5 data to train the NN. Meteorological data (temperature, relative humidity, geopotential height, pressure, 10 m wind) and vertical hydrometeor distributions from ERA5 (specific cloud liquid, ice, rain and snow content) have been used as input to simulate TBs with PAMTRA. The ERA5 skin temperature was used for the sea ice and sea surface temperatures. The TBs were simulated with PAMTRA’s default gaseous absorption, hydrometeor absorption and scattering models as described in Mech et al. (2020).

Four years of simulated TBs and ERA5 data (2001, 2006, 2011, and 2015) were held back from the retrieval development for the final evaluation (ERA5 evaluation data set). With the remaining 14 years of data, we trained the NN and validated its performance (11 and 3 years for the training and validation data sets, respectively). The number of training (validation) samples is roughly 192000 (52000). To avoid training near-surface temperature and humidity biases from ERA5 into the retrieval, a small subset of about 5 % of level 2 MOSAiC radiosondes (Maturilli et al., 2021) was also included in the validation process. For the retrieval development and evaluation, atmospheric profiles have been interpolated onto the same height grid used in the standard HATPRO retrieval (Löhnert, 2023; Marke et al., 2024) and in Walbröl et al. (2022), ranging from 0 to 10000 m with the vertical spacing increasing from 50 m at the surface to 500 m at the top. The height grid was limited to 8000 m for temperature profiles to avoid the tropopause. Additionally, to imitate measurement uncertainties, random Gaussian noise with a mean of 0 and standard deviations of 0.5, 0.75, and 2.5 K has been added to the simulated TBs at K–V, G, and 243–340 GHz, respectively. We intentionally used a higher noise level for the higher frequencies to account for the higher PAMTRA simulation uncertainties due to scattering from hydrometeors and water vapour continuum absorption.

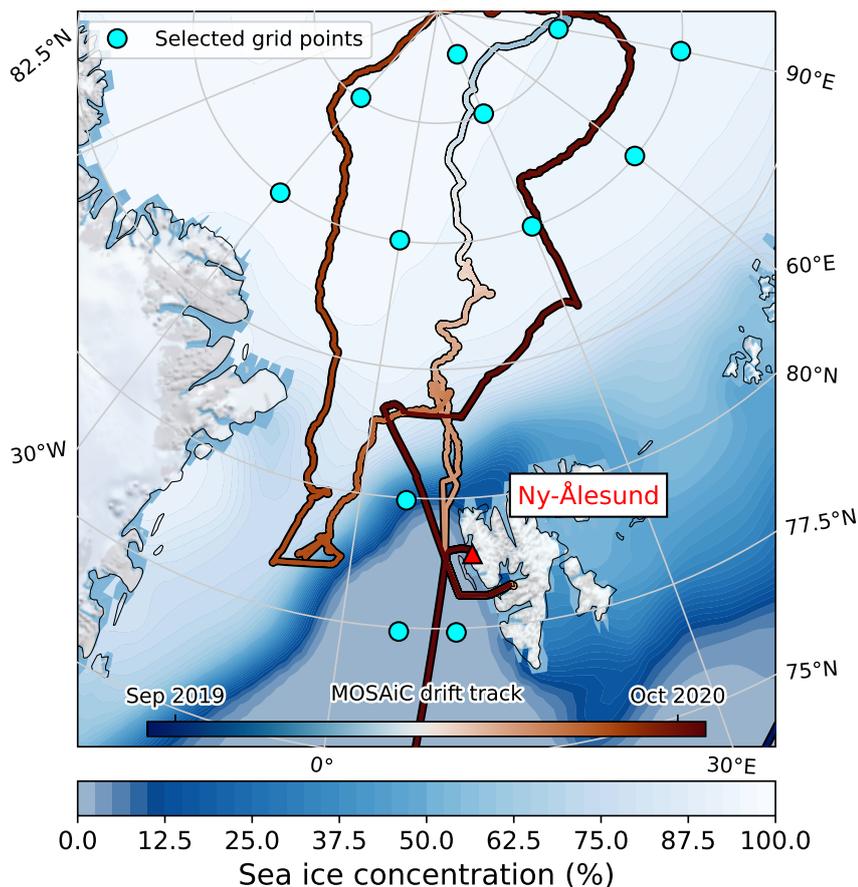


Figure 5.2: Mean sea ice concentration in the Arctic over the years 2001—2018 based on daily ERA5 data at 12 UTC. Light blue circles mark the position of the 12 grid points selected for the retrieval development. The MOSAiC drift track is marked as coloured line with black outline.

5.4.2 Retrieval setup

This study used multilayer perceptron NNs (fully connected layers) to retrieve IWV, specific humidity, and temperature profiles. To optimally use HATPRO’s boundary layer observations, we retrieved temperature profiles from zenith and boundary layer observations separately. The challenge is to develop retrievals that are not overfitted and can therefore adapt well to new data. Overfitting occurs when the retrieval does not only learn the relation between the observations and the atmospheric state but also the (synthetic) noise. Additionally, we wanted to ensure that the retrievals are robust by training an ensemble of 20 NNs with identical settings but with different random number seeds. The random number seeds affect the selection of years for the training and validation data, as well as the NN initialization (weight coefficients). The NNs are considered robust when the errors in the validation data show a small spread over the ensemble of 20 NNs. For example, the spread should be smaller than a given threshold (e.g., 0.2 kg m^{-2} for IWV) or smaller than the magnitude of the error.

To meet the retrieval performance requirements, we developed four NNs with different settings (see Appendix 5.A), one for each retrieved quantity (IWV, specific humidity, and temperature profiles from zenith and boundary layer observations). The retrievals of profiles required deeper networks and stronger regularization measures (e.g., dropout layers, batch normalization, see Appendix 5.A) to avoid overfitting and to achieve a good performance. Besides TBs at different frequencies, we also included seasonal information in the form of the cosine and sine of the day of the year as input to all NNs except for the boundary layer temperature profile (inspired by Billault-Roux and Berne, 2021). Additionally, adding the 2 m temperature and the retrieved IWV as input to the specific humidity profile retrieval slightly reduced errors during validation. Therefore, the specific humidity retrieval can only be performed after the IWV retrieval. For the boundary layer temperature profile, the input vector consists of V-band TBs at various elevation angles (90.0, 30.0, 19.2, 14.4, 11.4, 8.4, 6.6 and 5.4°), which are measured during HATPRO’s boundary layer scan. TBs at other frequencies were not included because they were not measured at these elevation angles. Also, adding other parameters to the input vector did not improve errors. Therefore, the input vector is identical to the one used in the HATPRO regression retrieval described in Walbröl et al. (2022). Further details of the NN retrieval principles and settings can be found in Appendix 5.A.

5.4.3 Metrics for retrieval evaluation and vertical information content

The retrieved state vector \mathbf{x} (e.g., specific humidity profile) is evaluated using the reference $\tilde{\mathbf{x}}$ provided by ERA5 (ERA5 evaluation data set) or MOSAiC radiosondes (MOSAiC evaluation data set). For each component j of the state vector (i.e., j -th height level), we calculate the bias, the root mean squared error (RMSE) and the bias-corrected RMSE:

$$\text{Bias}_j = \frac{1}{N_s} \sum_{i=0}^{N_s} (x_{ij} - \tilde{x}_{ij}) \quad (5.1)$$

$$\text{RMSE}_j = \sqrt{\frac{1}{N_s} \sum_{i=0}^{N_s} (x_{ij} - \tilde{x}_{ij})^2} \quad (5.2)$$

$$\text{RMSE}_{\text{corr } j} = \sqrt{\frac{1}{N_s} \sum_{i=0}^{N_s} ((x_{ij} - \text{Bias}_j) - \tilde{x}_{ij})^2} \quad (5.3)$$

N_s is the number of data samples of the respective evaluation data set. For IWV, we also compute the Pearson product-moment correlation coefficient

$$R = \frac{\sum_{i=0}^{N_s} (\tilde{x}_i - \tilde{\bar{x}})(x_i - \bar{x})}{\sqrt{\sum_{i=0}^{N_s} (\tilde{x}_i - \tilde{\bar{x}})^2 \sum_{i=0}^{N_s} (x_i - \bar{x})^2}}, \quad (5.4)$$

where \bar{x} ($\tilde{\bar{x}}$) is the mean retrieved (reference) state vector.

The vertical information content of passive microwave observations was computed following the ideas of physical retrievals of Rodgers (2000). Due to computation time, the information content was only computed for a randomly selected subset of 4% of the ERA5 evaluation data set (2803 samples). Firstly, we interpolated the vertical grid from the ERA5 model levels to the retrieval height grid and simulated new reference observation vectors \mathbf{y} (here, TBs) with PAMTRA. For these simulations, the retrieval grid has been extended to 45000 m to simulate emissions from gases (mainly oxygen) beyond the retrieval height grid. Secondly, each state vector component is perturbed step by step. We multiply the respective height level by 1.01 for specific humidity profiles, similar to Ebell et al. (2013). Thirdly, we simulate new TBs with PAMTRA for each perturbed state vector. Fourthly, the Jacobian Matrix \mathbf{K} is calculated with entries $K_{aj} = \partial y_{ia} / \partial x_{ij}$ where ∂y_{ia} is the a -th component of the difference between the perturbation-based and reference observation vector of the i -th data sample. ∂x_{ij} is the j -th component (j -th height level) of the difference between the perturbed and reference state vector. Fifthly, the Averaging Kernel matrix \mathbf{A} is computed with $\mathbf{A} = (\mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} \mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}$ where \mathbf{S}_a and \mathbf{S}_ε are the covariance matrices of the state and observation vectors, respectively. \mathbf{S}_ε contains the TB noise on the main diagonal while the remaining entries are 0. \mathbf{S}_a is calculated as full covariance matrix from the ERA5 evaluation data set. Finally, the degrees of freedom (DOF) are inferred from the trace of the Averaging Kernel \mathbf{A} .

5.5 Retrieval evaluation

We applied the retrievals to both the ERA5 evaluation data set and MOSAiC observations (MOSAiC evaluation data set), for which the radiosondes serve as the reference data set. The retrieval evaluation with respect to the ERA5 data allows us to assess the retrievals' theoretical best performance because it is an idealized world without measurement problems. Here, we compute errors for all 20 NNs to get an idea of the spread among the NNs. For the evaluation with the MOSAiC radiosondes, we selected the NN that has a low RMSE and bias in the validation data set while also having the lowest RMSE in the 5% MOSAiC

radiosonde subset that we included in the validation process. Hereafter, this NN is referred to as the final NN.

5.5.1 IWV

The performance of the IWV retrieval applied to the ERA5 and MOSAiC evaluation data sets can be seen in Fig. 5.3. For the ERA5 data, we can evaluate the robustness of the NN through the spread of the errors among all 20 NNs. The RMSE of IWV varies little over the 20 NNs for IWV up to 24 kg m^{-2} , indicated by the low spread ($< 0.3 \text{ kg m}^{-2}$). Only for higher IWV, the spread increases significantly to 0.8 kg m^{-2} . However, only 41 of 70080 ($< 0.1 \%$) of the synthetic data set samples have an IWV above 24 kg m^{-2} . Therefore, errors are computed over a very low fraction of the data and tend to vary more for different NNs. Most importantly, statistical retrievals such as NNs struggle to capture extreme conditions not well represented in the training data set. This can also be seen in the bias, which is close to zero for IWV below 20 kg m^{-2} as expected for a well trained NN, but deviates from zero for higher IWV. However, biases are still small for both the ERA5 and MOSAiC evaluation data sets, staying below 2%.

The RMSE of the final NN, which was selected based on errors in the validation data set, is about 2% of the IWV, and therefore also at the lower end of the 20 NN ensemble for the ERA5 evaluation data set. This shows that the retrieval is well trained because it performs similarly well on the evaluation data set as on the validation data set. For the comparison with MOSAiC observations, where we also use the final NN, the RMSE is slightly higher in most IWV regimes, reaching up to 3–4%. In absolute terms, the RMSE increases from 0.1 to 0.7 kg m^{-2} with IWV increasing from 1 to 29 kg m^{-2} . Here, the additional uncertainties in the radiosonde measurements and matching with the MWR data must be considered.

5.5.2 Specific humidity profiles

We evaluate the retrieved specific humidity profiles (q) in terms of bias and $\text{RMSE}_{\text{corr}}$ for the ERA5 and MOSAiC evaluation data sets (Fig. 5.4). The $\text{RMSE}_{\text{corr}}$ values are similar to $\text{RMSE}_{\text{corr}}$ because of a small bias. For the MOSAiC data, the $\text{RMSE}_{\text{corr}}$ increases from 0.25 g kg^{-1} at the surface to 0.5 g kg^{-1} at 1500 m, which is 15 to 30% of the mean specific humidity (Fig. 5.4b). At higher altitudes, the $\text{RMSE}_{\text{corr}}$ is lower but the relative error increases because the mean specific humidity also decreases. While the $\text{RMSE}_{\text{corr}}$ are generally smaller for the ERA5 data, the shape is similar with the highest $\text{RMSE}_{\text{corr}}$ of about 0.25 g kg^{-1} (15% of the mean q) at 1000 m and even lower values at the surface with 0.15 g kg^{-1} (8%). The $\text{RMSE}_{\text{corr}}$ spread across all 20 NNs is negligible, mostly ranging from

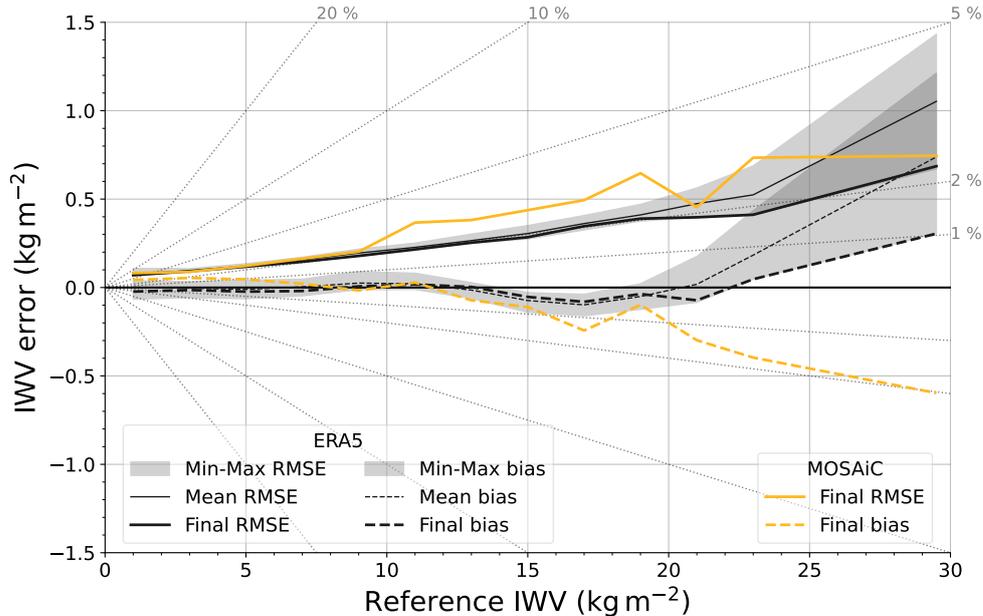


Figure 5.3: IWV errors (RMSE and bias) for certain bins of reference IWV (0–2, 2–4, ..., 22–24, 24–35 kg m⁻²). IWV errors based on the ERA5 (MOSAiC) evaluation data set are displayed in black (yellow). The maximum and minimum spread of RMSE and bias over the 20 Neural Networks are indicated by grey shading. The RMSE (bias) of the mean over the 20 Neural Networks is displayed as a thin solid (dashed) black line. The RMSE (bias) of the final NN is shown as a thick solid (dashed) black line.

0.01 to 0.02 g kg⁻¹.

The mean MOSAiC radiosonde q profile shows the maximum value about 250 m lower than the mean retrieved q profile (Fig. 5.4). Because of the different heights of the humidity inversion, we find the highest RMSE_{corr} and bias slightly above the height level of the maximum q value (at 1500 m). At this height, the retrieved q profile overestimates the radiosonde measurement by up to 0.15 g kg⁻¹ (see bias in Fig. 5.4a). Above 3500 m, the bias remains negative with values up to -0.04 g kg⁻¹ at 5500 m. On the ERA5 evaluation data set, the final NN, which was also used to derive the q profile for MOSAiC, denotes much smaller biases and is slightly negative for all heights (only up to -0.025 g kg⁻¹). However, in the lowest 2000 m, the bias varies much more than the RMSE_{corr}, ranging from -0.1 to +0.1 g kg⁻¹ depending on the chosen NN.

The smaller magnitude of the error profiles in the ERA5 evaluation data set is likely due to the lower complexity of q profiles in ERA5 compared to radiosonde observations. Specific humidity profiles in reanalyses are typically much smoother and do not resolve small inversions (Chellini and Ebell, 2022). Passive microwave observations cannot resolve small inversions and average out strong vertical gradients. Therefore, errors of retrieved profiles are large when compared to radiosonde

data in the presence of strong vertical gradients (i.e., humidity inversions), while the smoother profiles of reanalyses can be captured better. As the retrieval has been trained with reanalysis data, it is also expected to perform best when applied to the same reanalysis. Furthermore, the errors of the evaluation based on real observations can be higher due to measurement errors of radiosondes (noise, sonde drift, systematic errors due to sensor response time, etc.) and of the MWRs (noise, systematic errors).

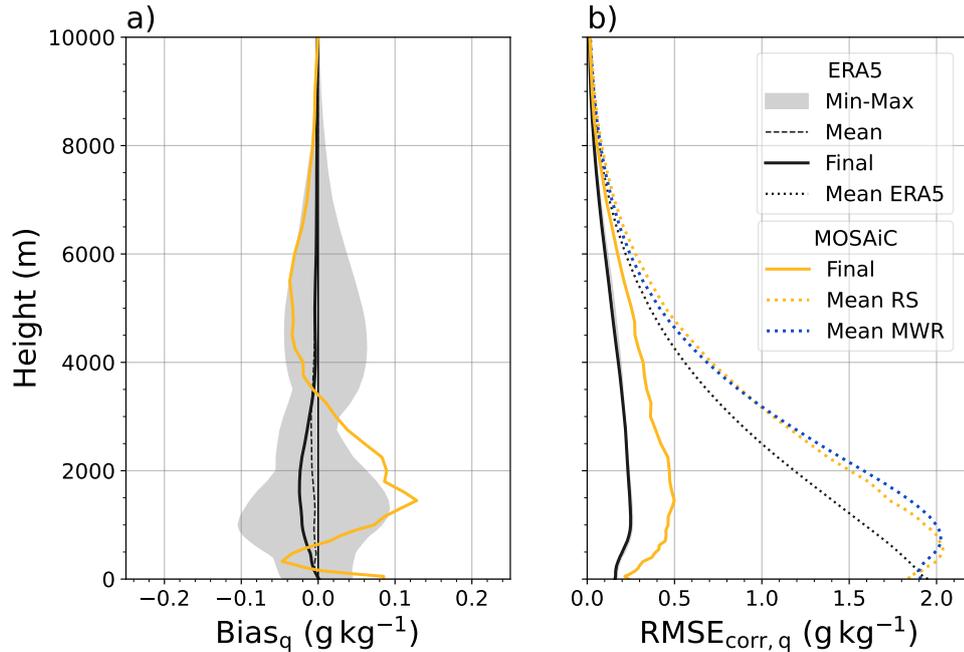


Figure 5.4: Specific humidity q error profiles showing (a) the bias and (b) the bias-corrected RMSE with respect to the reference from the ERA5 and MOSAiC evaluation data sets. The dashed black line in each panel shows the mean over the 20 Neural Networks while shading indicates the min–max spread. The prediction of the final Neural Network is indicated by the thick black (yellow) lines for the ERA5 (MOSAiC) evaluation data set. The mean MOSAiC radiosonde (RS) profile and ERA5 profile are shown as yellow and black dotted lines, respectively, and serve as reference for the absolute error values. The mean retrieved profile from MOSAiC microwave radiometer observations (MWR) is also included as blue dotted line.

5.5.3 Temperature profiles

For the evaluation of the retrieved temperature profiles, we also analyze the bias and RMSE $_{\text{corr}}$ (Fig. 5.5) but distinguish between profiles retrieved from zenith observations (henceforth, zenith temperature profiles) and boundary layer scan (henceforth, BL temperature profiles). As for specific humidity, the spread over the 20 NNs is larger for the bias than for RMSE $_{\text{corr}}$ but generally quite small (especially for BL temperature profiles).

Firstly, we evaluate the zenith temperature profiles: The biases and $\text{RMSE}_{\text{corr}}$ of zenith temperature profiles are larger for the MOSAiC compared to the ERA5 evaluation data set below 1500 m but mostly similar at higher altitudes (see Fig. 5.5a and b). Within the lowest 150 m, the MOSAiC data $\text{RMSE}_{\text{corr}}$ decreases rapidly from 2.9 to 1.4 K. This large $\text{RMSE}_{\text{corr}}$ is associated with near-surface temperature inversions that typically occur in the Arctic. In the ERA5 evaluation data set, this steep error gradient is less pronounced because near-surface temperature inversions over sea ice are not well represented in ERA5. Between about 200 and 2000 m, the $\text{RMSE}_{\text{corr}}$ is between 1.2 and 1.6 K for the MOSAiC and between 0.8 and 1.6 K for the ERA5 evaluation data set. At the top of the retrieval grid at 8000 m, the $\text{RMSE}_{\text{corr}}$ increases to 2.5 K for MOSAiC and 3 K for ERA5.

In the lowest 500 m, the bias of the zenith temperature profiles lies between -1 and $+1$ K for the MOSAiC and between -0.2 and $+0.2$ K for the ERA5 evaluation data set (final NN, see Fig. 5.5a). Here, also the strong surface temperature inversions, which are not well resolved by the retrieved profile, are responsible for the large bias. Above 1500 m, the bias in both data sets is generally smaller than ± 0.2 K. However, the MOSAiC observation bias varies over the seasons: In winter (22 October 2019–30 April 2020), the bias is mostly negative in the mid-troposphere, ranging from -0.4 to -0.8 K, while they are positive in summer (01 May–01 October 2020), ranging from $+0.5$ to $+0.9$ K (not shown).

As expected, biases and $\text{RMSE}_{\text{corr}}$ are smaller for the BL temperature profiles in the lowest 1500 m compared to the zenith temperature profiles (see Fig. 5.5c and d). This result is consistent with the findings of Crewell and Löhnert (2007). For the MOSAiC data, the $\text{RMSE}_{\text{corr}}$ is 2 K at the surface (0.9 K at 100 m) and smaller than 1.2 K up to 1 km height. The error is therefore 1 K (0.5–0.6 K) lower compared to the zenith temperature profile error. Based on the ERA5 evaluation data set, the near-surface $\text{RMSE}_{\text{corr}}$ values are only 0.4–0.5 K, which is lower than for the MOSAiC data because of the less complex temperature profile and the absence of measurement uncertainties. In the lowest 1500 m, also the bias is reduced, being nearly 0 K in the ERA5 evaluation data set (with the final NN), and between -0.6 and $+0.4$ K in the MOSAiC data. Also, the seasonal variation of the MOSAiC BL temperature profile bias is smaller than that of the zenith temperature profiles. Above 2000 m, the $\text{RMSE}_{\text{corr}}$ is similar for both the zenith and BL temperature profiles but the bias above 2000 m is stronger (more negative) in BL temperature profiles, especially for the MOSAiC data (up to -2 K).

We conclude that if the 30 minute temporal resolution is sufficient for the user, a combination of BL profiles and zenith profiles provides optimal performance. We recommend that BL temperature profiles should be used in the lowest 1500 m,

followed by a linear transition to the zenith temperature profile between 1500 and 2000 m and only the zenith temperature profile above 2000 m.

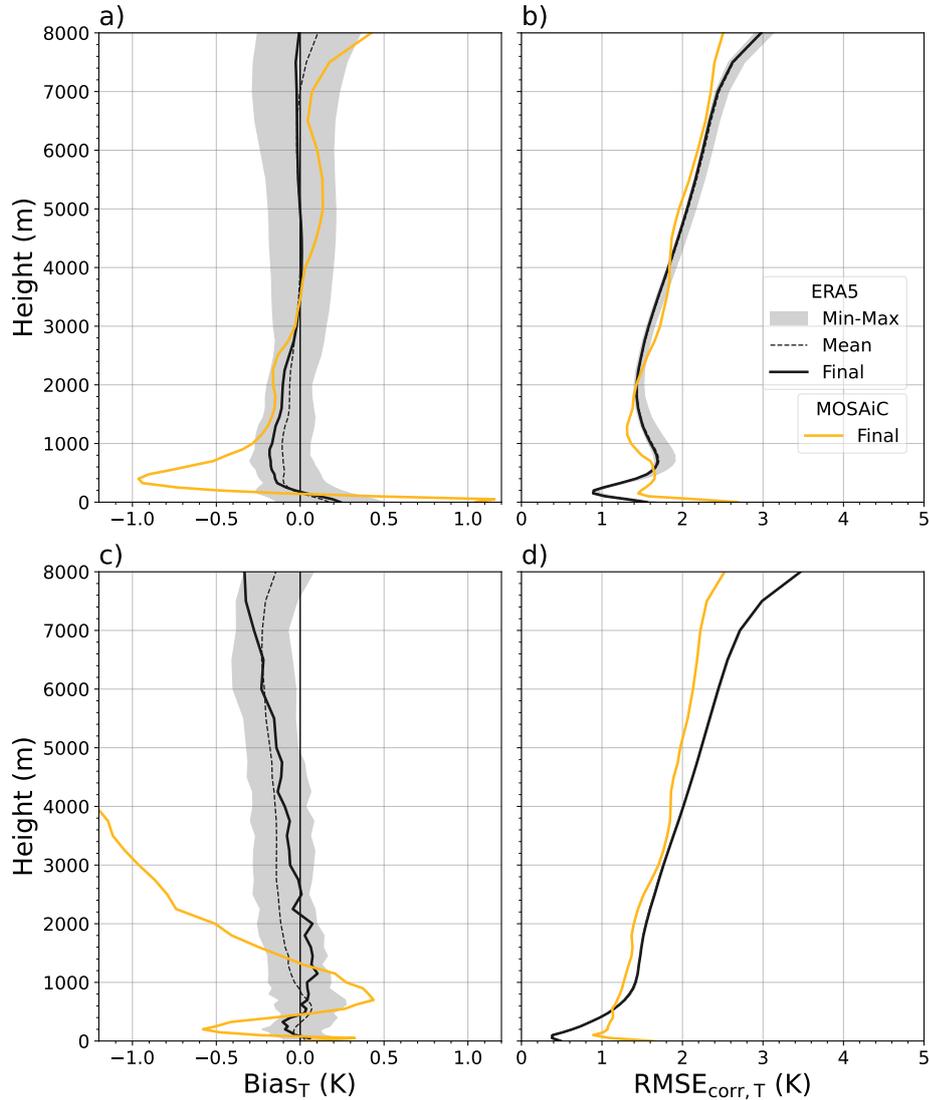


Figure 5.5: Error profiles of (a,b) zenith and (b,d) boundary layer temperature T profiles. Panels (a) and (c) show the bias and panels (b) and (d) the bias-corrected RMSE with respect to the reference from the ERA5 and MOSAiC evaluation data sets. Shading and different line types are similar to Fig. 5.4.

5.6 Information benefit analysis

After introducing the combined HATPRO and MiRAC-P retrieval, it still has to be demonstrated that the synergy is beneficial compared to single instrument retrievals. The benefit is quantified through error reduction and gain in vertical information content. We compare the errors of the synergy with the single instrument retrievals by Walbröl et al. (2022) for MOSAiC observations to present the

improvements for actual observations. As the retrieval methods also differ, we also analyzed the influence of different retrieval architectures (i.e., NN instead of regression) and training data sets (ERA5 instead of Ny-Ålesund radiosondes) on the error reduction compared to HATPRO-only retrievals. This helps to isolate the pure benefit of the combination of low and high frequency microwave observations from potential effects due to different retrieval methods. In Sect. 5.6.1 and 5.6.2, the error estimates for the synergy correspond to the ones shown with respect to MOSAiC radiosondes in Sect. 5.5.1 and 5.5.2.

5.6.1 IWV

Figure 5.6 shows the RMSE and bias of IWV obtained from single instrument observations (HATPRO-only, MiRAC-P-only) and from the synergy of both instruments, with radiosonde IWV as reference. As found in Walbröl et al. (2022), the HATPRO-only IWV retrieval denotes high relative errors and a positive bias ($> 20\%$) for IWV below 5 kg m^{-2} , while having lower relative errors (2–4 %) for IWV greater than 10 kg m^{-2} . For MiRAC-P, the error behaviour is reversed: Small biases and RMSE are found for extremely dry conditions and errors become much larger than the HATPRO-only retrieval for IWV greater than 10 kg m^{-2} .

As expected, the synergy performs similarly well or even better than the single instrument retrievals. For IWV below 5 kg m^{-2} , the RMSE of the synergy is reduced by 75 % compared to HATPRO while being similar to MiRAC-P. The RMSE of the synergy is also smaller by up to 0.2 kg m^{-2} compared to HATPRO-only when IWV is above 5 kg m^{-2} , corresponding to a RMSE reduction of 15–50 %. However, the improvement of RMSE for high IWV is mainly due to the bias reduction from more than -0.5 for HATPRO to -0.1 to -0.5 kg m^{-2} for the synergy. When considering the bias-corrected error ($\text{RMSE}_{\text{corr}}$), the synergy shows up to 20 % higher errors than the HATPRO regression retrieval for IWV above 10 kg m^{-2} (not shown). The error reduction compared to MiRAC-P is even higher in this IWV range.

To study the influence of the different retrieval methods and training data sets, we trained one NN with identical settings as used in the final synergy (see Appendix 5.A, Table 5.1), but included only K-band TBs as input vector. Therefore, the only difference between this NN and the HATPRO regression is the training data (ERA5 vs. Ny-Ålesund radiosondes) and the retrieval type (regression vs. NN). With this NN, we find that RMSE and biases of the retrieved IWV are similar to those of the HATPRO regression retrieval in almost the entire IWV range (see Appendix 5.B, Fig. 5.11). Only in very dry conditions (IWV below 2 kg m^{-2} , the K-band only NN shows 0.1 kg m^{-2} smaller bias and RMSE. Thus,

including the higher frequencies by MiRAC-P dominates the improvement of the error.

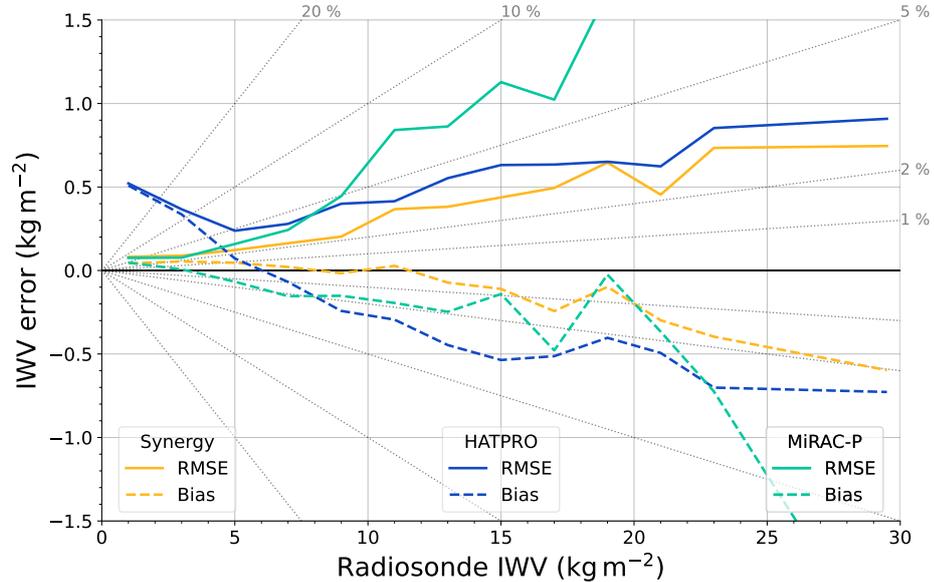


Figure 5.6: RMSE (solid lines) and bias (dashed lines) of IWV retrieved from MOSAiC MWR observations for certain bins of radiosonde IWV (0–2, 2–4, ..., 22–24, 24–35 kg m^{-2}). Yellow lines indicate retrieved IWV from the synergy of HATPRO and MiRAC-P, dark blue lines show HATPRO-only and cyan lines show MiRAC-P-only retrievals.

5.6.2 Specific humidity profiles

In Fig. 5.7, the bias and $\text{RMSE}_{\text{corr}}$ for the specific humidity profiles of the HATPRO regression retrieval and the synergy NN retrieval are shown with respect to MOSAiC radiosondes. At altitudes below 1500 m altitude, the $\text{RMSE}_{\text{corr}}$ is much smaller for the synergy compared to HATPRO. At the surface, the reduction of $\text{RMSE}_{\text{corr}}$ is most prominent, decreasing from 0.5 g kg^{-1} to less than 0.25 g kg^{-1} in absolute terms, and from 30 % to less than 15 % in relative terms (Fig. 5.7b). Above 1500 m, the $\text{RMSE}_{\text{corr}}$ difference between HATPRO and the synergy is marginal and the relative $\text{RMSE}_{\text{corr}}$ gradually increases from 25 to 80 % until the top of the retrieval grid (10000 m). Between the surface and 1000 m, the synergy also shows a much smaller bias (-0.05 to $+0.1 \text{ g kg}^{-1}$) than HATPRO (0.1 to 0.4 g kg^{-1}). The strongest improvement was found near the surface, where the bias is reduced by up to 75 %. Above 1000 m, the bias reduction of the synergy compared to HATPRO is less pronounced: The bias of HATPRO (the synergy) lies between -0.1 and $+0.1 \text{ g kg}^{-1}$ (-0.05 and $+0.15 \text{ g kg}^{-1}$). Therefore, combining both instruments is most beneficial in altitudes below 1500 m in the real-world application.

Because of the different magnitude of specific humidity and the different performances of HATPRO and MiRAC-P over the seasons, we also investigated seasonal differences in error reduction (not shown): In winter (here, 22 October 2019–30 April 2020), the $\text{RMSE}_{\text{corr}}$ is lower for both HATPRO and the synergy as also the water vapour amount is lower. However, the relative $\text{RMSE}_{\text{corr}}$ of the synergy is similar to the error for the full MOSAiC year in the lowest 1000 m while the relative error of the HATPRO retrieval is increased. Therefore, the benefit of the synergy in the lower troposphere is even more pronounced. The synergy also shows smaller errors than HATPRO in the middle and upper troposphere, which was not found for the entire MOSAiC year. The bias reduction of the synergy compared to HATPRO-only is also stronger in winter. In summer (here, 01 May–01 October 2020), the overall picture of the error profiles is similar to the full MOSAiC year, except that the $\text{RMSE}_{\text{corr}}$ s (relative $\text{RMSE}_{\text{corr}}$) for both retrievals are shifted to slightly higher (lower) values. The bias reduction of the synergy compared to HATPRO is also a little less pronounced.

As in Sect. 5.6.1, to identify whether the error reduction is mainly due to the inclusion of the higher frequencies or due to the different training data and retrieval method, we trained one NN with the same setup as the final synergistic retrieval but used only K-band TBs as input. We applied this K-band-only NN retrieval and the HATPRO regression to the ERA5 and MOSAiC evaluation data sets as in Sect. 5.5.2 and found that the $\text{RMSE}_{\text{corr}}$ was almost identical for both retrievals in all height levels (see Appendix 5.B, Fig. 5.12b). Only the bias is closer to 0 for the K-band-only NN than for the regression (Fig. 5.12a). As the results for both retrieval architectures are mostly similar when using the same input vector (K-band TBs), it follows that the inclusion of the higher frequencies contributes most to the overall error reduction.

We also investigated the influence of the additional input parameters (2 m temperature, IWV, day of the year) on the retrieved specific humidity profile. In one experiment, we excluded the MiRAC-P TBs from the input vector of the NN but kept the HATPRO TBs, as well as the day of the year, the IWV and the 2 m temperature. The resulting retrieved specific humidity also shows lower errors than the HATPRO-only regression at the surface (not shown). However, the vertical extent of the benefit is smaller, being mainly confined to the lowest 500 m, compared to the synergistic retrieval including the MiRAC-P TBs. Another experiment, where we used HATPRO and MiRAC-P TBs, as well as the IWV and day of the year as input but excluded the 2 m temperature, showed higher errors in the lowest 100 m. These experiments demonstrate that the MiRAC-P observations are needed to have a higher vertical extent of the error reduction and that the 2 m temperature effectively reduces errors at the surface.

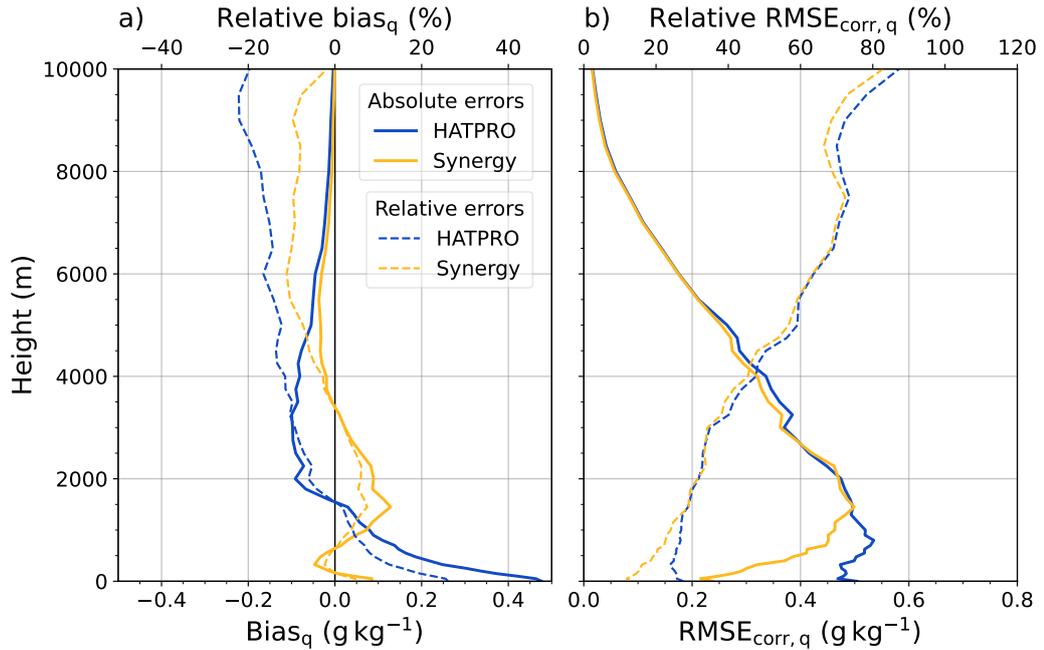


Figure 5.7: Specific humidity q error profiles showing (a) the bias and (b) the bias-corrected RMSE in absolute (solid lines) and relative terms (dashed lines) with respect to MOSAiC radiosondes. Specific humidity errors of the synergy (HATPRO) retrieval are shown in yellow (blue).

To quantify the synergy benefit, it is interesting to analyze not only the error of the retrieved profiles but also their vertical information content. This also offers the opportunity to investigate the impact of the different frequency bands. Thus, we computed the degrees of freedom (DOF) as a measure of the vertical information content for various frequency combinations as described in Sect. 5.4.3. In Fig. 5.8, the statistics of the DOF over a 4% subset of the ERA5 evaluation data set are visualized. When using only K-band frequencies, the specific humidity profile has about 1.7 DOF. Adding the V-band TBs only has a small effect as these frequencies are hardly sensitive to the water vapour amount. The largest increase in the DOF (from 1.7 to 2.4) is caused by the addition of G-band frequencies to the K-band frequencies. This increase is even more pronounced in cold, dry, and clear sky conditions, where the DOF is increased from 1.9–2.1 to 2.7–3.0 (Fig. 5.8). In contrast, the DOF hardly improved from 1.6 to 1.8–2.0 in warm and humid conditions. Clear sky scenes are typically associated with cold and dry conditions during the Arctic winter. The DOF are larger during cold and dry conditions than during warm and humid conditions because the G-band TBs are partly saturated. This means they no longer observe the entire tropospheric column and cannot add as much information. Adding V-band or the 243 and 340 GHz frequencies to K- and G-band TBs only has a minor impact on the DOF distribution.

Ebell et al. (2013) and Löhnert et al. (2009) analyzed the vertical information content of absolute humidity profiles from ground-based MWRs using K-band TBs at different mid-latitude sites and found 2.4 and 1.6 DOF, respectively. Additionally, Löhnert et al. (2009) obtained 2.7 DOF for a tropical site with a much higher mean IWV. Thus, the DOF depends strongly on the frequencies used to derive the humidity profile and the atmospheric conditions. In the Arctic, humidity profiling is more challenging with K-band frequencies due to the lower sensitivity, which is why the higher frequency observations are needed to obtain similar DOF (see also Fig. 5.1).

Based on the Averaging Kernel and the vertical height grid spacing, we can also estimate the theoretical vertical resolution of the specific humidity profiles (e.g., dz_j/A_{jj} where dz is the height grid spacing and A_{jj} the diagonal entries of the Averaging Kernel at height level j). In Fig. 5.9, the estimated vertical resolution (effective resolution) is shown for K-band only and for all frequencies. The other frequency combinations are not discussed as their Averaging Kernel values lie in between those of the K-band and all frequencies. The effective resolution at a certain height level indicates to which vertical resolution the specific humidity profile is smoothed by the microwave observations. Generally, larger values of the effective resolution are found at higher altitudes, consistent with the decreased sensitivity of ground-based microwave observations at these altitudes. The jump of the effective resolution at 5000 m height is due to a strong change in height grid spacing. At the surface, using all frequencies instead of just the K-band improves the effective resolution by a factor of 2 (from 1200 m for K-band to 600 m for all frequencies). At higher altitudes, the relative improvement is smaller, but the absolute resolution improvement is still mostly between 1000 and 2000 m.

5.6.3 Relative humidity profiles

Relative humidity is an important parameter, particularly for cloud processes, and a desired variable for the modeling community. We computed relative humidity from the retrieved temperature, specific humidity profiles and surface air pressure measured by the weather station attached to HATPRO using the hypsometric equation. For HATPRO, the conversion from absolute humidity to relative humidity profiles was straight forward. Due to the bias reduction that we achieved with the new NN retrievals in both the retrieved temperature and specific humidity profiles, we also expect to see lower biases in relative humidity. In the following, we compare the relative humidity bias and $\text{RMSE}_{\text{corr}}$ of HATPRO and the synergy with respect to the MOSAiC radiosondes, which are shown in Fig. 5.10.

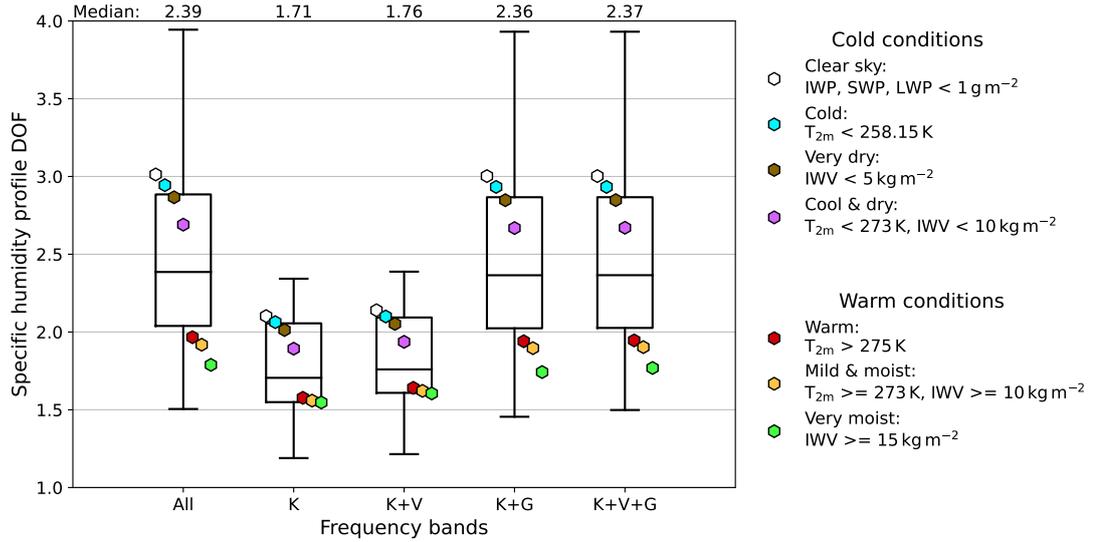


Figure 5.8: Distribution of the degrees of freedom (DOF) over 2803 samples visualized as boxplot for different frequency combinations (all frequencies, K- and V-band, K- and G-band, K-, V- and G-band). The box indicates the interquartile range (IQR, 1–3rd quartile) of the distribution and the horizontal line within the box shows the median. The whiskers extend from below the 1st quartile and above the 3rd quartile by $1.5 \times \text{IQR}$, respectively. Additionally, the median DOFs of different atmospheric conditions have been highlighted.

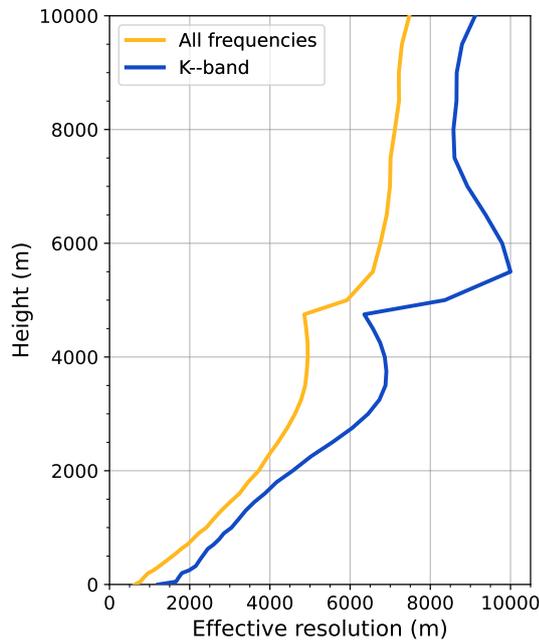


Figure 5.9: Vertical resolution of the specific humidity profiles estimated with the mean Averaging Kernel over the 2803 samples and vertical grid spacing for all frequencies (yellow) and for K-band only (blue).

The bias of the synergy (5%) is much smaller compared to HATPRO (40%) in the lowest 1000 m (Fig. 5.10a). Similarly strong improvements can be found in the lowest 1000 m of the $\text{RMSE}_{\text{corr}}$ profile (Fig. 5.10b), where errors are reduced

from more than 60 to 15% at the surface and from 35–45 to 15% at higher altitudes. Above 2000 m, the $\text{RMSE}_{\text{corr}}$ of HATPRO and the synergy are similar (about 20%), but the bias is closer to 0% while HATPRO shows a negative bias up to -10% .

In cold and clear sky conditions, where IWV and 2 m temperatures were below 10 kg m^{-2} and 273.15 K, respectively, and no clouds were detected by Cloudnet as described in Sect. 5.3.2, the bias reduction is even stronger below 1500 m (Fig. 5.10a). In warm conditions ($\text{IWV} \geq 10 \text{ kg m}^{-2}$, 2 m temperature $\geq 273.15 \text{ K}$), both retrievals perform similarly well, suggesting no benefit of the synergy compared to the HATPRO-only retrieval. If low-level stratus clouds were not respected in the clear sky detection, the $\text{RMSE}_{\text{corr}}$ s of the HATPRO retrieval are up to 10 percentage points higher in the lowest 1000 m, while the errors of the synergy only slightly increased (not shown). In general, the relative humidity errors of the synergy are much less sensitive over these two types of atmospheric conditions (or over the seasons, not shown).

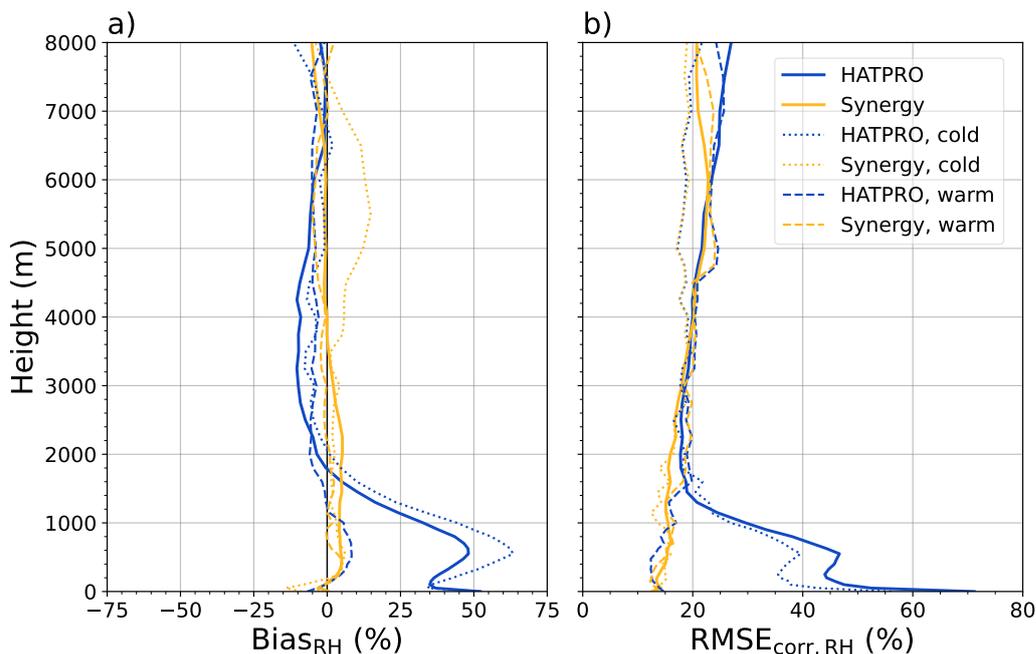


Figure 5.10: Relative humidity error profiles showing (a) the bias and (b) the bias-corrected RMSE with respect to MOSAiC radiosondes. Relative humidity errors of the synergy (HATPRO) retrieval are shown in yellow (blue). Errors are also displayed for different atmospheric conditions: Cold and clear sky (integrated water vapour (IWV) $< 10 \text{ kg m}^{-2}$, 2 m temperature ($T_{2\text{m}}$) $< 273.15 \text{ K}$) as dotted lines and warm ($\text{IWV} \geq 10 \text{ kg m}^{-2}$, $T_{2\text{m}} \geq 273.15 \text{ K}$) as dashed lines.

5.7 Conclusions

In this study, we demonstrate the benefit of combining low (22–58 GHz, HATPRO) and high frequency (175–340 GHz, MiRAC-P) microwave radiometer (MWR) observations for humidity profiling and integrated water vapour (IWV) estimates in Arctic conditions. The newly developed Neural Network (NN) retrievals for IWV and for specific humidity and temperature profiles have been applied to synthetic measurements based on ERA5 and real observations from the MOSAiC expedition. Subsequently, they have been evaluated with ERA5 data and MOSAiC radiosondes, respectively, and compared to the retrievals by Walbröl et al. (2022). Retrieved temperature and specific humidity profiles were used to compute relative humidity together with the surface air pressure from the weather station attached to HATPRO.

We illustrate the sensitivity of the NN to random perturbations with an ensemble of 20 NNs. The spread of errors over the 20 NNs is generally small, except for specific humidity biases. We selected one NN, whose errors were on the lower end of the spread during the retrieval development, as the final NN. Also in the final evaluation, the final NN denoted one of the smallest errors of all 20 NNs. In the following paragraphs, we only summarize retrieval errors with respect to MOSAiC radiosondes as these errors are typically larger than the theoretical ones based on the ERA5 evaluation data set: For IWV, the RMSE is about 3–4 % and biases are smaller than 2 % over a wide range of IWV conditions. Specific humidity is overestimated by up to $+0.15 \text{ g kg}^{-1}$ at 1500 m relative to radiosondes. In other height levels, the biases are smaller. The bias-corrected RMSE ($\text{RMSE}_{\text{corr}}$) is also highest at 1500 m with 0.5 g kg^{-1} (about 30 %). Temperature profile $\text{RMSE}_{\text{corr}}$ (biases) from zenith MWR observations lie between 1.4 and 2.9 K (-1 and $+1$ K) in the lowest 1500 m. Temperature profiles retrieved from boundary layer MWR observations showed much smaller errors in that height range, which is consistent with the findings of Crewell and Löhnert (2007).

In the next step, we compared the errors of the new synergistic NN retrievals to the single MWR retrievals of Walbröl et al. (2022) to estimate the information benefit. Additionally, we computed the vertical information content of specific humidity profiles as degrees of freedom (DOF). The information benefit is only shown for MOSAiC observations to obtain the benefit for the real measurements. IWV errors of the synergy are generally smaller than or similar to those of the single MWR retrievals. In cases when IWV is greater than 10 kg m^{-2} , the RMSE of the synergy is at least 15 % smaller than the HATPRO-only retrieval, which is mainly due to the lower biases of the synergy.

For specific humidity profiles, the largest information benefit was found. The

combination of HATPRO and MiRAC-P increased the DOF from 1.7 to 2.4 and reduced the $\text{RMSE}_{\text{corr}}$ by up to 50 %. Through the synergy, strong positive biases below 1000 m could also be reduced by up to 75 %. The benefit is most distinct in the lowest 1500 m because here, the error reduction is the strongest. At these heights, the synergy enhanced the effective vertical resolution of the specific humidity profile by a factor of up to 2 compared to the HATPRO-only retrieval (from 1200 m to 600 m). In cold and dry conditions, the DOF increase and the error reduction were even more pronounced.

We also analyzed the influence of additional NN input parameters (2 m temperature, day of the year, and IWV) on the specific humidity profile errors and found that including the 2 m temperature is important to minimize errors at the surface. Because of the improvements in specific humidity (and temperature) profiles, the synergy also results in lower relative humidity errors compared to the HATPRO-only retrieval, which is particularly evident in the lowest 1500 m. Additionally, the errors of the relative humidity profiles from the synergy vary much less over different atmospheric conditions than those from the HATPRO-only retrieval.

Coming back to the research questions listed in Sect. 5.2, we can conclude:

1. For specific humidity profiles, the bias-corrected RMSE could be reduced by up to 50 %. Bias reductions are partly even higher. The information benefit is mainly attributed to the combination of HATPRO and MiRAC-P. The different retrieval training data and methods only had a small influence.
2. The vertical information content in the specific humidity profile was increased by 40 %.
3. The combination of HATPRO and MiRAC-P frequencies increased the vertical information content in particular during cold and dry conditions and the least during moist and warm conditions.

HATPROs are used at different sites worldwide (polar, mid-latitude, and subtropical regions). In dry regions (high altitude or polar sites), the observation network would clearly benefit from an instrument that includes the G-band frequencies for IWV and humidity profiling (relative and specific humidity) as these frequencies increased the DOF the most. It is planned to install MiRAC-P at Ny-Ålesund again in 2025 to enhance the continuous atmospheric observations at the German–French research station AWIPEV. We are confident that adding MiRAC-P to the already installed HATPRO will improve humidity profiling similarly as demonstrated for the MOSAiC expedition. The low specific humidity profile errors give us confidence that the synergy is suitable for gaining insights

into the general structure of Arctic humidity profiles (i.e., inversions). However, a detailed analysis of the ability of the synergy to identify humidity inversions is still missing.

In the next step, the enhanced water vapour products from the synergy of HATPRO and MiRAC-P, as well as the radiosonde measurements from MOSAiC, will be used to quantify IWV and specific humidity errors of satellite products and reanalyses. As reanalyses assimilated the MOSAiC radiosonde observations, this comparison likely does not reflect the true performance of the reanalyses in the central Arctic. With the considerable specific humidity profile improvements of the synergy compared to HATPRO, the question arises how well humidity inversions, which are important for cloud formation and maintenance, are captured. This question will be answered with a statistical analysis for the entire MOSAiC period. We will then evaluate the representation of humidity inversions in satellite products and reanalyses compared to observations from the MWRs and radiosondes. Radiative transfer simulations allow us to assess how biases in humidity inversion characteristics affect the downwelling thermal infrared radiation.

Code and data availability

The retrieved synergistic profiles of temperature, specific humidity and relative humidity, as well as integrated water vapour are available on PANGAEA (Walbröl et al., 2024b,a). The retrievals are based on brightness temperature observations from HATPRO (Engelmann et al., 2022) and MiRAC-P (Walbröl et al., 2022a). We used the single instrument retrievals of temperature, absolute humidity and IWV from HATPRO (Ebell et al., 2022) and IWV from MiRAC-P (Walbröl et al., 2022b) for the benefit estimation. Radiosonde measurements from MOSAiC (Maturilli et al., 2021) and the Polarstern track data (Rex, 2020; Haas, 2020; Kanzow, 2020; Rex, 2021a,b) are also available on PANGAEA. Cloudnet target classification, as well as the low-level stratus mask and the additional quality flag data, are available and can be accessed via Engelmann et al. (2023), Griesche et al. (2023) and Griesche and Seifert (2023), respectively. Met City observations have been downloaded from Cox et al. (2023a). On Zenodo, we published the retrieval training, test and evaluation data (Walbröl and Mech, 2024), the information content estimation output (Walbröl, 2024b), and the ERA5 evaluation data predictions and reference (Walbröl, 2024a). A snapshot of the GitHub repository containing the scripts is also archived (Walbröl, 2024d). The PAMTRA code can be accessed via Mech et al. (2019b). The simulated brightness temperatures of the two radiosoundings shown in Fig. 5.1 can be found at Walbröl (2024c).

5.A Neural Network retrieval details

As noted in Sect. 5.4.2, all NNs in this manuscript are multilayer perceptrons (fully connected layers), but some include dropout layers and batch normalization (see Table 5.1), and have been created with Python’s Keras module (contained in Tensorflow, Abadi et al., 2015). The forward propagation of a simple, fully connected NN starts with an input layer whose number of nodes equals the number of components of the input vector. The mathematical operations to propagate to the next layer of the network are similar to multiple linear regression: Each node is multiplied by a randomly initialized weight before being summed up and a bias coefficient is added. Afterwards, the result is used as input to a so-called activation function (e.g., exponential or rectified linear unit, also known as relu). The output of the activation function is then forwarded to each node of the next layer where the process is repeated until the output layer is reached. We always use a linear activation function between the last hidden layer and the output layer. The output layer represents the prediction of the NN and is compared to the truth of the training and validation data sets using a certain loss function (here, mean squared error).

To minimize the loss function, an optimization algorithm (e.g., gradient descent) adapts the weights of each node in a backpropagation process. In this study, we used the Adam optimization algorithm (Kingma and Ba, 2017). The learning rate can be adjusted to reduce or enhance the magnitude of the gradient during backpropagation, leading to slower and smoother or faster and more erratic learning. The NN typically processes a specific number of training data samples, determined by the chosen batch size, before updating the weights. The epoch number determines the maximum number of times the training data set is cycled through. In our retrievals, we activated the EarlyStopping function implemented in Keras that monitors the loss of the validation data set over the epoch numbers. The training was terminated if the validation loss did not improve by more than the minimum delta value for a certain number of epochs (callback patience).

Dropout and batch normalization layers are tools to regularize the NN to make it less prone to overfitting. If batch normalization is set to True for a retrieval (see Table 5.1), we included a batch normalization layer after each hidden layer. It normalizes the output of the preceding hidden layer so that its mean (standard deviation) is close to 0 (1). The dropout chance noted in Table 5.1 indicates the chance that the value of a node is set to 0 during training. If the dropout chance is > 0.0 , we added a dropout layer after each hidden layer or, if applicable, after each batch normalization layer.

Table 5.1: Neural Network settings for each retrieved variable (IWV, specific humidity (q), zenith and boundary layer temperature profiles (T_{zenith} , T_{BL})). DOY_1 and DOY_2 are the cosine and sine of the Day Of the Year and T2m is the 2 m air temperature. Details can be found in the text.

Settings	IWV	q	T_{zenith}	T_{BL}
Input vector	TBs at K,G,243,340, DOY_1, DOY_2	TBs at K,V,G,243,340, T2m, IWV, DOY_1, DOY_2	TBs at K,V,243,340, DOY_1, DOY_2	TBs at V, different elevation angles
N hidden layers	2	3	2	2
N nodes per layer	(16,16)	(64,64,64)	(256,256)	(256,256)
Activation function	exponential	softmax	relu	linear
Dropout	0.0	0.2	0.1	0.0
Batch normalization	False	True	True	True
Batch size	64	256	256	256
Epoch number	15	100	150	800
Learning rate	0.0005	0.0005	0.0003	0.00005
Callback patience	3	30	15	80
Minimum Delta	0.001	0	0	0

5.B Information benefit: Influence of different method

Figure 5.11 shows the IWV error with respect to MOSAiC radiosondes for the old single instrument retrievals (HATPRO regression, MiRAC-P only NN) and the new NN retrieval. However, in this case, the input vector of the NN consists of K-band TBs only. This demonstrates that the different retrieval method and training data compared to the HATPRO regression is not responsible for the error reduction in dry conditions seen in Fig. 5.6 and discussed in Sect. 5.6.1.

Similarly, the specific humidity error profiles for the HATPRO regression and the NN using only K-band TBs are shown in Fig. 5.12. The $RMSE_{corr}$ of both retrievals is comparable for all height levels but the lower tropospheric bias of the NN, labeled as synergy, is smaller. Therefore, the strong $RMSE_{corr}$ reduction is solely caused by including the higher frequencies in the retrieval. However, the different method and training data set seem to contribute a little to the bias reduction.

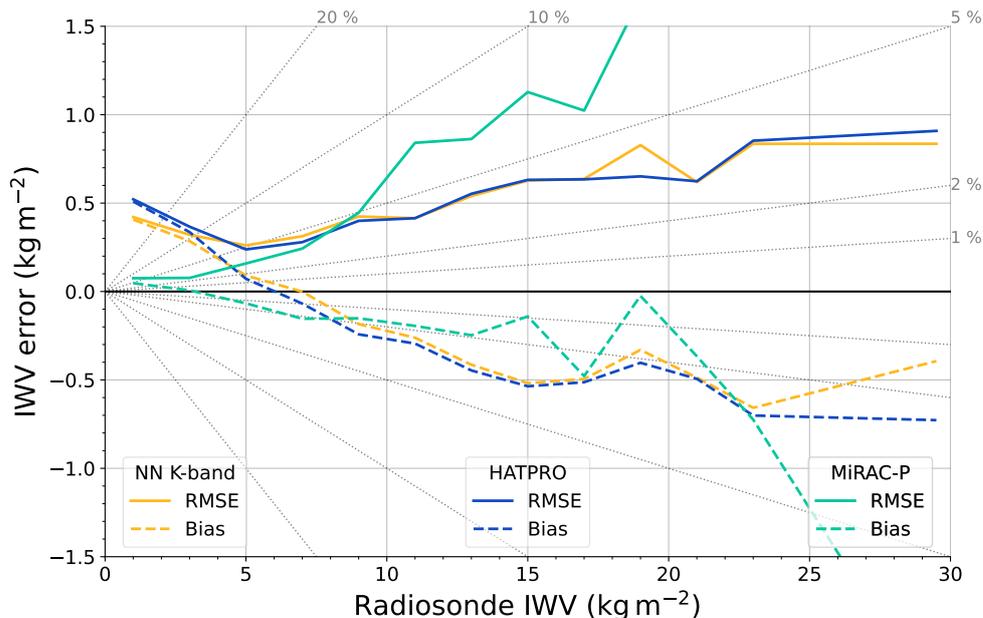


Figure 5.11: As Fig. 5.6 but using only K-band TBs as input vector to the new NN retrieval.

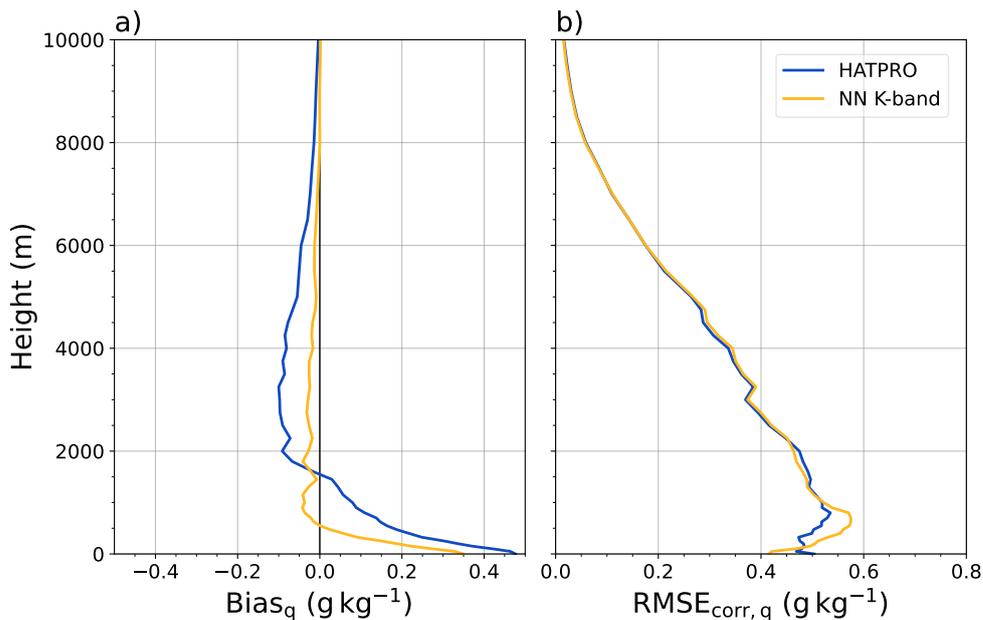


Figure 5.12: As Fig. 5.7 but using only K-band TBs as input vector to the NN retrieval (yellow).

5.C Comparison with smoothed radiosonde profiles

For a fair comparison of the retrieved and the radiosonde specific humidity profiles, the latter can be smoothed to the retrieval height resolution when the Averaging Kernel (AK) is available. Following Löhnert and Maier (2012), we compute

the smoothed specific humidity profiles

$$\mathbf{q}_{\text{smoothed}} = \mathbf{q}_{\text{ret}} + \mathbf{A}(\mathbf{q}_{\text{rs}} - \mathbf{q}_{\text{ret}}), \quad (5.5)$$

where \mathbf{q}_{ret} and \mathbf{q}_{rs} are the retrieved and radiosonde specific humidity profiles, respectively, and \mathbf{A} is the AK. In Fig. 5.13, the specific humidity errors with respect to the smoothed radiosonde profiles are shown. The displayed errors are therefore resolution-corrected. For HATPRO, the smoothing-based errors are much smaller compared to the true errors shown in Fig. 5.7 because the smoothing filtered out the humidity inversions. At the resolution of the retrieved HATPRO profile, the HATPRO-only retrieval can extract more information than the synergistic retrieval at the resolution of the synergy profile because the errors are slightly smaller (yellow and blue solid lines in Fig. 5.13a, b). However, when comparing the specific humidity from HATPRO with the radiosonde profile smoothed with the synergy (thus, slightly higher resolution), the errors are again similar to Fig. 5.7. Thus, the radiosonde profile smoothed with the synergy seems to represent the average true radiosonde profile relatively well. At the surface, the resolution-corrected $\text{RMSE}_{\text{corr}}$ (Fig. 5.13b) of the synergy is similar to the true $\text{RMSE}_{\text{corr}}$ (Fig. 5.7b). At heights where the resolution-corrected errors are lower than the true errors, e.g., around 1500 m, the low vertical resolution of the retrieval is a significant limitation.

In the specific humidity profile example (Fig. 5.13c), the effect of the different smoothing strengths can be seen. The synergistic retrieval and the radiosonde profile at the resolution of the synergy can both identify the inversion observed by the radiosonde well. However, the specific humidity retrieved by HATPRO does not sense a strong humidity inversion near the surface. The strong overestimation of the HATPRO specific humidity compared to the HATPRO-smoothed radiosonde profile near the surface suggests that the resolution could only partly explain this deviation to the true radiosonde profile.

Author contributions

AW and MM were involved in the retrieval preparation. AW has performed retrieval development, evaluation, information benefit analysis. All visualizations have been created by AW. AW, KE, HG and SC conceptualized this study and discussed the results. All authors reviewed this manuscript.

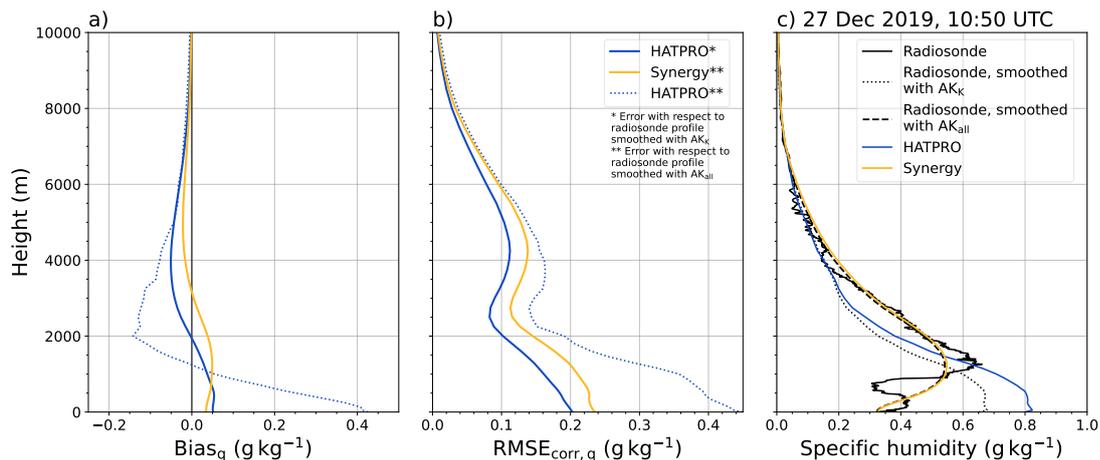


Figure 5.13: Specific humidity q error profiles showing (a) the bias, (b) the bias-corrected RMSE of q retrieved from HATPRO (blue) and the synergy (yellow), and (c) an example specific humidity profile from 27 December 2019 at 10:50 UTC. In (a) and (b), the errors are computed with respect to radiosonde profiles that have been smoothed with the K-band based Averaging Kernel (AK_K) and the all-frequency Averaging Kernel (AK_{all}), respectively. In (c), the original radiosonde profile is shown as black solid line, while the profile smoothed with AK_K (AK_{all}) is visualized as dotted (dashed) black line. The retrieved profiles from HATPRO (synergy) are displayed as blue (yellow) solid lines.

Acknowledgements

We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for the ArctiC amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)³ Project Number 268020496 — TRR 172 within the Transregional Collaborative Research Center. Data used in this manuscript was produced as part of the international Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAiC20192020 and the Polarstern expedition AWI_PS122_00. We thank all those who contributed to MOSAiC and made this endeavour possible (Nixdorf et al., 2021). The microwave radiometer HATPRO was funded by Federal Ministry of Education and Research (BMBF) under FKZ: 01LKL1603A. We acknowledge the support from the OCEANET-Atmosphere project, funded by the German Federal Ministry for Education and Research (BMBF) via the SCiAMO project (MOSAIC-FKZ 03F0915A), in which frame the two microwave radiometers were operated. Radiosonde data were obtained through a partnership between the leading Alfred Wegener Institute (AWI), the atmospheric radiation measurement (ARM) user facility, a US Department of Energy facility managed by the Biological and Environmental Research Program, and the German Weather Service (DWD). ERA5 data (Hersbach et al., 2018) were downloaded from the Copernicus Climate Change Service (C3S) Climate

Data Store. The results contain modified Copernicus Climate Change Service information 2022. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bb1320. Finally, I appreciate the discussions within the working group and with my coauthors.

Chapter 6

Evaluation of water vapour products and assessing the importance of humidity inversions

6.1 Introduction

Incoming radiation from the sun (solar or shortwave radiation) is partly absorbed and partly scattered by the Earth's surface and atmosphere. The absorption of solar radiation heats the Earth's surface (and, to a lesser extent, the atmosphere), which then emits thermal infrared radiation according to its temperature (terrestrial or longwave radiation). If the atmosphere was absolutely transparent in the thermal infrared, the radiative equilibrium temperature of the Earth would be approximately 255 K (Liou, 2002, Chapter 4.1). However, due to the presence of gases in the atmosphere that strongly absorb in the thermal infrared (Tyndall, 1861), radiation emitted by the surface (and lower atmosphere) is absorbed and re-emitted back towards the surface, trapping a part of the emitted infrared radiation (e.g., Raval and Ramanathan, 1989). This effect is known as the greenhouse effect, which is responsible for rather habitable global average temperatures of 288 K (Liou, 2002, Chapter 4.1). The strongest greenhouse gas is water vapour, which absorbs infrared radiation at several resonant absorption lines (most importantly the vibrational-rotational band at $6.25 \mu\text{m}$ and two purely vibrational bands at around $2.7 \mu\text{m}$) and due to the water vapour continuum absorption (Liou, 2002, Chapter 4.2). Even in dry conditions, water vapour dominates the downwelling longwave radiation (DLR) compared to other greenhouse gases (Town et al., 2005).

Due to the high reflectivity of the sea ice and missing solar radiation during polar night, the surface and tropospheric temperatures in the Arctic Ocean are

mainly governed by advection and longwave radiation. Therefore, strong correlations between the DLR, water vapour, temperature and cloud cover are expected and have been confirmed by several studies (Doyle et al., 2011; Ghatak and Miller, 2013; Maturilli et al., 2015; Clark et al., 2021). For example, intrusions of warm and moist air (e.g., due to extratropical cyclones) can increase the DLR on the order of $50\text{--}100\text{ W m}^{-2}$ (Woods and Caballero, 2016; Bresson et al., 2022), which enhances sea ice melt in spring or reduces sea ice growth in winter (Tjernström et al., 2015; Johansson et al., 2017; Graham et al., 2019c). Kapsch et al. (2013) found the sea ice extent at the end of the melt season to be lower because of warm air intrusions during winter and spring. Further, Alekseev et al. (2019) found a clear negative correlation between sea ice extent and DLR over the years 1979–2014. All these findings demonstrate the importance of DLR for the Arctic climate system and the strong connection between DLR and water vapour.

DLR increases approximately logarithmically with the integrated water vapour (IWV), but quantification of this relationship significantly varies between studies (e.g., Zhang et al., 2001; Ghatak and Miller, 2013). Ghatak and Miller (2013) estimated the IWV–DLR relationship with reanalyses over the years 1979–2011 and found $\Delta\text{DLR}/\Delta\text{IWV}$ of $18\text{--}31\text{ W kg}^{-1}$ in dry conditions (Arctic winter) and much lower values in humid conditions due to a saturation of this relationship. As their analysis also included temperature changes, they expect the relationship to be overestimated. Zhang et al. (2001) obtained even larger $\Delta\text{DLR}/\Delta\text{IWV}$ values (up to 100 W kg^{-1}) with radiative transfer calculations but considered a wider IWV range, including also drier conditions than Ghatak and Miller (2013). Their results also suggest that even in humid conditions, the sensitivity of DLR to a unit change in IWV is greater than the sensitivity of DLR to a unit change in the mean atmospheric temperature. Zhang et al. (2001) concluded that the IWV impacts snow melt more than the mean atmospheric temperature.

Throughout the Arctic, the vertical water vapour distribution is characterized by the frequent presence of inversions, where the water vapour concentration (here, specific humidity) increases with height (Devasthale et al., 2011). Surface-based humidity inversions are formed due to radiative cooling in clear sky conditions during polar night, resulting in the condensation of water vapour (Curry, 1983; Brunke et al., 2015). The condensates are then removed from the atmosphere by gravitational settling. Due to the strong connection between the water vapour concentration and the temperature through the Clausius-Clapeyron equation, such surface-based temperature and humidity inversions often coincide (Sedlar et al., 2012; Nygård et al., 2014). Elevated inversions, which occur more frequently in summer, are mainly formed by moisture advection from lower latitudes and the air mass transformation processes in the boundary layer. As a

warm and moist air mass propagates from the relatively warm ocean over the sea ice, the low altitude air is cooled and condensation occurs (Tjernström et al., 2019). The resulting low-level clouds release latent heat and increase the downwelling longwave radiation, warming the sub-cloud layer (Tjernström et al., 2019; You et al., 2021). Humidity (and temperature) inversions are usually located near the top of the cloud layer (Tjernström et al., 2004; Brunke et al., 2015) and act as a moisture source for such low-level clouds via entrainment (Solomon et al., 2011). Naakka et al. (2018) found that temperature inversions rarely occurred above 2 km, while humidity inversions did occur due to moist air advection.

Humidity inversion characteristics have a seasonal cycle due to the different formation mechanisms: In winter, inversions occur most frequently ($> 90\%$ of the time) and are often surface-based while in summer, inversions tend to occur less frequently (70–90%), and are elevated and stronger (higher difference between minimum and maximum specific humidity) (Devasthale et al., 2011; Nygård et al., 2014; Naakka et al., 2018). The frequency of occurrence of humidity inversions generally decreases with height (e.g., $< 40\%$ above 2 km, Naakka et al., 2018). Devasthale et al. (2011) found that radiosonde profiles frequently denoted 2–3 humidity inversions (up to 8), of which the weak ones were usually missed by coarse resolution products (e.g., satellites).

In dry conditions ($IWV < 5 \text{ kg m}^{-2}$), humidity inversions can contain up to 50% of the IWV and are therefore expected to strongly contribute to the DLR (Devasthale et al., 2011). Tjernström et al. (2019) also found that the presence of humidity inversions increased the surface energy budget by 10–25 W m^{-2} due to increased turbulent and longwave radiative heat fluxes. However, in their analysis, humidity inversions were accompanied by clouds and higher IWV, which likely dominated the surface energy budget response. Nygård et al. (2014) suggested that weak inversions have an indirect radiative effect on the DLR due to cloud formation and maintenance.

Several studies have investigated the effect of water vapour at different altitudes on longwave radiation but instead focused on upwelling longwave radiation at the top of the atmosphere (TOA) instead of the DLR (e.g., Shine and Sinha, 1991; Colman, 2001; Held and Soden, 2000; Inamdar et al., 2004). However, due to the stable stratification in the Arctic, the DLR is much more sensitive to low altitude changes in greenhouse gas concentrations than the upwelling longwave radiation at TOA (Colman, 2001; Graversen and Wang, 2009; Bintanja et al., 2011). Ohmura (2001) quantified the contribution of different atmospheric layers to the DLR (60% from the first 100 m, 90% from the first 1 km) but did not disentangle the effects of temperature and water vapour. Thus, an investigation of the impact of the vertical water vapour distribution on the DLR in

the Arctic is missing. Considering the high occurrence of humidity inversions, such an investigation is important to understand the radiative effect of humidity inversions.

Current reanalyses and forecast systems still struggle with the representation of the stable boundary layers in polar regions (Ganeshan and Yang, 2019; Graham et al., 2019a; Solomon et al., 2023), resulting in biases of the near-surface thermodynamic profiles. Additionally, reanalyses have higher uncertainties in the Arctic Ocean due to the sparsity of ground stations and difficulties in satellite observations (Gelaro et al., 2017; Avila-Diaz et al., 2021). For example, the monthly mean IWV varies by 30 % across satellite products and by 20 % across four modern reanalyses in the central Arctic in early summer (Crewell et al., 2021). When estimating IWV trends with reanalyses, the differences across reanalyses can even exceed the magnitude of the trend itself in specific regions and seasons (especially summer) (Rinke et al., 2019).

The European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis v5 (ERA5) is one of the best-performing global reanalysis, especially concerning cloud representation. However, over sea ice, ERA5 has a distinct warm bias near the surface in cold and stable conditions during polar night (partly > 3 K Graham et al., 2019a; Wang et al., 2019; Di Biagio et al., 2021; Herrmannsdörfer et al., 2023). The warm bias is caused by the lack of snow on the sea ice and an underestimation of the sea ice thickness, leading to excessive conductive heat fluxes from the warm ocean through the sea ice to the much colder lower atmosphere (Batrak and Müller, 2019). For the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) from NASA’s Global Modeling and Assimilation Office (GMAO), Gelaro et al. (2017) and Graham et al. (2019a) also found a warm bias of 1.2–3 K over sea ice in winter compared to field campaign data. Graham et al. (2019a) and Di Biagio et al. (2021) also identified longwave radiation biases in current reanalyses (also ERA5 and MERRA-2), which were related to uncertainties in cloud and surface temperature inversion representations.

The detailed water vapour measurements gathered during the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAIC) expedition provide excellent reference data to evaluate state-of-the-art models and satellite products in all seasons. This study particularly focuses on the representation of humidity inversions and their longwave radiative effect. High-quality measurements from radiosondes, as well as the two microwave radiometers (MWRs) HATPRO (Humidity and Temperature PROfiler) and MiRAC-P (Microwave Radiometer for Arctic Clouds - Passive) are used as reference data. To our knowledge, humidity inversions have not yet been studied from ground-

based MWRs due to their limited vertical resolution. However, in our past study, we improved the resolution by a factor of 2 compared to standard single-MWR retrievals (Chapter 5 of this thesis or Walbröl et al., 2024), providing a better framework to resolve the main humidity inversion. Thus, the following research questions are addressed:

1. How well do current models represent the IWV and the vertical water vapour distribution?
2. Are water vapour profiles and the typical Arctic humidity inversions sufficiently well captured by satellite and ground-based remote sensing to correctly assess the DLR?
3. Is an accurate IWV estimate sufficient to calculate DLR? To what extent does the vertical water vapour distribution, especially with regard to inversions, affect the DLR?

The models evaluated in this study include the reanalyses ERA5 and MERRA-2, as well as two state-of-the-art weather forecast systems: The ICOSahedral Non-hydrostatic (ICON) model developed by the German Weather Service and the Max Planck Institute for Meteorology is a global weather forecast model that is also frequently used in Arctic research (Kretzschmar et al., 2019; Schemann and Ebell, 2020; Bresson et al., 2022). The Coupled Arctic Forecast System (CAFS) model from the National Oceanic and Atmospheric Administration Physical Sciences Laboratory is a regional forecast model specialized for Arctic conditions (Solomon et al., 2023, 2024). From the satellite perspective, we evaluate the sounding product of the Metop satellite series, which performed best over sea ice (Crewell et al., 2021), because it combines microwave and infrared observations. For the IWV comparison, we also include a new multi-parameter retrieval based on microwave radiances from AMSR2 that is a step towards tackling the sea ice problem in satellite-based microwave remote sensing.

After describing the reference data, as well as the evaluated models and satellite products in Sect. 6.2 and the methods in Sect. 6.3, the results of the IWV and specific humidity evaluation are presented in Sect. 6.4. Subsequently, Sect. 6.5 follows with an analysis of the detectability of humidity inversions and the representation of inversion characteristics with respect to radiosondes. Finally, radiative transfer simulations are used to analyze the sensitivity of DLR to water vapour within humidity inversions in Sect. 6.6.

6.2 Data

In this study, we analyze water vapour products from satellite observations and different types of numerical models with respect to reference observations from the MOSAiC expedition. Additional MOSAiC observations are used as input to a radiative transfer model to validate the setup and to distinguish clear sky from cloudy conditions. Therefore, it is convenient to separate the data sets into MOSAiC observations (radiosondes, MWRs, lidar), numerical models and satellite products. The data sets are presented in the following sections.

6.2.1 MOSAiC observations

Radiosondes

In contrast to Chapters 4 and 5, an advanced radiosonde product developed by Dahlke et al. (2023) is used in this study. The radiosonde observations at 30 m height are merged with measurements at 0, 2, 6 and 10 m of the tower (Shupe et al., 2022) that has been operated on the sea ice during most parts of the MOSAiC expedition. The tower was located at the measurement site 'Met City', which was established in 300–600 m distance from RV *Polarstern* to minimize the influence of the ship on the measurements (Shupe et al., 2022). Vaisala PTU300 and HMT330 were deployed for temperature and relative humidity measurements at the different heights of the tower. According to the manufacturer, measurement uncertainties in the conditions encountered during MOSAiC (near-surface air temperatures down to -40°C) are not higher than $\pm 0.4\text{ K}$ and $\pm 1.5\%$ for temperature and relative humidity, respectively (Vaisala, 2023b,a). Dahlke et al. (2023) also corrected unrealistic heat and moisture peaks in the lowest 100 m, caused by the ship and its exhaust. The extended radiosonde profile better represents the boundary layer structure than the level 2 data used in previous chapters. The need for a good boundary layer representation to better capture surface-based humidity inversions motivated the use of this data set.

Microwave radiometer synergy

We use the IWV and specific humidity profiles retrieved from the synergy of HATPRO and MiRAC-P (Chapter 5, Walbröl et al., 2024). The high quality of the retrieved IWV, which was demonstrated in the same chapter, and its high temporal resolution of approximately 1 s are beneficial for the evaluation of other data sets compared to the 6-hourly resolution of the radiosondes. With the higher temporal resolution, more temporal IWV variability (e.g., due to storms) is resolved, and more temporal overlaps with satellite data can be identified. We

only use data where quality flags indicate good quality, as discussed in Sect. 5.3.2.

Polly^{XT} Raman lidar

During MOSAiC, the multiwavelength Raman lidar Polly^{XT} (Portable Lidar system, next generation) from the Leibniz Institute for Tropospheric Research (TROPOS) was operated on board RV *Polarstern* in the OCEANET container to measure optical properties of clouds and aerosols, and for water vapour profiling at the bow of the ship, where also HATPRO and MiRAC-P were installed (Engelmann et al., 2016, 2021). The lidar’s laser emits linear polarized radiation at 355, 532 and 1064 nm with a pulse repetition rate of 20 Hz at an off-zenith angle of 5°. When electromagnetic radiation with an initial frequency ν_0 interacts with e.g., a water vapour molecule, the rotational and vibrational energy state of the molecule is changed to a lower or higher level, resulting in a frequency increase or decrease of the scattered radiation, respectively (Raman frequency shift, Chapter 25.1 and 25.3 of Foken, 2021). Two receivers, one near-range (for altitudes between 120 m and a few kilometres) and one far-range (800 m to 40 km), are installed to extend the range of the derived profiles (Engelmann et al., 2021). Polly^{XT} receives the radiation backscattered from nitrogen at 387 nm and water vapour at 407 nm. The backscattered radiation can be used to derive water vapour mixing ratio profiles (Dai et al., 2018). The inelastic Raman backscattering is weak compared to the elastic backscattering e.g., from cloud droplets (no frequency shift of the radiation Diaz et al., 2017). Therefore, measurements based on Raman backscattering have a lower signal-to-noise ratio and are only available below cloud base and in the dark season during the MOSAiC expedition (28 September 2019–29 February 2020). The lowest height above the instrument is limited to 30 m because of the overlap of the laser beam and the receivers. The range or height resolution of the lidar water vapour profile is 7.5 m but has been averaged over 5 height bins in the data set used in this study (Engelmann et al., 2023). The data has also been temporally smoothed over 10 minutes.

Cloudnet

For the clear sky detection, we use the same Cloudnet products (target classification, low-level stratus mask, Cloudnet issue data set, Engelmann et al., 2023; Griesche et al., 2023; Griesche and Seifert, 2023) that were described in Sect. 5.3.2. For this study, Dr. Hannes Griesche provided us with an experimental data set of the cloud microphysical properties (liquid and ice water content (LWC and IWC), and effective radii of liquid droplets and ice crystals) via personal communication. This data set extends the Cloudnet retrieval by the microphysical properties of

low-level stratus not detected in the original retrieval (Illingworth et al., 2007; Griesche et al., 2024b). In Cloudnet, the presence of liquid water inside the cloud is detected using the attenuated backscatter coefficient of the lidar. The LWC and droplet effective radii are derived from cloud radar reflectivities, which can miss the low-level stratus when located below the lowest cloud radar height bin (Griesche et al., 2024a). Cloudnet does not identify a liquid phase if the lidar signal is completely attenuated below the lowest cloud radar height bin (integrated LWC would be 0). With the near-range receiver of Polly^{XT}, the presence of low-level stratus can be determined, and LWC is scaled to the HATPRO liquid water path (LWP). Griesche et al. (2024b) found a linear relationship to estimate the liquid droplet effective radius from LWP for the low-level stratus case.

Thermal infrared radiative flux measurements

Several stations measuring solar and thermal infrared (shortwave and longwave) radiation, but also surface conductive heat flux and turbulent heat fluxes (indirectly), were installed on the sea ice during MOSAiC. Their goal was to study the surface energy budget for the full life cycle of an ice floe and to capture the kilometre-scale spatial variability at the surface around RV *Polarstern* (Cox et al., 2023b). While the Met City site provides detailed radiation observations in the 'Central Observatory' close to RV *Polarstern* (Shupe et al., 2022), three autonomous Atmospheric Surface Flux Stations (ASFS-30, ASFS-40, ASFS-50) were deployed at a distance of 12 km in a triangular pattern around the ship. In this study, we mainly use the DLR measurements performed with Eppley Precision Infrared Radiometer pyrgeometers at Met City, which measure the radiative flux of the hemisphere in the spectral range of 4.5–40 μm every second with an uncertainty of 2.6 W m^{-2} (Cox et al., 2023b).

However, the Met City radiation observations were temporarily unavailable due to power outages (e.g., caused by sea ice dynamics), maintenance or other logistical reasons. Sea ice dynamics, the harsh Arctic weather and wildlife attacks caused breakdowns of the ASFS stations, which were then brought back to RV *Polarstern* for maintenance (Cox et al., 2023b). As some ASFS stations were not returned to their original measurement sites but stayed within the Central Observatory, their data could be used to close most of the Met City data gaps. For spring and summer 2020, we use data from ASFS-30 and ASFS-50 when Met City data was unavailable and the respective ASFS station was within the Central Observatory.

The DLR measurements of the ASFS were performed with Hukseflux IR20 pyrgeometers, which have a similar uncertainty but a higher measurement rate and slightly different spectral range (4–50 μm). The pyrgeometers were internally

or externally ventilated and heated to reduce biases originating from the heating of the sensor and icing (Cox et al., 2021). In the data set used in this study, radiative flux observations have been averaged over 1 min (Cox et al., 2023a). From the downwelling and upwelling longwave radiation observations, Cox et al. (2023b) also computed the surface skin temperature following the procedure of Persson et al. (2002).

6.2.2 Numerical models

Reanalyses

Section 5.3.1 already provided a brief overview of the European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis v5 (ERA5). In addition, it is essential to mention that MOSAiC radiosondes have been assimilated with a 4D-Var assimilation scheme (Hersbach et al., 2020). In this study, we used the hourly temperature, pressure, specific humidity and ozone profiles on the 137 vertical model levels (up to a height of 0.01 hPa, approximately 80 km) as well as IWV for the entire MOSAiC period.

The global Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) is produced by NASA’s Global Modeling and Assimilation Office (GMAO) with a horizontal resolution of $0.5^\circ \times 0.625^\circ$ (latitude, longitude) and 72 vertical levels up to a height of 0.01 hPa (Gelaro et al., 2017). MERRA-2 uses a 3D-Var assimilation scheme and therefore does not respect the exact time of a measurement when collected for the analysis. Also, MERRA-2 assimilates only clear sky satellite radiances while ERA5 includes all-sky scenes (Gelaro et al., 2017; Hersbach et al., 2020). The temporal resolution of the data is 1-hourly for single level data (such as IWV) and 3-hourly for model level data (temperature, specific humidity profiles).

Coupled Arctic Forecast System

The National Oceanic and Atmospheric Administration Physical Sciences Laboratory (NOAA-PSL) developed the Coupled Arctic Forecast System (CAFS), which is a fully coupled atmosphere–ocean–ice–land regional forecast model for the Arctic and based on the climate model ‘Regional Arctic System Model’ (RASM) (Solomon et al., 2023, 2024). Initial and boundary conditions are provided by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS), as well as satellite data for sea ice concentration and sea surface temperatures. A direct assimilation of atmospheric observations is not implemented in CAFS. CAFS is initialized at 00 UTC daily and produces 6-hourly output for forecast lead times of up to 10 days. The horizontal resolution is 10 km, and 40

vertical levels are distributed between 12 m and about 19 km. In this study, we used CAFS output published on the data repository PANGAEA at 6, 12, 18 and 24-hour forecast lead times to achieve a 6-hourly resolution (Solomon and Jozef, 2024).

Global ICOSahedral Non-hydrostatic model

The ICON model developed by the German Weather Service (Deutscher Wetterdienst, DWD) and the Max Planck Institute for Meteorology (MPI-M) is a state-of-the-art model for numerical weather prediction and climate modelling (Zängl et al., 2015). The global weather forecast configuration uses a 3D-Var assimilation scheme and provides a 7.5 day forecast and 1-hourly output on an icosahedral-triangular grid with a horizontal resolution of 13 km and 90 vertical levels between 10 m and 75 km (Zängl et al., 2022; Reinert et al., 2024). By using 6–17 hour forecast lead times of the model runs initialized at 00 and 12 UTC each day, we obtain an hourly output while allowing for a 6-hour model spin-up time. The data were downloaded from the [Arctic Data Centre](#) (Frank, 2023).

6.2.3 Satellite products

IASI combined sounding product

IASI measures radiances in the thermal infrared between 3.7 and 15.5 μm with a high spectral resolution and a spatial resolution of 12 km at nadir on board the polar orbiting (sun-synchronous) Metop-A, -B and -C satellites (Blumstein et al., 2004). The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Central Facility derives various surface parameters and profiles, including IWV, temperature and humidity profiles in the combined sounding product IASI Level 2 Product Processing Facility (PPF) v6 (PPF v5 is presented in August et al., 2012). For the combined sounding product, the infrared radiances are combined with microwave observations at frequencies between 23 and 190 GHz from the Advanced Microwave Sounding Unit (AMSU) and the Microwave Humidity Sounder (MHS), which are also on board the Metop satellites. Temperature and humidity profiles are stored on 101 pressure levels between 1100 and 0.005 hPa according to the product documentation (EUMETSAT, 2017).

Initially, a first guess statistical retrieval (piecewise linear regression) is performed using the microwave and infrared observations from IASI, AMSU and MHS. The cloud conditions are assessed with observations from IASI and the Advanced Very High Resolution Radiometer (AVHRR), and numerical weather

prediction. In clear sky scenes, a physical (optimal estimation) retrieval is attempted, incorporating the IASI radiances to improve the accuracy and vertical resolution of the derived temperature and humidity profiles. The physical retrieval is initiated with the first guess of the statistical retrieval. If the physical retrieval does not converge to a solution or fails other quality criteria (e.g., valid range), the first guess is used. The data has been obtained from EUMETSAT (2024).

Integrated water vapour from AMSR2

Rückert et al. (2023a) developed a physical (optimal estimation) retrieval of atmospheric (IWV and LWP), sea ice (sea ice concentration, multi-year ice fraction, snow depth, snow-ice interface temperature and snow-air interface temperature), and open ocean parameters (sea surface temperature and wind speed) from microwave radiances (6.9–89 GHz) measured by the Advanced Scanning Microwave Radiometer 2 (AMSR2). The AMSR2 instrument is on board the Japan Aerospace Exploration Agency (JAXA) Global Change Observation Mission 1st-Water (GCOM-W1) satellite in a sun-synchronous sub-recurrent orbit and measures with a horizontal resolution between 62×35 km at 6.9 GHz and 5×3 km at 89 GHz. Combining surface and atmospheric parameters allows the retrieval to adapt the parameters more freely to obtain a physically consistent solution. Compared to fixed surface properties, the retrieval can thus better distinguish the signal contribution from the surface and the atmosphere. Currently, their retrieval is only designed for Arctic winter conditions, partly due to high uncertainties in the surface parameters during the melt season and because melt ponds on the sea ice could not be distinguished from open ocean (Rückert et al., 2023a). The retrieved IWV colocated with RV *Polarstern* has been provided by Janna E. Rückert, the author of the retrieval, via personal communication.

6.3 Methods

6.3.1 Data processing

For the humidity comparisons and radiative transfer simulations, temporal and spatial overlaps of the tested data sets with the MOSAiC observations are needed. The temporal overlap criteria depend on the data set and application and will be briefly described in the following. ERA5 and MERRA-2 were selected at the nearest grid point to RV *Polarstern* following the recent example of Solomon et al. (2023). The output of CAFS was produced for the first floe of the MOSAiC expedition and therefore did not always follow RV *Polarstern*, e.g., during the

logistical interruption from mid-May to mid-June 2020 or when the first floe had to be abandoned as it dissolved at the end of July 2020. The ICON output of the data set used in this study was generally colocated for RV *Polarstern*, but the distance could be large during transit. ICON and CAFS data were therefore limited to cases when the geodesic distance to RV *Polarstern* was less than 50 km. When temporal overlaps with the radiosonde data were asked (Sect. 6.4.2, 6.5 and 6.6.3), the nearest time step after each radiosonde launch was used if the time step was within a certain range (60 minutes for ERA5 and ICON, 90 minutes for CAFS and MERRA-2).

IASI data has been averaged over all pixels within a 50 km radius of RV *Polarstern* and over 60 minutes around each radiosonde launch for the temporal overlaps. We also computed height levels for the retrieved profiles, stored on pressure levels, using the radiosonde data. Specific humidity profiles derived with the optimal estimation procedure have been included where they were available (see description in Sect. 6.2.3). However, the first guess profiles had to be used more than two-thirds of the time. The creator of the AMSR2 IWV product performed the same spatial overlap we did for IASI.

Consistent with the preceding chapters, we confine the MWR data to times when quality flags indicate good quality, applied a 5-minute running mean to the IWV, and averaged the specific humidity profiles over 15 minutes (starting at the time of each radiosonde launch) for comparisons with radiosondes. For the humidity inversion characteristics, specific humidity profiles with a 5-minute running mean are taken every 15 minutes (compromise between resolution and data amount).

The lidar data were averaged over only 5 minutes, from the time of each radiosonde, to avoid including more artefacts and outliers near the surface and above 2000 m that escaped the quality control of Engelmann et al. (2023). Further lidar data processing would be required to distinguish outliers and artefacts from actual signals. The processing would include the identification of RV *Polarstern*'s exhaust using additional wind data and filtering unrealistically strong vertical moisture gradients. As the lidar data was mainly used to cross-validate the radiosonde data, the additional processing was not performed at this point but may be considered in the future.

Furthermore, the lidar data were limited to heights between 30 and 4000 m because of the low data availability above 4000 m, caused by the low signal-to-noise ratio above this height level. We considered lidar data only at times when no cloud was detected below 4000 m with the Cloudnet target classification and the low-level stratus mask (as described in Sect. 5.3.2). For the humidity inversion analysis, the radiosonde and lidar profiles were interpolated to a height

grid with 50 m spacing using weighted averages to keep the vertical water vapour distribution close to the original profile (see Appendix 6.A). The relatively coarse height grid filters out most of the smaller fluctuations in the humidity data to enhance the focus on the main vertical features (here, inversions).

6.3.2 Humidity inversion detection

Specific humidity q inversions can be detected with different methods and different minimum strength and depth thresholds (e.g., Devasthale et al., 2011; Vihma et al., 2011; Nygård et al., 2013; Brunke et al., 2015; Maturilli and Kayser, 2017a; Naakka et al., 2018). The basic principle is to find heights where the specific humidity increases with height. While this method is straightforward for 'smooth' profiles (e.g., numerical models, microwave and infrared retrievals), more work is needed to distinguish small humidity fluctuations of a radiosonde profile from inversions.

This study has revised the humidity inversion detection algorithm of Chellini and Ebell (2022). Figure 6.1 illustrates the detection algorithm. The iteration through the height grid starts at the surface and ends at 7000 m. Initially, the lowest height index where q increases with height is determined (inversion base height z_{base} and specific humidity at the base q_{base}). Then, the next height level above the base where q falls below q_{base} again marks the 'extended inversion top' ($z_{\text{top,ext}}$, see Fig. 6.1). The formally defined inversion top (z_{top} and q_{top}), where q starts decreasing with height again, is the maximum q between the base z_{base} and the extended top $z_{\text{top,ext}}$. The search for the next inversion starts at the extended inversion top of the current inversion. If the next inversion was based above 7000 m, the detection would be aborted. Once all inversions have been determined, the depth ($z_{\text{top}} - z_{\text{base}}$), extended depth ($z_{\text{top,ext}} - z_{\text{base}}$), strength ($q_{\text{top}} - q_{\text{base}}$) and relative strength ($(q_{\text{top}} - q_{\text{base}}) / q_{\text{top}}$) of each inversion is calculated. As in Nygård et al. (2013), an inversion is considered surface-based if the base height is equal to or below 50 m.

Small-scale specific humidity fluctuations are filtered out by minimum strength, relative strength, depth and extended depth thresholds. For humidity inversion detection with radiosondes, the minimum inversion strength threshold depends on the specific humidity error determined via error propagation of the measured relative humidity, air temperature and air pressure. If the maximum specific humidity error of a profile is less than 0.5 g kg^{-1} (always fulfilled during MOSAiC), the maximum humidity error is used as a threshold for the minimum inversion strength. Otherwise, the minimum strength threshold is 0.5 g kg^{-1} for radiosondes and 0.05 g kg^{-1} for the remote sensing products, reanalyses and forecast models,

whose profiles are usually much smoother.

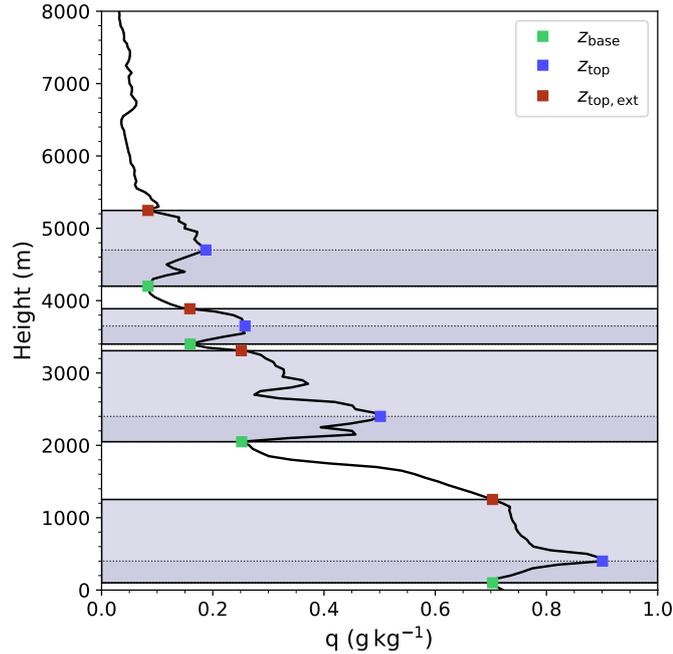


Figure 6.1: Demonstration of the humidity inversion detection scheme for a specific humidity profile measured by the radiosonde launched on 01 November 2019 at 10:56 UTC. Green, blue and red squares indicate the inversion base z_{base} , top z_{top} and extended top $z_{\text{top,ext}}$, respectively.

The minimum relative strength, depth and extended depth thresholds were chosen to be 5%, 100 m and 200 m, respectively. The depth thresholds filter specific humidity fluctuations while enabling more precise base height and top estimates because we use a finer height grid of 50 m spacing. The relative strength threshold is similar to the absolute threshold used by Vihma et al. (2011) in winter conditions. For further filtering, inversions are separated into two types: Firstly, if the inversion depth is < 100 m, the (relative) inversion strength must be equal to or exceed twice the minimum (relative) strength threshold (allowing very shallow but strong inversions). Secondly, if the inversion depth is ≥ 100 m, the (relative) inversion strength must be equal to or exceed the minimum (relative) strength threshold (allowing deeper but weaker inversions).

This algorithm focuses more on the main humidity inversions than other algorithms. Here, smaller inversions nested within a larger one are not identified as single inversions. For example, other algorithms may consider the humidity increase at approximately 2600–3000 m (between the top and extended top of the second lowest inversion, Fig. 6.1) as a single inversion while it is nested within the main inversion in this algorithm. Thus, inversion number and depth statistics will differ from previous studies. This algorithm tends to determine smaller inversion numbers but greater depths. This algorithm also respects the humidity contained

between the inversion top and the extended top, allowing us to consider the total amount of water vapour within an inversion (fraction of the IWV between z_{base} and $z_{\text{top,ext}}$). This water vapour amount is obviously higher than that between the base and the formal inversion top (z_{base} and z_{top}) and will therefore have a greater radiative effect in the thermal infrared.

6.3.3 Contingency table for inversion detectability

Contingency tables are often used to compare the occurrence or absence of an event in a reference data set to another data set that should be evaluated (e.g., occurrence of strong precipitation, Saouabe et al., 2020). Therefore, it is a valuable tool for assessing the ability of the different observational or model data sets to detect humidity inversions with respect to radiosondes. Table 6.1 shows the setup of the contingency tables for this study. Correct positives (cp) refer to inversions that are identified by both the radiosonde and the tested data (e.g., IASI), while false positives (fp) indicate that no inversion is present in the radiosonde data but exists in the tested data. False negatives (fn) correspond to cases where the tested data does not show an inversion while the radiosonde does. Correct negatives (cn) occur when neither the tested data nor the radiosonde detects an inversion. Several statistics, whose meanings will be discussed in the analysis, are computed from the contingency table to assess the performance of the tested data set:

- Accuracy = $(cp + cn) / total$
- Bias = $(cp + fp) / (cp + fn)$
- Probability of false detection POFD = $fp / (cn + fp)$
- Success ratio SR = $cp / (cp + fp)$
- Heidke skill score HSS = $(cp + cn - e_c) / (total - e_c)$
with $e_c = ((cp + fn)(cp + fp) + (cn + fn)(cn + fp)) / total$

6.3.4 Radiative transfer model

In this study, DLR is simulated with the Tropospheric Research (TROPOS) Cloud and Aerosol Radiative effect Simulator (T-CARS) (Barlakas et al., 2020; Barrientos-Velasco et al., 2022)). T-CARS is a Python-based version of the 1-D rapid radiative transfer model for general circulation models (RRTMG) (Mlawer et al., 1997; Iacono et al., 2008) and calculates broadband upwelling and downwelling fluxes in the solar and infrared spectrum (wavelength ranges of 0.2–12 μm

Table 6.1: Contingency table used to assess the detectability of humidity inversions in a test data set compared to the reference data set (here, radiosondes). True and false indicate the occurrence and absence of humidity inversions, respectively.

		Radiosonde		
		True	False	Total
Test data	True	Correct positives	False positives	True in test data
	False	False negatives	Correct negatives	False in test data
		True in radiosondes	False in radiosondes	Total

and 3–1000 μm). Although the focus lies on DLR, parameters affecting the upwelling and solar spectrum fluxes are also included.

The simulations were set up similarly to Barrientos-Velasco et al. (2022) with the help of Dr. Carola Barrientos-Velasco. However, slightly different input data is used: The radiosonde data described in Sect. 6.2.1 provide the profiles of air temperature, pressure and specific humidity. The specific humidity has been converted to volume mixing ratio. Skin temperature estimates are taken from the Central Observatory. The Cloudnet LWC, IWC, and effective radii of liquid droplets and ice crystals were implemented following Barrientos-Velasco et al. (2022). Ozone mass mixing ratio profiles from ERA5 model level data, colocated with RV *Polarstern*, are also converted to volume mixing ratios. Other trace gas volume mixing ratios (carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), chlorofluorocarbon gases (CFC-11, CFC-12) and carbon tetrachloride (CCl_4)) are taken from measurements performed by the NOAA Global Monitoring Laboratory and historical records (Lan et al., 2022b,a; Dutton et al., 2023a,c,d,b). Carbon monoxide and CFC-22 volume mixing ratios are used from Anderson et al. (1986). All profile data has been interpolated onto the Cloudnet height grid (30 m spacing), which has been extended with similar spacing to 20 km.

The solar zenith angle at RV *Polarstern* at a given time is determined using Python’s *skyfield* module (Rhodes, 2019). The surface emissivity is assumed to be 0.9999 throughout the expedition, although Barrientos-Velasco et al. (2022) noted that 0.9907 should be used when the sea ice concentration around RV *Polarstern* is below 50%. The surface albedo (near-infrared, visible and ultraviolet spectrum) was set to 0. These assumptions do not affect our results as the surface albedo, as well as the solar zenith angle and surface emissivity, do not influence the simulated DLR.

An example of the simulated thermal infrared radiative fluxes is given in Fig. 6.2. The simulated radiative fluxes overall agree well with the observations at Met city. Clouds clearly influence the DLR until approximately 06 UTC and after 21 UTC in the observations, and after 22 UTC in the simulations. The sudden peaks during the cloudy periods in the simulated DLR are caused by missing or 0 values of LWC in the Cloudnet data set. Thus, further data processing would be required when this analysis is extended to all sky conditions. The relatively constant clear sky bias ($\approx -5 \text{ W m}^{-2}$) in DLR between 06 and 21 UTC can be partly attributed to the neglect of aerosols in the simulations (Barrientos-Velasco et al., 2022), but may also be due to errors in the near-surface temperature and humidity data. However, as the focus lies on the differences between simulations with original and modified humidity profiles, the clear sky bias cancels out.

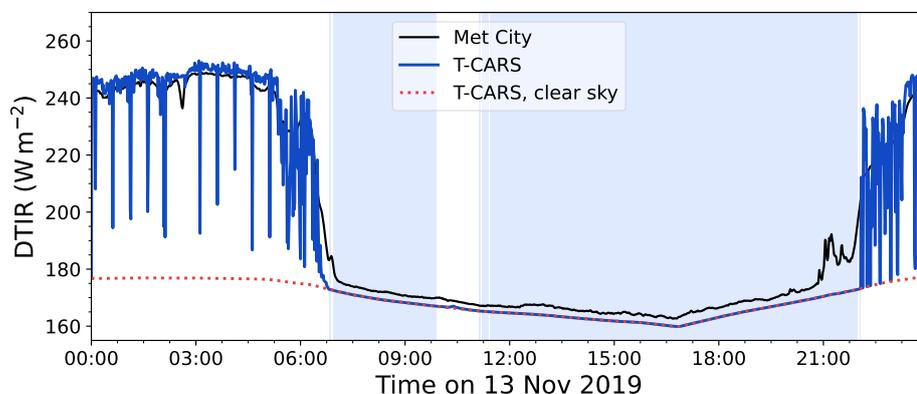


Figure 6.2: Simulated and measured downwelling thermal infrared radiative fluxes (DLR) on 13 November 2019. Measured (simulated) radiative fluxes are shown as blue (black) solid lines, and simulated clear sky radiative fluxes are displayed as red dotted lines. Blue patches indicate clear sky according to Cloudnet.

6.3.5 Specific humidity profile modifications

To analyze the sensitivity of DLR to water vapour within humidity inversions, the strength and base heights of humidity inversions are modified. Figure 6.3 illustrates each of the modifications for an elevated inversion whose original base was located at approximately 1800 m height. In the following, the procedure of inversion modifications between its base and extended top are described:

- **Strength:** The strength factor a determines by what magnitude the strength of the inversion is modified. Between the base and the extended top, the specific humidity profile is modified according to $q'_{\text{inv}} = q_{\text{base}} + a(q_{\text{inv}} - q_{\text{base}})$, and thus depends on the difference between the specific humidity within the inversion (q_{inv}) and at base (q_{base}). The dependence on the difference

ensures a smooth transition at the base and extended top of the inversion while accurately modifying the inversion strength. For example, at the formal inversion top, where q is maximum within the inversion, the difference $q_{\text{inv}} - q_{\text{base}}$ equals the inversion strength, which is therefore directly changed according to the required strength factor. Figure 6.3a presents the inversion modifications for strength factors a between 0 and 1.

- **Base height:** The inversion base height is shifted according to an offset Δz . Initially, the height indices of the original and perturbed bases and the original extended top are determined. The specific humidity profile is shifted using the height index difference between the perturbed and original base. If Δz is positive (base height increased), the specific humidity between the original and perturbed base height is set to the value at the original base q_{base} . At the extended top of the perturbed inversion, a linear transition zone with a height of 250 m is used to merge the perturbed and original profiles. Compared to a fine height grid, the actual perturbed base heights may differ more from the target base heights on a coarser height grid because the height grid is unchanged. For example, $\Delta z = 65$ m will not shift the inversion by 65 m on a 50 m grid. Figure 6.3b shows base height shifts Δz between -500 and $+500$ m.

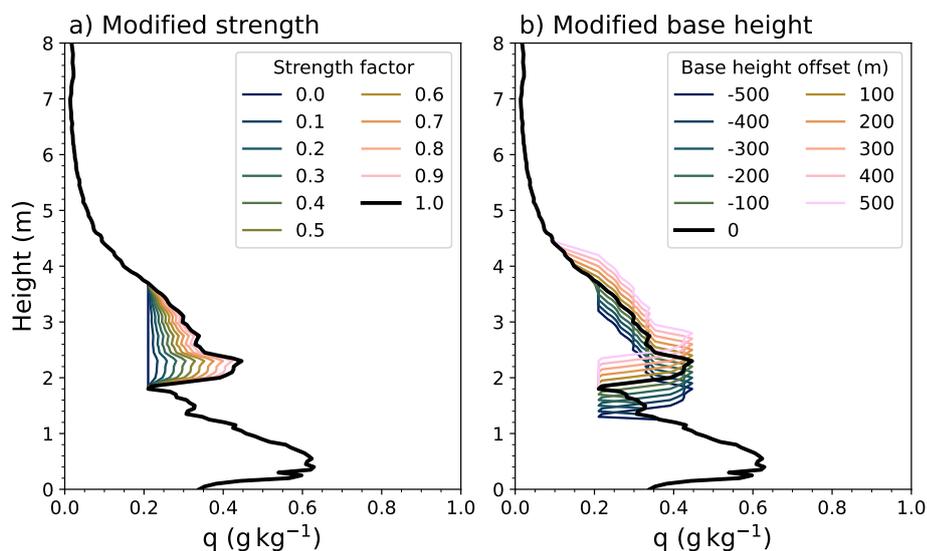


Figure 6.3: Specific humidity profile with modifications of the inversion strength (a) and base height (b). In the example (radiosonde launched on 26 February 2020, at 10:55 UTC), only the strongest inversion with a base height above 1000 m is modified.

6.4 Water vapour comparison

Due to the strong relationship between DLR and water vapour (e.g., Ghatak and Miller, 2013), it is important to first identify the water vapour biases of each data set. The findings will be used to interpret the DLR sensitivity to the different specific humidity profiles of the data sets in Sect. 6.6.3. With the reference measurements from MOSAiC, biases and root mean squared deviations (RMSD) of the IWV and specific humidity profiles were calculated using Eq. 5.1 and 5.2. All data sets have been restricted to times between 22 October 2019 and 05 August 2020 because of the data availability and the spatial overlap constraint of CAFS. Therefore, we ensure a similar time range in the comparison. However, most data sets are available for the entire MOSAiC expedition. Data sets may include different numbers of samples in the comparison due to different temporal resolutions. For example, IASI has fewer temporal overlaps with radiosondes than ERA5 due to orbit limitations.

For the IWV evaluation, the synergistic retrieval of HATPRO and MiRAC-P was used as reference because of the high quality and temporal resolution. The high temporal resolution allows the inclusion of more samples (especially for ERA5, ICON, AMSR2 and IASI) and thus more IWV variability in the comparison. Additionally, the MWRs provide point measurements of IWV, while the radiosondes can drift tens of kilometres during ascent and descent. The Polly^{XT} was excluded from the IWV comparison because of biases due to the limited height range, which was affected by the sensitivity. Specific humidity profiles of the data sets were evaluated against the radiosonde observations.

6.4.1 Integrated water vapour comparison

The MOSAiC expedition featured a large range of IWV conditions, from less than 1 kg m^{-2} to approximately 30 kg m^{-2} as shown in the IWV histogram (Fig. 6.4a). For more than 50 % of the time, IWV was below 5 kg m^{-2} , while only 10 % of the IWV data was above 15 kg m^{-2} . Consequently, fewer data samples were included in the error computation for moist conditions than for dry conditions.

All data sets generally denote higher RMSD with increasing IWV, while the relative RMSD is constant or even decreases (Fig. 6.4b). Throughout the entire IWV range, ERA5 shows the smallest RMSD with $< 0.2 \text{ kg m}^{-2}$ (in relative terms, 15 %) at the lower IWV limit and up to 1.2 kg m^{-2} ($< 5 \%$) at the upper IWV limit. At the dry limit, the RMSD of ERA5 is within the uncertainty range of the MWR synergy. However, the 4D-Var assimilation of the MOSAiC radiosondes is likely responsible for the good performance. MERRA-2 and ICON, who use 3D-var assimilation, partly have twice as high RMSD as ERA5 with

0.3 kg m⁻² or 20 % at the lower IWV limit and up to 2.4 kg m⁻² or 10 % at the upper limit. Despite not being directly influenced by the MOSAiC radiosondes, CAFS shows similarly low RMSD as ICON and MERRA-2. For IWV below 10 kg m⁻², IASI performs similarly well regarding the RMSD as ICON, MERRA-2 and CAFS. However, for larger IWV, the RMSD of IASI is about 5–10 percentage points higher than the models. With more than 20 %, the AMSR2 retrieval shows the highest RMSD among all presented data sets for IWV below 12 kg m⁻², which demonstrates the persisting challenges regarding satellite-based microwave remote sensing. Also, during moist air intrusions in summer, when IWV increased to values above 20 kg m⁻², the AMSR2 retrieval denotes the highest deviations. However, the retrieval was designed for the cold season and was not specifically optimized for IWV. Rapid changes of the sea ice emissivity during and after such intrusions can explain increased uncertainties during melting conditions (Rückert et al., 2023b).

The biases of the models are mostly smaller than ± 0.5 kg m⁻², except for CAFS, where it is up to ± 0.8 kg m⁻² (Fig. 6.4c). For IASI, a distinct negative bias was identified, with a magnitude of -3 kg m⁻² (in relative terms, -15 %) at an IWV of 20 kg m⁻². When considering only IASI’s physical retrieval (with good quality flags), the RMSD and magnitude of the negative biases of IASI is up to 5 percentage points smaller for IWV between 5 and 20 kg m⁻² (not shown). However, the number of samples included in the comparison with the MWR data is reduced by two-thirds. The AMSR2 IWV retrieval also underestimates IWV by 5–15 % in scenes where IWV exceeds 3 kg m⁻². Vaquero-Martínez et al. (2020) also found that all of the satellite products they included in their comparison, none of which are included here, underestimate the IWV in humid conditions. The strong positive bias of up to 1 kg m⁻² for IWV below 3 kg m⁻² and a time series analysis (not shown) revealed that 3 kg m⁻² seems to be the sensitivity limit of the AMSR2 IWV retrieval. Around 15 kg m⁻², the AMSR2 retrieval performs best and has smaller biases than IASI.

When applying 7-day running means to all data sets to smooth out higher frequency IWV variability (e.g., due to storms), all models show much smaller RMSD (< 5 %) but similar biases for almost the entire IWV range (Appendix 6.B, Fig. 6.18). With the 7-day running mean, potential errors, which can be caused by missing the exact timing of moisture intrusions (related to storms, Viceto et al., 2022), are removed. Thus, a significant fraction of the RMSD of the models is caused by synoptic-scale variability of IWV. For example, the timing and strength of IWV changes during storms are often not well represented (Crewell et al., 2021). Despite using the assimilated radiosondes and filtering IWV variability, MERRA-2, ERA5 and ICON still have one of the strongest negative biases of -5 to -10 %

for IWV below 5 kg m^{-2} . IASI also has a similarly strong negative bias in humid conditions as in the unfiltered IWV comparison. More details regarding the IWV evaluation with a 7-day running mean can be found in Appendix 6.B.

The most distinct deviations, which also appeared in the 7-day running mean, were found for the satellite data in humid conditions. The deviations are likely related to clouds and the rapidly changing sea ice characteristics during the melt period. The fraction of the data with IWV above 15 kg m^{-2} may seem low (10%), but as we included almost the entire MOSAiC expedition, 10% of data represents about 30 days. Additionally, all models except CAFS showed slightly negative biases when IWV was below 10 kg m^{-2} . It is important to stress that ERA5, MERRA-2 and ICON are likely strongly influenced by the MOSAiC radiosondes. A comparison without the assimilation of MOSAiC radiosondes would be of high interest but is not possible at the time of writing this study.

6.4.2 Specific humidity profile comparison

Specific humidity profiles were evaluated for each season with the bias, RMSD and relative RMSD for each height level, as shown in Fig. 6.5. Higher absolute deviations were generally found in the warmer seasons because of higher specific humidity values. Due to the exclusion of data after 05 August 2020 (see above), summer only spans from 01 June to 05 August 2020, and autumn ranges from 22 October to 30 November 2019. More than 1050 radiosondes were incorporated in the comparison and interpolated onto the other data sets' height grids. The vertical spacing of the height grids is illustrated in Fig. 6.19 in Appendix 6.C. Lidar profiles were included for clear sky scenes up to 4000 m height until 29 February 2020 for cross-validation of the radiosonde humidity measurements.

The Polly^{XT} agrees very well with the radiosonde observations in the lowest 2 km as the RMSD is less than 10% (Fig. 6.5c, f) and therefore within the uncertainty range of Polly^{XT} (Dai et al., 2018). Above 2 km, radiosonde drifts and lidar measurement outliers (artefacts), which were not caught in the quality control, contribute to the increasing errors. As the strength of the Raman lidar signal depends on the amount of water vapour and the distance to the instrument (mainly due to extinction), higher altitudes can only be sensed in more humid conditions. Consequently, the comparison between radiosondes and lidar observations above 3 km is limited to the more humid cloud-free conditions, while lower altitudes can also be sensed in dry conditions. Together with the outliers and the radiosonde drift, this may explain the considerable increase in (relative) RMSD with height in winter. Near the surface, the plume of the ship's exhaust may be responsible for the sharp RMSD and bias peak. Dahlke et al. (2023) corrected this poten-

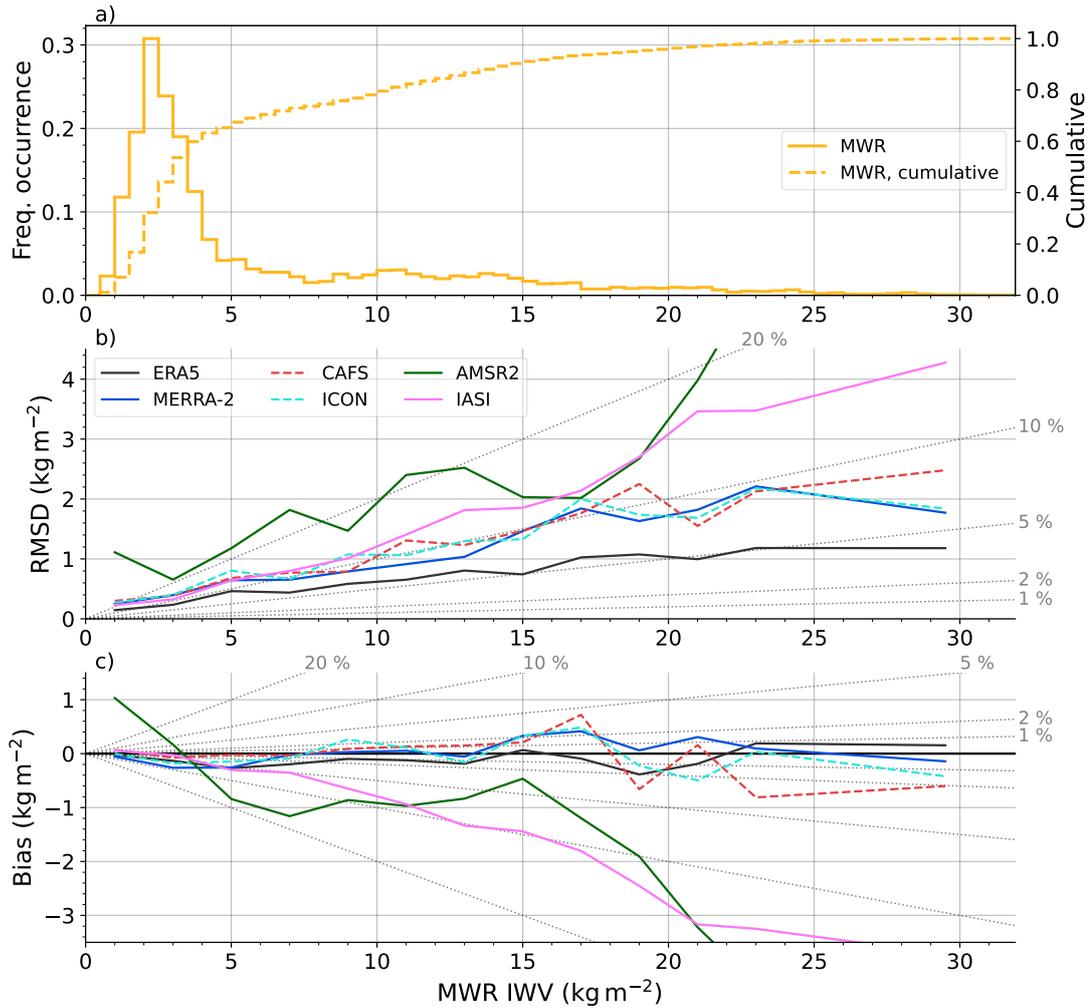


Figure 6.4: IWV histogram of the MWR data showing the frequency of occurrence (solid) and cumulative occurrence (dashed) (a), and IWV root mean squared deviation (RMSD, b) and biases (c) of different data sets with respect to the MWR data. ERA5, MERRA-2, IASI and AMSR2 errors are shown as solid black, dark blue, pink and green lines. CAFS and ICON are shown as dashed red and light blue lines. Relative error lines are included as grey dotted lines in the background of the RMSD and bias panels.

tial source of error in the radiosonde data set. However, this error has not been corrected in the lidar data.

In all seasons, the synergy of HATPRO and MiRAC-P shows similar or even lower RMSD than MERRA-2, ICON and CAFS below 4 km with relative RMSD of 10–30 % (Fig. 6.5c, f, i, l). These low deviations demonstrate that the specific humidity profiles from the MWR are, on average, at least as good as the models considered, despite the assimilation of the radiosonde data in ICON and MERRA-2. Relative errors generally increased with height due to the decreasing specific humidity. However, it is likely that radiosonde drifts also contributed to the increase in relative deviations with height because all other data sets represent

point measurements or spatial averages. The MWR profiles also show biases with magnitudes smaller than most other data sets in the lowest 1 km in all seasons except summer. In winter, the MWR synergy has the smallest biases among all data sets (except the lidar). Above 1 km, the MWR synergy overestimates the specific humidity more than the models in autumn, spring and summer (Fig. 6.5a, d, g, j). These enhanced biases are related to the representation of humidity inversions, as already discussed in Sect. 5.5.2 and shown later.

All models, including ERA5, but to a lesser extent, slightly underestimate the specific humidity between a few hundred meters above the surface and 2 km in all seasons except summer. For ERA5 and MERRA-2, this is in accordance with the findings of Graham et al. (2019a) in cold conditions (their Fig. 4). Because a high fraction of the IWV is located at these heights, this result is consistent with the negative IWV bias that we have identified in dry conditions. In spring and summer, CAFS overestimates the near-surface specific humidity by up to 0.15 g kg^{-1} . MERRA-2 has a considerable positive specific humidity bias of up to 0.1 g kg^{-1} at the surface in spring, and above 1.5 km in summer. The latter findings agree with the results of Graham et al. (2019b). ERA5 also overestimates the near-surface specific humidity in summer.

The MWR synergy has lower relative RMSD below 3 km (10–25 %) than IASI (15–30 %) in all seasons. Above 3 km, IASI has smaller relative RMSD than the MWR synergy and similar RMSD as MERRA-2 and ICON (30–45 %). In the cold seasons, IASI overestimates the specific humidity in the lowest 1 km but underestimates it at higher altitudes. In summer, IASI is strongly negatively biased in the lowest 5 km (up to -0.35 g kg^{-1}), except for a $\approx 250 \text{ m}$ deep layer of strong positive bias at the surface (up to $+0.5 \text{ g kg}^{-1}$). This underestimation is consistent with the negative IWV bias seen for humid conditions in Fig. 6.4c. When confining the IASI data to the physical retrieval, which reduces the number of samples in the comparison from 645 to 140, the specific humidity bias is much smaller in summer (not shown).

Figure 6.6 provides an example of a specific humidity profile in winter to illustrate the origin of specific humidity biases. The strong surface-based humidity inversion seen by the radiosonde (0–2500 m height according to the detection algorithm) is generally captured by all data sets, but its strength is underestimated. Due to the inversion strength underestimation, CAFS, MERRA-2 and ICON underestimate the specific humidity in the lowest 2000 m. MERRA-2 and ERA5 slightly overestimate the specific humidity in the lowest 100 m, which can be found for many profiles in winter, resulting in the positive bias seen in Fig. 6.5d. IASI strongly overestimates q in the lowest 1000 m, while the MWR denotes one of the highest deviations above 4500 m. The lidar profile shows higher specific

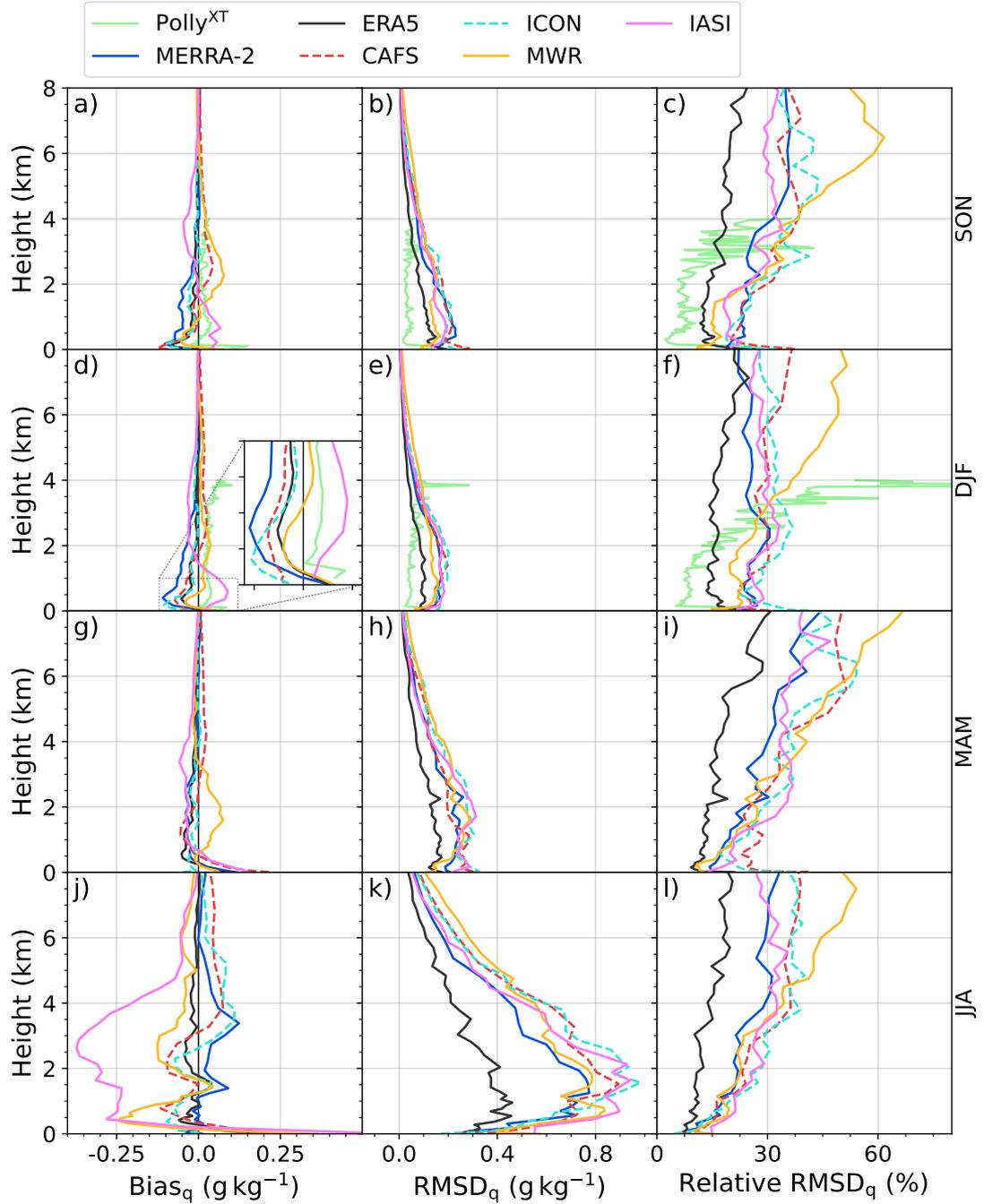


Figure 6.5: Specific humidity biases (left column), root mean squared deviation (RMSD, middle column) and relative RMSD (right column) of different data sets compared to radiosonde measurements for autumn (SON, a–c), winter (DJF, d–f), spring (MAM, g–i) and summer (JJA, j–l). The Polly^{XT} lidar, ERA5, MERRA-2, IASI and MWR data are visualized as solid green, black, dark blue, pink and yellow lines, while ICON and CAFS are indicated as dashed light blue and red lines. The inset axis in panel (d) zooms into the biases of the lowest 1 km.

humidity than the radiosondes at the top of the lidar profile (3500–4000 m), which is an artefact rather than an atmospheric signal.

The specific humidity evaluation shows that ERA5 generally has the smallest

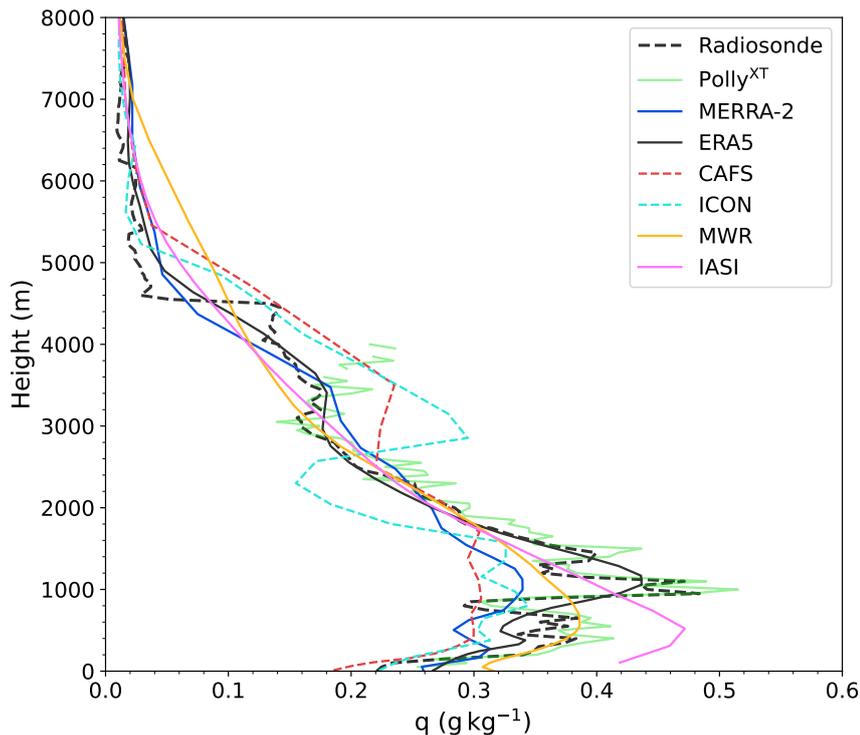


Figure 6.6: Specific humidity profiles of all data sets, temporally colocated with the radiosonde launched on 01 January 2020 at 05:44 UTC. Colours are similar to Fig. 6.5 with the addition of the radiosonde profile, which is illustrated with a black dashed line. The radiosonde profile was interpolated to a vertical grid with 50 m resolution for the visualization.

RMSD. All models are negatively biased in the lower troposphere during the cold seasons, except near the surface. The synergy of HATPRO and MiRAC-P denotes similarly low RMSD as MERRA-2, ICON and CAFS in the lowest 4 km in all seasons despite the limited vertical resolution.

6.5 Humidity inversions

6.5.1 Inversion detectability

Before analyzing the humidity inversion characteristics of the data sets, it is of interest to test the ability to detect inversions compared to radiosondes using the contingency table and statistics described in Sect. 6.3.3. Due to the temporal overlap constraints (similar to the Sect. 6.4.2), the number of cases for each tested data set is often lower than the total number of radiosondes considered in this comparison (1096). In total, 98 % (in numbers: 1076) of the 1096 radiosonde launches contain at least one humidity inversion. Similarly high fractions of cases with at least one inversion (N_{inv}/N , see Table 6.2) are achieved by the models with 91 % for ERA5 and MERRA-2, 93 % for CAFS and 95 % for ICON.

Table 6.2 summarizes the contingency table statistics for the different data sets. The highest accuracy was found for the models (> 0.92), followed by the lidar (0.82) and the MWR synergy (0.66), while IASI has the lowest accuracy (0.53). Therefore, the MWR synergy identifies the presence and absence of inversions better than IASI but worse than the models. If the tested data set (e.g., ICON) detects more (less) inversions than the radiosondes, the bias values are > 1 (< 1). Bias values close to 1 indicate that the number of inversions in the tested data is similar to that in the radiosonde data. Surprisingly, CAFS, ICON and MERRA-2 have similar or even better accuracy and bias values than ERA5 despite having coarser vertical resolutions. As for the IWV and specific humidity profile evaluation, CAFS performs well compared to the other models despite the disadvantage regarding the assimilation of MOSAiC radiosondes. Thus, it is unclear to what extent the assimilation of the radiosondes contributes to the high accuracy and good bias values of ERA5, MERRA-2 and ICON.

The MWR shows the lowest and therefore best probability of false detection (0.0) among all data sets, followed by IASI (0.14) and ERA5 (0.15). These data sets identify the absence of inversions well. However, IASI and the MWR also miss numerous inversions (hence the low bias values). Consequently, it is likely that the low probability of false detection is due to random chance rather than skill in detecting the absence of inversions. As each of the 114 cases included in the comparison with Polly^{XT} features at least one inversion, the probability of false detection could not be computed (therefore set to not a number, *nan*). CAFS and ICON perform worst regarding the probability of false detection as they predict an inversion in 50 % of the cases when no inversion is observed in the radiosonde profile. All data sets show high success ratios (≥ 0.99) due to the small number of false positive cases (the tested data set detected an inversion when the radiosonde did not). For example, when the MWR identified an inversion, an inversion was also present in the radiosonde profile in 100 % of the cases.

The Heidke skill score, which expresses the accuracy of the detection over random chance of correct detection, is generally low for all data sets because of the relatively high number of false negatives (misses) compared to correct negatives. Typically, the Heidke skill score is low when the number of false predictions (misses and false positives) is high compared to the number of true positives (hits) or true negatives. For Polly^{XT}, the fraction of misses to the hits dominates the skill score because no false positives or correct negatives occurred. Thus, due to the relatively high fraction of misses to the hits (21 to 93) the Heidke skill score is the lowest (0.0). Although MWR and IASI have a higher fraction of misses to all cases, their skill score is slightly higher (0.02 and 0.05) because they were also able to correctly detect the absence of inversions. ERA5 has the

Table 6.2: Statistics computed from the contingency table (Table 6.1) to assess the detectability of humidity inversions in a tested data set compared to radiosondes. N , N_{inv} , POFD, SR and HSS refer to the number of cases, number of cases where the tested data set detected inversions, probability of false detection, success ratio and the Heidke skill score.

	N	N_{inv}	Accuracy	Bias	POFD	SR	HSS
MWR	1064	682	0.66	0.65	0.00	1.00	0.05
Polly ^{XT}	114	93	0.82	0.82	<i>nan</i>	1.00	0.00
IASI	645	336	0.53	0.53	0.14	1.00	0.02
ERA5	1096	1000	0.93	0.93	0.15	1.00	0.27
MERRA-2	1096	1001	0.92	0.93	0.25	1.00	0.24
CAFS	991	917	0.93	0.93	0.50	0.99	0.10
ICON	1075	1018	0.95	0.96	0.50	0.99	0.20

highest Heidke skill score with 0.27 as it has the lowest fraction of misses and false positives to the true positives and negatives.

The presence of inversions is mostly missed when the relative strength and extended depth are low ($< 40\%$ and $< 1500\text{ m}$), and when no surface inversion occurred (not shown). However, all data sets also missed some strong, deep and surface-based inversions. The contingency table statistics yield similar results when incorporating the entire MOSAiC year where possible (i.e., MWR, ERA5, MERRA-2, IASI, ICON). A comparison between the MWR synergy and the HATPRO-only retrieval (Appendix 6.D) demonstrates that the profiling capabilities concerning humidity inversions were significantly improved through the synergy.

6.5.2 Inversion characteristics

In this part, humidity inversion characteristics (number of inversions in a profile, base height, relative strength, extended depth) of the models, MWR and IASI were compared to radiosondes to identify biases in the representation of humidity inversions. Here, all data sets were analyzed using their native height grids (except the radiosondes and lidar, which were interpolated to a grid with 50 m spacing) and were colocated with radiosondes as in the previous analyses. The Polly^{XT} data was included for completeness but biases may exist with respect to the other data sets due to the low data coverage. However, the lidar data can still be used to verify the radiosonde statistics.

Figure 6.7 shows the statistics of the number of inversions per profile, relative strength, base height and extended depth for each data set and each inversion. The radiosondes and the lidar detect the highest number of inversions, mostly 1–4 and 0–3, respectively, which is expected due to the higher vertical resolution

(Fig. 6.7a). The limited height range and data coverage of Polly^{XT} may explain the lower number of inversions compared to radiosondes because this bias also occurs when limiting the radiosondes to the times when the lidar was available (not shown). All models similarly underestimate the occurrence of cases with 2 or more inversions and show almost no occurrences of 4 or more inversions per profile. IASI and MWR rarely observed more than 1 inversion (Fig. 6.7a). They also seem unable to resolve elevated inversions with bases above 600 m (Fig. 6.7b). Additionally, IASI misses more inversions than any other data set, showing the highest fraction of 0 inversions (consistent with the contingency table statistics shown before). All data sets agree that most inversion bases are near or at the surface (Fig. 6.7b). However, the fraction of elevated inversions with base heights between 1500 and 7000 m is underestimated by all data sets compared to the radiosondes.

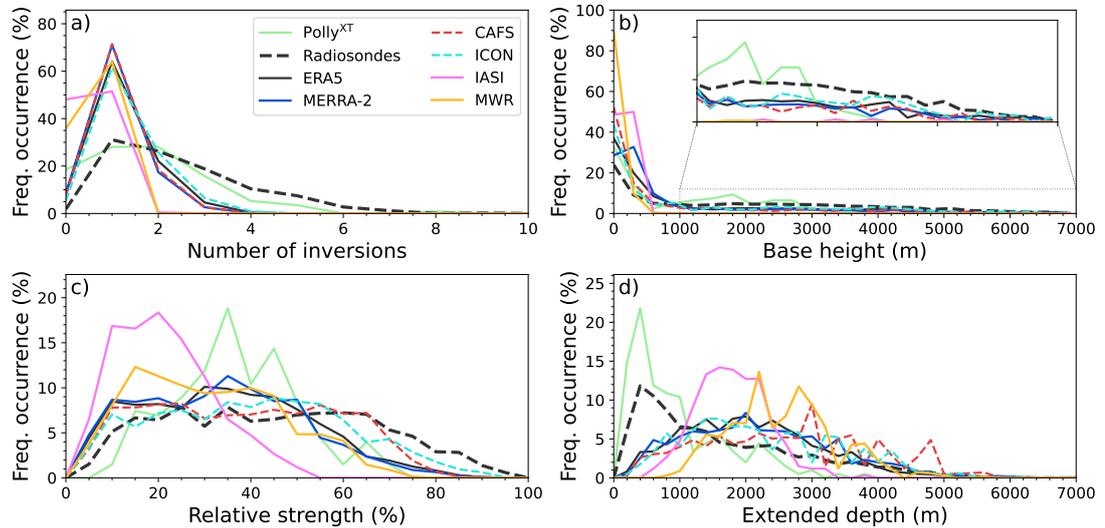


Figure 6.7: Histograms showing the distributions of the number of inversions per profile (a), the inversion base heights (b), their relative strengths (c) and extended depths (d) of different data sets. ERA5, MERRA-2, IASI, MWR and Polly^{XT} are illustrated as solid black, dark blue, pink, yellow and green lines, while the radiosondes, CAFS and ICON are indicated by dashed black, red and light blue lines. The inset axis in (b) zooms into the distribution of the base heights between 1000 and 7000 m.

The inversions detected by the radiosondes mostly have relative strengths between 15 and 65 % (Fig. 6.7c). Relative strengths outside this range occur less frequently. Compared to the radiosondes, all data sets tend to underestimate the inversion strength (Fig. 6.7c). However, the models have a smaller bias towards weaker inversions than MWR and IASI. ICON and CAFS perform slightly better concerning the representation of inversion strengths (especially relative strengths > 50 %) than ERA5 and MERRA-2. Inversions are weakest in IASI with relative

strengths mostly between 10 and 40 % and no relative strengths above 55 %. The disagreement between the radiosondes and Polly^{XT} was also found when limiting the radiosonde data to the lidar time grid. Therefore, it is likely that Polly^{XT} underestimates the strength because of its limited height range, missing upper-level inversions, which tend to have higher relative strengths (see also Fig. 6.8).

The radiosondes show a high occurrence of low extended depths of 500–1000 m and a gradual decrease of the occurrence of higher depths (Fig. 6.7d). Extended depths above 4500 m were rarely observed by the radiosondes. All data sets (except Polly^{XT}) underestimate the occurrence of inversions with extended depths below 1000 m but instead slightly overestimate the occurrence of higher depths (Fig. 6.7d). The distribution of the inversion extended depth is similar among the models and generally closer to the radiosondes than the depth distributions of IASI and the MWR. Despite the higher vertical resolution and a potential advantage regarding the assimilation of radiosondes, ERA5 does not perform better than the other models. Possibly, a much higher vertical resolution over a wider height range is needed (e.g. on the order of 10 m) to better represent sharp inversions.

If the relative strength was replaced by absolute strength in Fig. 6.7, the results would be similar to Nygård et al. (2014) (their Fig. 6) and Devasthale et al. (2011) (their Fig. 10) with some notable discrepancies. In their studies, 2–3 more inversions per profile were detected, and the depths were much smaller. Different vertical resolutions in the data sets and detection methods likely explain most differences. As mentioned in Sect. 6.3.2, this detection algorithm focuses more on the major inversions and therefore tends to detect less but deeper inversions.

Figure 6.8 shows the seasonal statistics of the frequency of occurrence, extended depth and relative strength of inversions split into height layers of 1 km depth following Chellini and Ebell (2022) and Naakka et al. (2018). Throughout the year, the radiosondes denote a high frequency of occurrence of near-surface inversions (base height below 1 km) with a slight maximum in winter (95 %) compared to the other seasons (85–90 %). Devasthale et al. (2011) and Naakka et al. (2018) also found that the highest occurrence of inversions is in winter and near the surface. The differences between Polly^{XT} and the radiosondes mainly result from the limitation to cloud-free scenes and heights below 4 km. Most differences could be eliminated when removing radiosonde data above 4 km height and restricting them to the lidar times (not shown). The only remaining notable disparity is that the extended depths of humidity inversions for Polly^{XT} are 200–600 m smaller than those in the radiosondes observations (not shown).

In winter, all data sets except IASI and Polly^{XT} agree best with the radiosondes regarding the occurrence of near-surface inversions, with occurrence rates

above 90 % (Fig. 6.8d). The excellent agreement of the occurrence, extended depth and strength between the MWR synergy and the radiosondes in winter is consistent with the small biases in the specific humidity profiles (see Fig. 6.8d–f, and Fig. 6.5d). In spring and summer, the MWR synergy (IASI) underestimates the occurrence of near-surface inversions by up to 50 (60) percentage points, which explains the relatively low accuracy of humidity inversion detection shown in Sect. 6.5.1.

In all seasons, the occurrence of inversions generally decreases with height (Fig. 6.8a, d, g, j). For radiosondes, the occurrence ranges from $< 50\%$ for inversions with base heights between 1 and 2 km to $< 10\%$ for inversions between 6 and 7 km. The highest (smallest) occurrence of inversions with base heights above 1 km was found in summer (winter and autumn). Thus, the seasonal cycle of the occurrence of elevated inversions contrasts with that of near-surface inversions. The models underestimate the occurrence of elevated inversions by 10–30 percentage points, which is consistent with the findings of Naakka et al. (2018) for other reanalyses (their Fig. 7). Both IASI and MWR detect almost no inversion above 1 km and therefore underestimate the occurrence by up to 45 percentage points, likely due to the coarse vertical resolution. In summer, the high occurrence of elevated inversions, which the MWR and IASI do not capture, could explain the stronger specific humidity biases detected in Sect 6.4.2.

The median of the extended depth of the radiosonde inversions lies between 600 and 2200 m in all seasons (Fig. 6.8b, e, h, k). The highest extended depths were found for near-surface inversions and those in the upper troposphere (above 5 km). Inversions with bases between 1 and 2 km show the smallest median extended depths in all seasons. The seasonal cycle of the extended depth is small compared to the intraseasonal variability (within a season, indicated by shading in Fig. 6.8). As already seen in Fig. 6.7, the extended depths of Polly^{XT} and the radiosonde are lower than for the other data sets, probably due to the higher resolution (Fig. 6.8b, e, h, k). All data sets generally overestimate the extended depth by at least 500–1000 m (up to 2000 m) compared to the radiosondes, except for autumn and winter where most data sets capture the depth of near-surface inversions well. Data sets with coarser vertical resolution (CAFS, MWR) tend to overestimate the extended depth more than those with a finer resolution (ERA5, ICON).

The relative strength generally increases with height as the specific humidity drops (Fig. 6.8c, f, i, l). However, in winter, the water vapour concentration is also low at the surface, and radiative cooling can generate strong surface inversions (Brunke et al., 2015; Naakka et al., 2018), resulting in a high relative strength. The relative strength of near-surface inversions observed by radioson-

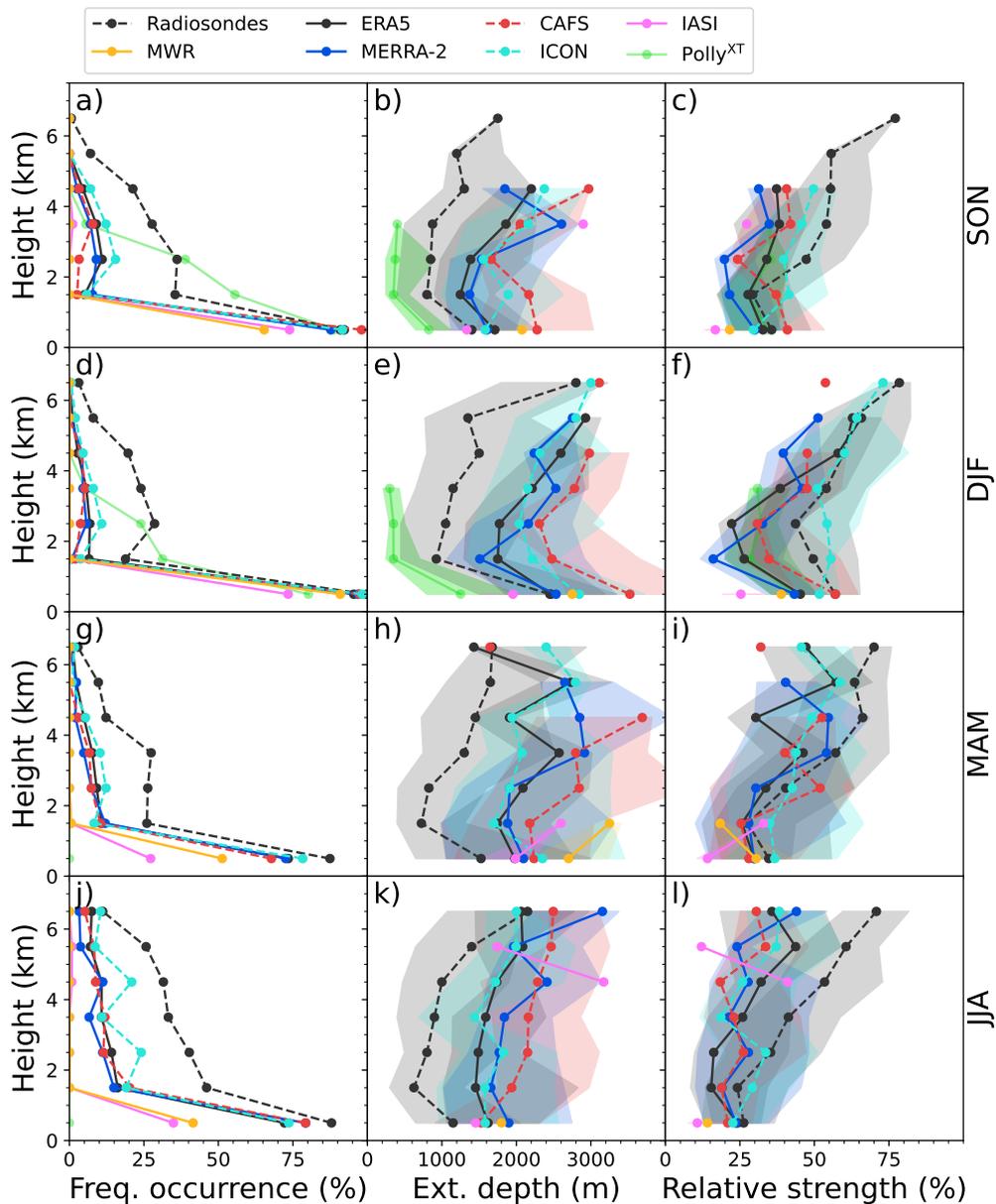


Figure 6.8: Seasonal frequency of occurrence (left column), extended depth (middle column) and relative strength (right column) of specific humidity inversions for 1000 m thick height layers. The inversion characteristics are shown for autumn (SON, a–c), winter (DJF, d–f), spring (MAM, g–i) and summer (JJA, j–l), respectively. The seasonal medians of the data sets are illustrated as in Fig. 6.7. Shading indicates the intraseasonal standard deviation.

des has a clear seasonal cycle with the highest values in winter (55 %) and lowest values in summer (25 %). The models and MWR show a similar seasonal cycle for the near-surface inversions. Near the surface, the relative strength agreement between models and radiosondes is within 10 % in spring and summer and slightly worse in winter and autumn. The bias towards weaker inversions seen in Fig. 6.7c can partly be explained by underestimating the frequency of occurrence of high altitude inversions, which often have a higher relative strength. However, a gen-

eral underestimation of the relative strengths by 5–30 percentage points above 3000 m was also found for all seasons except winter. Consistent with the findings above, ICON and CAFS perform slightly better regarding the relative strength, being closer to the radiosonde data, than ERA5 and MERRA-2.

Biases regarding the humidity inversion characteristics can partly explain specific humidity biases. For example, underestimating the strength while correctly capturing the extended depth and frequency of occurrence results in negative specific humidity biases, which can be seen for ERA5 in winter (Fig. 6.8d–f and Fig. 6.5d). In summer, IASI and the MWR synergy underestimate the occurrence and strength of inversions, causing negative specific humidity biases within the inversion layer and positive biases below or in between inversions due to smoothing.

The findings above suggest that inversions are smoother (lower strength, higher depth) in the models, MWR and IASI than in the radiosondes. Models with a finer vertical resolution (ERA5, ICON) capture the extended depth of inversions slightly better than coarse resolution models (CAFS). The number of inversions above 1 km is underestimated by all data sets, especially by MWR and IASI, which may also be attributed to the vertical resolution as suggested by Naakka et al. (2018) and Devasthale et al. (2011). However, despite the coarser vertical resolution, CAFS represents the strength of the inversions better than or at least similarly well as ERA5 and MERRA-2. Thus, the vertical resolution is probably not the only important aspect, but model physics and parametrizations associated with the formation of humidity inversions may also play a role (mostly atmospheric dynamics and moist physics, Brunke et al., 2015). Naakka et al. (2018) found that uncertainties in surface latent and sensible heat fluxes in reanalyses could be attributed to inversion strength underestimations. Further research would be needed to analyze the cause of the different performances of the models.

6.6 Downwelling longwave radiation sensitivity

After having analyzed the humidity inversion characteristics during the MOSAiC expedition, the question is to what extent humidity inversions affect the downwelling longwave radiation (DLR). It was also shown that humidity inversion characteristics differ across the data sets, resulting in specific humidity biases. Therefore, this sensitivity study also quantifies the DLR discrepancy resulting from the different specific humidity profiles of the data sets.

The T-CARS radiative transfer model (Sect. 6.3.4) was used to simulate DLR with modified specific humidity profiles as input. All DLR simulations were

performed in clear sky conditions and for cases when the Met City or ASFS skin temperature measurements were available within 5 minutes of a radiosonde launch. Cloudy scenes were excluded as their emitted radiation could conceal the water vapour changes (Curry et al., 1995). In total, 120 cases, most of which are in winter and early spring, are suitable for the simulations. The T-CARS input (e.g., temperature profiles) was not altered except for the specific humidity profiles. The radiative effect of humidity inversion modifications is quantified by the difference between the DLR based on the modified specific humidity profile and DLR simulated with the original profile $\Delta\text{DLR} = \text{DLR}_{\text{mod}} - \text{DLR}_{\text{orig}}$. The sensitivity studies start with an analysis of the DLR changes for humidity inversions at different altitudes. Next, the radiative effect of humidity inversions is quantified for all of the clear sky cases. Finally, the specific humidity profiles of the other observations (except the lidar) and models are used to estimate the direct DLR changes due to the different IWV and humidity profiles.

6.6.1 Sensitivity to inversion height and strength

To get a feeling how the DLR is affected by modifying the humidity inversion strength of inversions at different heights, a detailed analysis is presented for one case (04 February 2020 at 05 UTC). Initially, radiative transfer simulations were performed for a surface-based inversion, where the strength was modified as described in Sect. 6.3.5. Subsequently, the inversion base was shifted upwards in 500 m steps until 6500 m, and the same strength modifications were applied for each step. This procedure may be physically questionable, causing unrealistic super-saturation as the temperature profile was not altered. Therefore, this analysis is rather theoretical but helps to understand the relationship between DLR and water vapour within and below an inversion.

Figure 6.9 illustrates the modified specific humidity profiles. Below the inversion, the specific humidity is set to the value at the inversion base ($q(z_{\text{base}}) = q_{\text{base}}$), resulting in a deep layer of constant specific humidity for high inversion bases (e.g., Fig. 6.9d). As IWV is not conserved in the first part of this analysis, IWV increases when the inversion is either shifted to higher altitudes or the strength of the inversion is increased (see IWV values in Fig. 6.9).

The resulting DLR deviations (ΔDLR) due to the inversion strength and height modifications, simulated with T-CARS, are shown in Fig. 6.10. Without IWV conservation, ΔDLR clearly becomes more positive when increasing the inversion strength because the IWV increases (Fig. 6.10a). Similarly, the IWV increases when shifting the inversion to higher altitudes, causing more positive ΔDLR (vertical direction in Fig. 6.10a). However, when the inversion is lo-

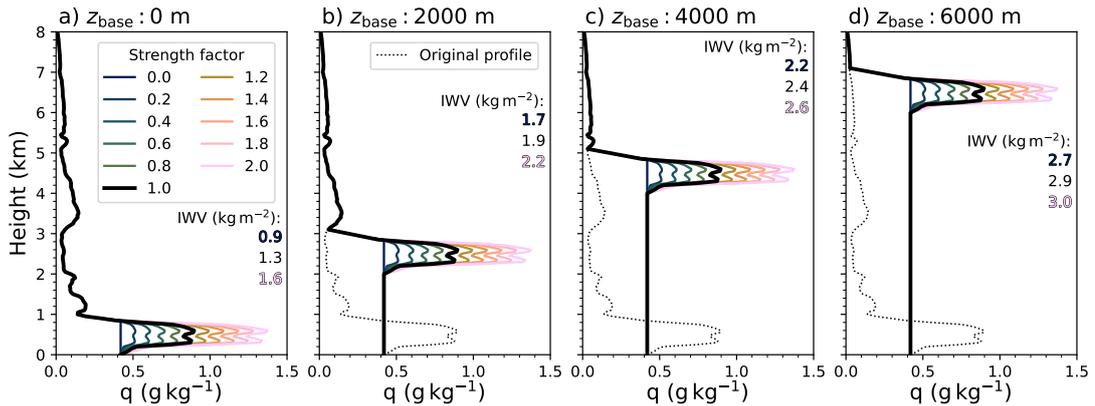


Figure 6.9: Specific humidity profiles q of a selected radiosounding (04 February 2020 at 05 UTC) with a distinct humidity inversion, whose strength and base height are modified. Only every fourth base height shift is visualized. Strength modifications with certain strength factors are indicated by colours while different base heights z_{base} are shown in each panel. Additionally, the IWW of the profiles with strength factors 0.0, 1.0 and 2.0 are displayed in their respective colours. The original q profile measured by the radiosonde is indicated with a dotted black line. The original IWW of this sounding is 1.3 kg m^{-2} .

cated at the surface ($z_{\text{base}} = 0 \text{ m}$), the increase of ΔDLR with inversion strength is much more pronounced than for an inversion with a base height of 6500 m. This behaviour is consistent with the expectations from radiative transfer: For an upward-looking ground-based observer, the transmissivity of the atmosphere, which is the fraction of an initial radiation intensity reaching the observer, decreases with height due to absorption and scattering (here, absorption) (Petty, 2006, Chapters 7.2 and 7.4). Thus, when the inversion is situated at high altitudes, only a small fraction of the emission signal from the water vapour within the inversion reaches the surface.

The sensitivity of ΔDLR to the water vapour amount within the inversion (IWW_{inv} , integrated from the base to the top) clearly decreases with height, as indicated by the weaker gradients in Fig. 6.10b. For the surface inversion, the $\partial\Delta\text{DLR}/\partial\text{IWW}_{\text{inv}}$ gradient is approximately 14.4 W kg^{-1} and is reduced to 2.1 W kg^{-1} for the elevated inversion at 6500 m (see also Table 6.3). The gradient decreases approximately exponentially with height and explains to what extent lower tropospheric water vapour contributes to the DLR compared to the upper tropospheric water vapour. This relationship can thus be used to weight specific humidity deviations between an original and a modified profile with a higher emphasis on near-surface deviations. With the two anchor points at the base heights 0 and 6500 m (Table 6.3), the weights \tilde{w} at a given altitude z_i can

be approximated by

$$\tilde{w}(z_i) = 14.4 \text{ W kg}^{-1} \exp\left(\frac{\ln\left(\frac{2.1}{14.4}\right)z_i}{6500 \text{ m}}\right), \quad (6.1)$$

$$w(z_i) = \frac{\tilde{w}(z_i)}{\sum_i \tilde{w}(z_i)}, \quad (6.2)$$

where $w(z_i)$ is the normalized weight at altitude z_i . Note that the weights may only be valid for this rather theoretical case study because the $\partial\Delta\text{DLR}/\partial\text{IWV}_{\text{inv}}$ gradient also depends on the IWV and the temperature. Nevertheless, the weights will prove to be useful to vertically weight specific humidity deviations in Sect. 6.6.3 to explain DLR deviations.

Table 6.3: Changes of ΔDLR per change of the integrated water vapour within the inversion (IWV_{inv}) without IWV conservation (as in Fig. 6.10b) for seven base heights z_{base} .

Base height z_{base} (m)	0	1000	2000	3000	4000	5000	6000	6500
$\frac{\partial\Delta\text{DLR}}{\partial\text{IWV}_{\text{inv}}}$ (W kg^{-1})	14.4	11.1	7.9	5.8	4.3	3.2	2.3	2.1

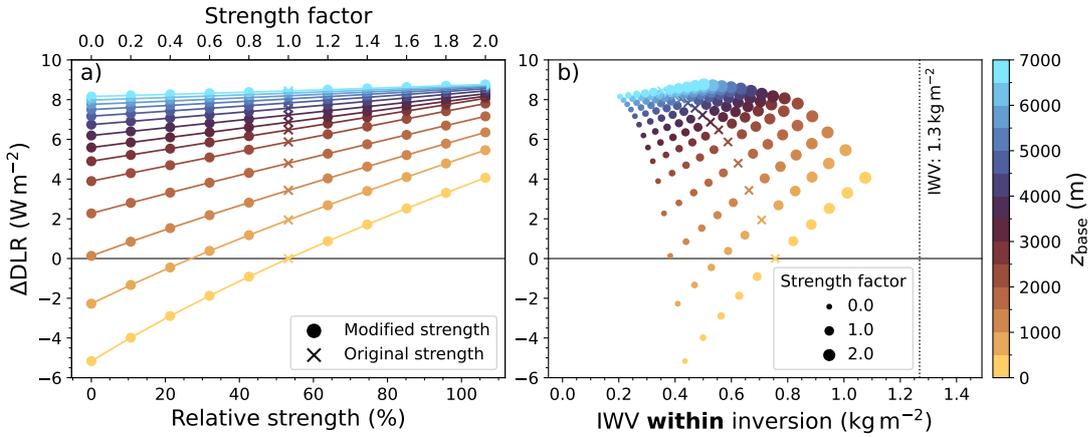


Figure 6.10: Downwelling longwave radiation differences (ΔDLR) between simulations with modified and original specific humidity profiles. The humidity profile modifications include humidity inversion base height shifts (colours) and strength modifications as shown in Fig. 6.9. ΔDLR is once illustrated against the relative inversion strength and strength factor (a) and once against the total amount of water vapour between the base and the top of the inversion IWV_{inv} (b). In the latter, the inversion strength is indicated by the size of the markers. The dotted line in (b) indicates the IWV of the original specific humidity profile.

The following part focuses on the vertical water vapour distribution to identify the height at which the water vapour emission signal of the inversion is minor compared to the water vapour below the inversion (between the surface and the base of the inversion). The effect of the IWV on the DLR is eliminated by con-

serving the IWV during humidity inversion modifications. The humidity inversion modifications illustrated in Fig. 6.11 are similar to those shown in Fig. 6.9 but the IWV is 1.3 kg m^{-2} for all profiles. IWV was conserved by removing or adding the excess or deficit of water vapour uniformly over all heights while ensuring that no negative specific humidity values occurred. For example, to conserve IWV for an inversion with a high base height (z_{base}) and whose strength has been doubled, the specific humidity has to be reduced below and above the inversion (e.g., 0–2 and 3–8 km height in Fig. 6.11b).

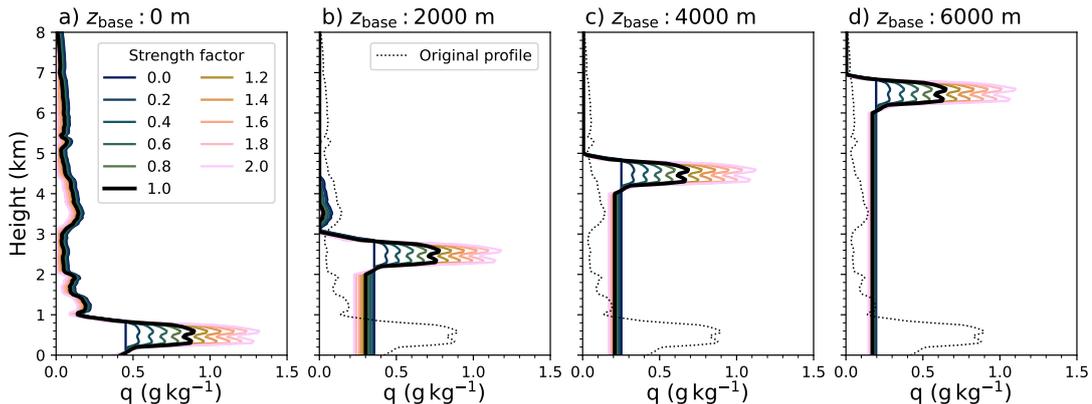


Figure 6.11: Specific humidity profiles q with humidity inversion modifications similar to those shown in Fig. 6.9 but with IWV conservation. IWV is 1.3 kg m^{-2} for all profiles.

Figure 6.12 shows ΔDLR for the humidity inversion modifications with IWV conservation. When considering inversions with base heights below 1500 m, ΔDLR also increases for higher inversion strengths (Fig. 6.12a), as found for the inversion modifications with varying IWV (Fig. 6.10a). However, the increase of ΔDLR with inversion strength is generally less pronounced because IWV does not increase. In fact, positive ΔDLR is mainly due the redistribution of water vapour to altitudes of higher temperatures. Appendix 6.E provides further details on the effect of the temperature inversion.

For inversions at altitudes above 2000 m, the differences between the simulations with and without IWV conservation are even more distinct. In contrast to the experiments with varying IWV, increasing the strength of elevated inversions results in more negative ΔDLR (Fig. 6.12a). The decrease of ΔDLR with increasing inversion strength can be explained by the vertical redistribution of water vapour and the amount of water vapour located *below* the inversion (Fig. 6.12b). An example is given for the inversion shifted to the highest altitude (6500 m): When the relative strength of the high altitude inversion is increased from 0 to twice the original strength (strength factor of 0.0 to 2.0), the amount of water vapour *below* the inversion is reduced from 1.15 to 0.9 kg m^{-2} (increasing

dot size and reduced I WV below inversion in Fig. 6.12b). Thus, the redistribution of water vapour reduces the water vapour concentration in the lower troposphere, which has the highest radiative effect as known from radiative transfer. The reduction of lower tropospheric water vapour due to the increased inversion strength therefore causes more negative ΔDLR (Fig. 6.12a).

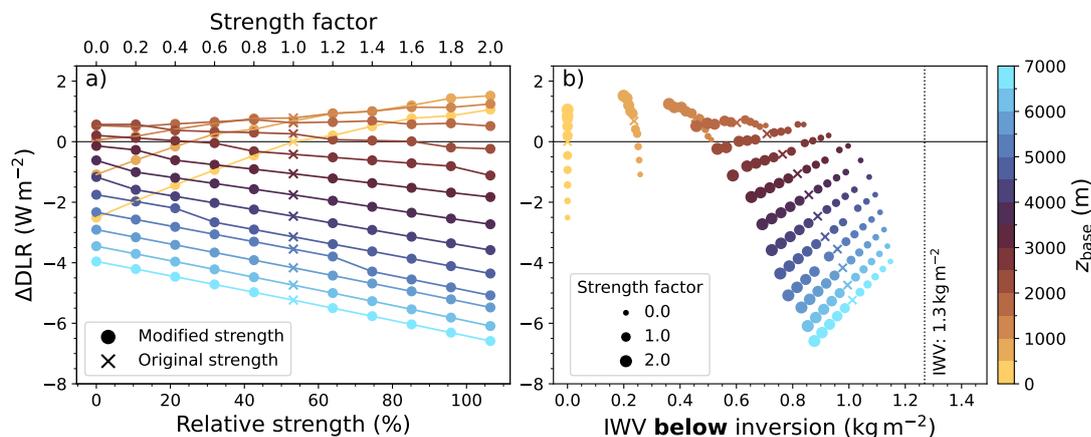


Figure 6.12: As Fig. 6.10 but showing ΔDLR for inversion modifications with IWV conservation. In (b), the relation between ΔDLR and the total amount of water vapour integrated from the surface to the inversion base (IWV below inversion) is illustrated instead of the IWV within the inversion.

In summary, the sensitivity of ΔDLR to the water vapour within the inversion decreases approximately exponentially with height, as expected from radiative transfer. For the inversion at 6500 m, most of the emission signal of the water vapour modifications within the inversion is attenuated before reaching the surface. This result is based on the findings of the first experiments (with varying IWV), where ΔDLR was barely sensitive to water vapour concentration changes within such a high-altitude inversion (see Table 6.3 and Fig. 6.10a–b).

The experiments with fixed IWV were used to identify the altitude until which the humidity inversion has a strong radiative effect at the surface. These experiments showed that at approximately 2000 m, the water vapour changes below the inversion dominate the DLR signal compared to the water vapour changes within the inversion (Fig. 6.12a). Thus, the radiative effect of inversions at 2000 m or higher is small compared to the water vapour below the inversion. In the next part, strength modifications are used for all clear sky cases to quantify the radiative effect of humidity inversions.

6.6.2 Radiative effect of humidity inversions

To obtain a statistical estimate of the radiative effect of humidity inversions in clear sky conditions, all 120 cases with available Metcity or ASFS measure-

ments were simulated with original and modified specific humidity profiles. In the modified profiles, the strength of specific humidity inversions were changed with strength factors between 0 and 1 (in steps of 0.1). A strength factor of 1 corresponds to the original inversion strength, while a strength factor of 0 removes the inversion. Humidity inversion modifications were performed once with varying and once with conserved IWV to disentangle the effect of the IWV on the DLR. The simulations with varying IWV provide the direct radiative effect of the water vapour within the inversion and therefore of the inversion itself. In contrast, the simulations with fixed IWV quantify the effect of the redistribution of water vapour from the humidity inversion to the entire profile. The simulations with fixed IWV thus imitate a data set which accurately captures the IWV but misses inversions.

In this analysis, IWV conservation is physically more consistent than in Sect. 6.6.1 to mimic more realistic specific humidity profiles. IWV is conserved by adding more of the lost water vapour (due to inversion strength reductions) at low altitudes than at high altitudes using the shape of a standard specific humidity profile: The standard specific humidity profile is computed using the air temperature of the International Civil Aviation Organization (ICAO) standard atmosphere (NOAA, NASA, and USFS, 1976), the air pressure computed via the barometric height formula and a relative humidity profile decreasing linearly from the global mean of 75 % at the surface (Peixoto and Oort, 1996) to 0 % at the standard tropopause (11 km).

In Fig. 6.13, Δ DLR is separately shown for near-surface ($z_{\text{base}} < 1000$ m) and the strongest elevated ($z_{\text{base}} \geq 1000$ m) inversions to respect the influence of the inversion base height on the Δ DLR. Note that some of the 120 cases did not have an inversion with a base above 1000 m and two cases did not feature a near-surface inversion. Thus, 118 near-surface and 99 elevated inversions are included in the analysis. Initially, simulations with varying IWV are considered (Fig. 6.13a, b). A nearly linear increase of Δ DLR with inversion strength (similar to Sect. 6.6.1) for the individual cases is observed. As water vapour is removed from the respective specific humidity profiles when reducing the strength, illustrated by the changing colours of each case (e.g., trace the outlier in Fig. 6.13b for different relative strengths), Δ DLR becomes more negative.

When the water vapour is completely removed from the near-surface inversions (relative strength set to 0), Δ DLR is up to -13 W m^{-2} (see boxplot in Fig. 6.13a). The most negative Δ DLR could be identified for near-surface inversions in dry conditions in winter when both the relative strengths and extended depths were high (≥ 70 % and > 2.5 km, respectively). For most of the 118 near-surface inversions, Δ DLR lies between -1 and -7 W m^{-2} when setting the strength to

0 (interquartile range (IQR) in Fig. 6.13a), with a median of approximately -3 W m^{-2} .

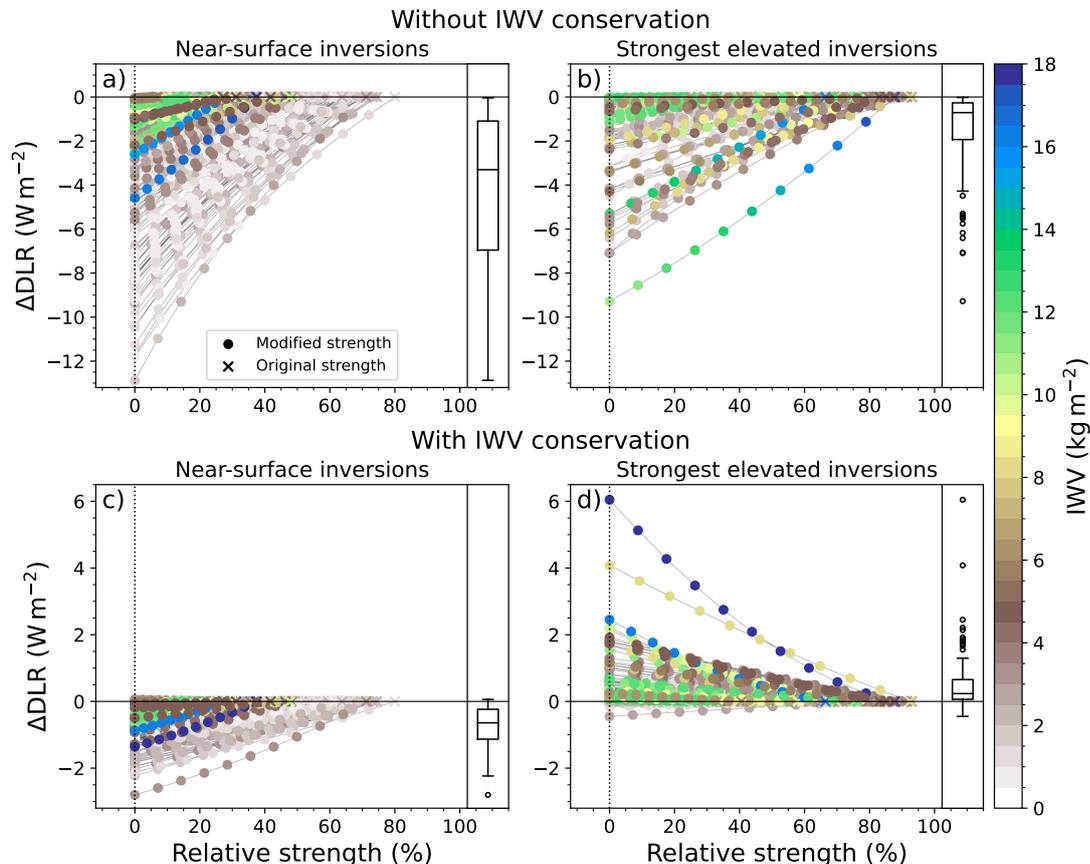


Figure 6.13: DLR deviations (ΔDLR) due to specific humidity inversion strength modifications with respect to the original specific humidity profiles for all available clear sky cases (see text for details). ΔDLR is shown separately for near-surface inversions (a, c) and the strongest elevated inversions (b, d). The modifications have been performed without (a–b) and with (c–d) IWV conservation. Original (modified) inversion strengths are indicated by "x" ("o"). Colours display the IWV of each case. The boxplot in each panel illustrates the distribution of ΔDLR when removing the inversion (strength set to 0 %, dotted line). The box indicates the interquartile range (IQR, 1–3rd quartile) of the distribution and the horizontal line within the box shows the median. The whiskers extend from below the 1st and above the 3rd quartile by $1.5 \times \text{IQR}$, respectively.

For the 99 elevated inversions, the DLR response to strength changes is weaker, consistent with the results from Sect. 6.6.1. Especially the IQR of ΔDLR is much smaller, spanning only from approximately 0 to -2 W m^{-2} with a median of -0.5 W m^{-2} (Fig. 6.13b). Outliers indicate that elevated inversions can also have a radiative effect of more than 4 W m^{-2} . Such cases mainly occurred in spring and summer, where most elevated inversions were found (Sect. 6.5.2) and when the relative strength and depth were above average.

Due to the strong relationship between DLR and IWV (up to 31 W m^{-2} per

1 kg m⁻² IWV change, Ghatak and Miller, 2013), simulations with varying IWV generally yield higher Δ DLR magnitudes than those with fixed IWV (Fig. 6.13a–b vs. c–d). In the following, the results for the simulations with fixed IWV are discussed. When eliminating the near-surface inversions with IWV conservation, the IQR of Δ DLR only extends from -0.25 to -1 W m⁻² with a median of -0.5 W m⁻² and is therefore much lower compared to the simulations with varying IWV (Fig. 6.13c). Even in the most extreme case, Δ DLR does not exceed -3 W m⁻² and is therefore just outside the uncertainty range of the pyrgeometers, which is 2.6 W m⁻² (Cox et al., 2023b).

For the strongest elevated inversions, the relationship between Δ DLR and inversion strength is generally reversed when conserving IWV (Fig. 6.13d) as observed in Sect. 6.6.1. When decreasing the inversion strength of an elevated inversion, Δ DLR is more positive because the IWV conservation method adds more water vapour to low altitudes to compensate for the loss within the inversion. In extreme cases, when the original inversion contains high amounts of water vapour, e.g., through a combination of high relative strength, depth and high specific humidity values, Δ DLR is up to $+6$ W m⁻² (Fig. 6.13d). However, in most cases, Δ DLR is negligible with values smaller than 1 W m⁻². Thus, the effect of the water vapour redistribution is negligible in most cases but can lead to considerable DLR deviations in some extreme cases.

So far, only the DLR effect of either the lowest or strongest elevated inversion has been discussed. To compute the total radiative effect of humidity inversions, all inversions of each radiosonde profile must be removed. The DLR difference between simulations with the original specific humidity profiles (with inversion, DLR_{orig}) and those where all inversions are removed ($\text{DLR}_{\text{no inv}}$) yields the radiative effect of humidity inversions

$$R_{\text{inv}} = \text{DLR}_{\text{orig}} - \text{DLR}_{\text{no inv}}. \quad (6.3)$$

Figure 6.14 shows the radiative effect of humidity inversions R_{inv} for all 120 cases. R_{inv} is mainly between 1 and 9 W m⁻² but can be up to 16 W m⁻². Comparing R_{inv} to the Δ DLR when removing only the lowest inversion reveals that the lowest inversion explains 60–100 % of the radiative effect of the inversions for all cases in winter and early spring (not shown). However, for the summer cases, which generally have more elevated inversions, the contribution of the lowest inversion is minor (20–30 %) compared to the elevated inversions, which contribute about 70–80 % in these cases (not shown). Higher R_{inv} occurred when either the total water vapour amount within all inversions (IWV_{inv}) is high, or when IWV_{inv} is high *relative* to the total IWV (Fig. 6.14b). Here, the results suggest that

the inversions can contain up to 70 % of the total IWV, which is much higher than the values found by Devasthale et al. (rarely more than 40 % 2011). The discrepancies can be attributed to the different integration limits and inversion detection methods.

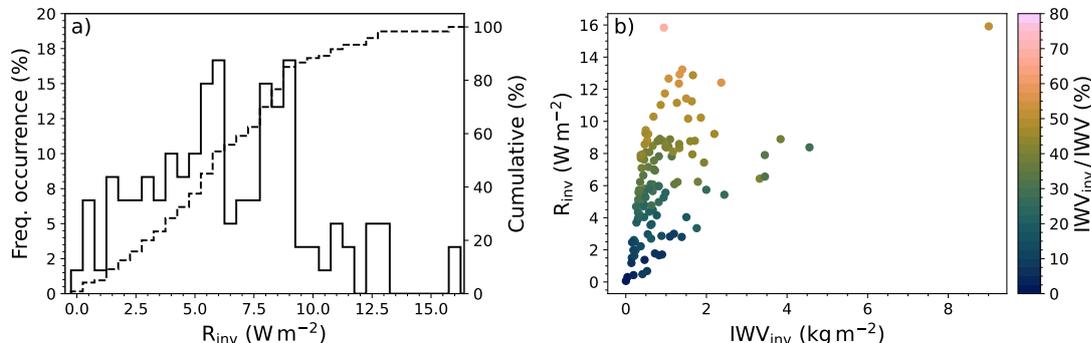


Figure 6.14: Frequency of occurrence (solid) and cumulative occurrence (dashed) of the radiative effect of humidity inversions (R_{inv}) for the 120 clear sky cases (a). Additionally, the relationship between the water vapour within all inversions (IWV_{inv}) and R_{inv} is visualized (b). Colours in (b) indicate the relative IWV_{inv} to the total IWV.

In summary, the results showed that near-surface inversions have a radiative effect of 1–9 W m^{-2} in most cases and up to 16 W m^{-2} for very deep and strong inversions. The radiative effect of elevated inversions is minor compared to the lowest inversion, except for summer when elevated inversions frequently occur. Additionally, these sensitivity studies were used to investigate the effect of the redistribution of water vapour (from a humidity inversion to the entire profile) on the DLR. The results confirm the expectation that the redistribution of water vapour generally has a negligible effect on DLR (mostly $< 1 \text{ W m}^{-2}$) and mainly depends on the water vapour distribution in the lower troposphere. However, in some cases with high amounts of water vapour within the inversion, the mere redistribution of water vapour can also result in DLR changes of more than 4 W m^{-2} . For example, if a model or remote sensing observation correctly captures the IWV but misses such an inversion, DLR simulations would be biased by 4 W m^{-2} . Having analyzed the effect of humidity inversions on DLR with synthetic specific humidity profile deviations, the next step is to quantify the effect of the actual specific humidity profile deviations of the different data sets on DLR.

6.6.3 Radiative impact of specific humidity deviations of different observations and models

The evaluation of the different models (ERA5, MERRA-2, ICON and CAFS) and remote sensing observations (IASI and MWR) revealed certain specific humidity

biases with respect to the radiosondes (Sect. 6.4.2), as well as discrepancies regarding the representation of humidity inversions (Sect. 6.5). This part explores the impact of the data sets' specific humidity deviations on DLR for all 120 clear sky cases (mostly winter and early spring). Here, the Polly^{XT} data was excluded because of low availability and limited height range. By using the different data sets' specific humidity profiles as input to T-CARS, the combined effect of different vertical water vapour distributions and IWV values is examined with respect to the radiosonde profiles. As the radiosonde profiles were used as reference, IWV differences were also computed with respect to the radiosonde IWV (contrasting Sect. 6.4.1). No other T-CARS input parameters were altered.

Figure 6.15 shows the DLR deviations (ΔDLR) of the different data sets, as well as IWV and specific humidity deviations with respect to the radiosonde observations. The IWV deviation (ΔIWV) histogram (Fig. 6.15c) focuses on the main deviations and was therefore cut at $\pm 2 \text{ kg m}^{-2}$. Note that the IWV for the MWR was computed with the retrieved specific humidity profile instead of using the dedicated IWV retrieval for consistency, which results in slightly larger deviations with respect to the radiosondes. Specific humidity deviations $\overline{\Delta q}$ were vertically averaged using the weights computed in Eq. 6.2 to emphasize lower tropospheric deviations (Fig. 6.15d). High correlations between $\overline{\Delta q}$ and ΔDLR (0.55–0.87, Fig. 6.15e) demonstrate that these exponential weights are a good approximation to account for the higher sensitivity of ΔDLR to lower tropospheric specific humidity changes.

The ΔDLR resulting from the different specific humidity profiles of the data sets are mostly within $\pm 2 \text{ W m}^{-2}$, as indicated by the interquartile ranges (IQRs) in Fig. 6.15a, and are therefore within the typical uncertainty range of the pyrgeometers used during MOSAiC (2.6 W m^{-2} Cox et al., 2023b). However, it is important to look at the details because some specific humidity and IWV biases are systematic and could therefore result in systematic DLR biases.

Positive and negative ΔDLR of more than $\pm 5 \text{ W m}^{-2}$ were found for MERRA-2, ICON, CAFS and IASI (Fig. 6.15a). For ICON, positive and negative ΔDLR occur similarly often, resulting in a median ΔDLR of nearly 0 W m^{-2} . The strongest median ΔDLR , which can be interpreted as DLR bias, were found for MERRA-2 (-1 W m^{-2}) and IASI (0.5 W m^{-2} , Fig. 6.15a). The DLR deviations can be explained by the IWV and weighted specific humidity deviations: Here, for the simulated cases, MERRA-2 (IASI) show more negative (positive) IWV deviations and even more pronounced negative (positive) $\overline{\Delta q}$ than the other data sets (Fig. 6.15c, d). The low inversion detection accuracy of IASI due to numerous misses (only 0.34 for the 120 cases simulated here, not shown) can explain the positive specific humidity deviations: When missing the humidity inversions,

which are mostly surface-based, IASI typically overestimates the specific humidity in the lower part and at the extended top of the inversion and only slightly underestimates the maximum specific humidity (not shown), resulting in overall positive $\overline{\Delta q}$.

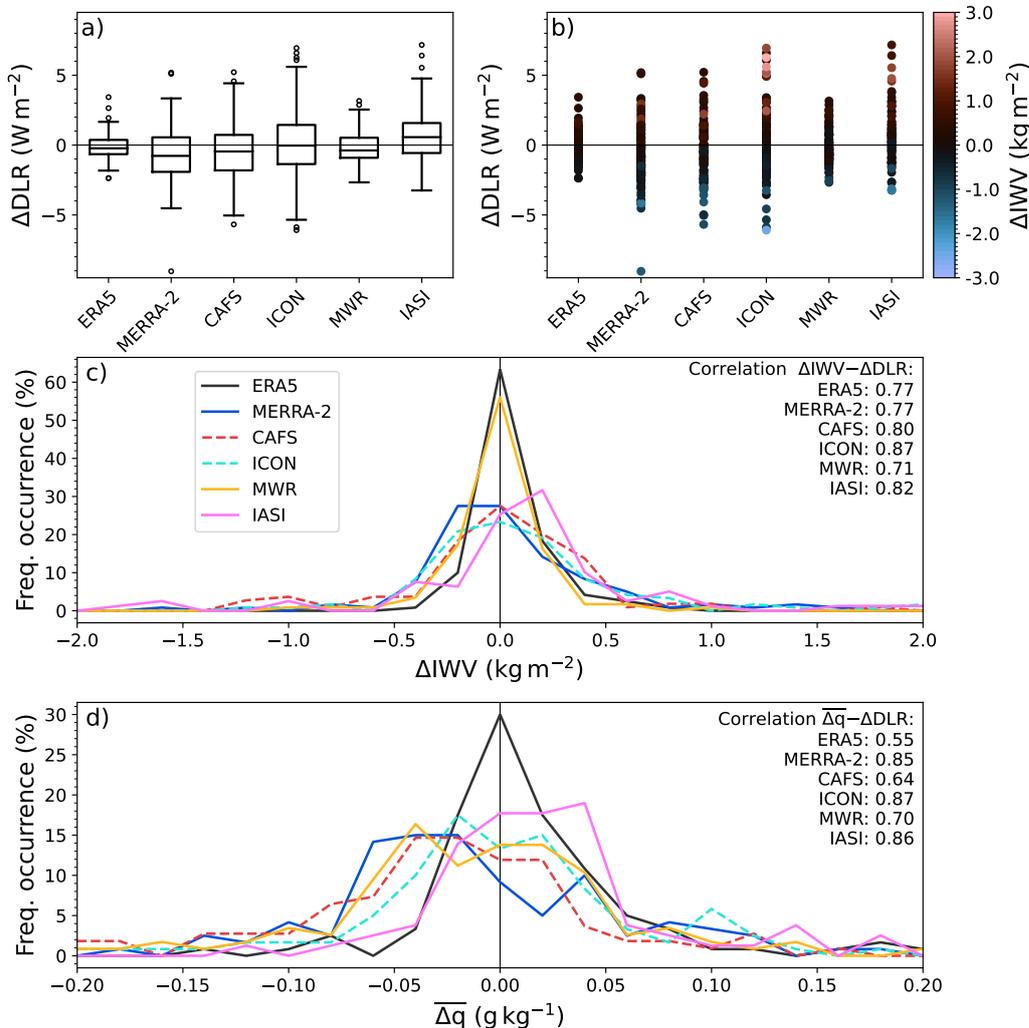


Figure 6.15: Difference between modified and original DLR (ΔDLR) for 120 clear sky cases using the specific humidity profiles of different data sets (ERA5, MERRA-2, CAFS, ICON, MWR and IASI). ΔDLR is visualized as boxplots (a) and scatter plots (b), where colours indicate IWV differences ΔIWV to the radiosonde observations. Additionally, histograms of ΔIWV (c) and the weighted vertically averaged specific humidity deviations $\overline{\Delta q}$ (d) are shown. Additionally, the Pearson correlation coefficients between ΔIWV and ΔDLR , as well as between $\overline{\Delta q}$ and ΔDLR are given for each data set. The different data sets in (c) and (d) are illustrated using the same colours as in Fig. 6.8.

ΔDLR outliers of MERRA-2, ICON, CAFS and IASI exceed $\pm 5 W m^{-2}$ and are mostly related to the strongest IWV deviations (-2 to $+3 kg m^{-2}$, Fig. 6.15b). However, strong IWV deviations cannot explain all ΔDLR outliers: For example, MERRA-2 has two outliers of $+5 W m^{-2}$, which are associated with minor IWV

deviations of only 0.5 kg m^{-2} . In these cases, MERRA-2 strongly overestimates the specific humidity in the lowest 2 km, resulting in one of the highest $\overline{\Delta q}$ of 0.17 g kg^{-1} , but agrees well with the radiosonde specific humidity or even underestimates it at higher altitudes. For IASI, the highest positive ΔDLR outlier can also be explained by the overestimation of the specific humidity in the lower troposphere.

Data sets with a more centered ΔIWV and $\overline{\Delta q}$ distribution (e.g., ERA5, ICON, MWR and CAFS) have a nearly 0 W m^{-2} median ΔDLR . ERA5 and the MWR synergy denote the smallest IQRs of ΔDLR with values less than 0.5 and 1 W m^{-2} , respectively (Fig. 6.15a), as well as the smallest ΔDLR outliers (below $\pm 3.5 \text{ W m}^{-2}$) and the smallest IWV deviations of all data sets (Fig. 6.15c). Thus, the MWR synergy yields similarly accurate DLR as ERA5 despite having much coarser specific humidity profiles. As the MWR synergy resolves surface-based inversion well, $\overline{\Delta q}$ is similar to that of ICON and CAFS and therefore only slightly higher than that of ERA5 (Fig. 6.15d).

ΔIWV and $\overline{\Delta q}$ are both highly correlated with ΔDLR (Fig. 6.15c, d) because they contain similar information and are also well correlated with each other (0.61–0.95 for the data sets, not shown). However, it was shown that ΔIWV could not explain all ΔDLR . Thus, a good representation of the near-surface specific humidity, which often features an inversion, is equally important as the IWV to avoid DLR biases.

In summary, for the 120 clear sky cases, specific humidity profile and IWV deviations from the radiosonde profiles mostly result in relatively small ΔDLR of less than $\pm 2 \text{ W m}^{-2}$ for all data sets. However, in cases when either the IWV or the lower tropospheric specific humidity is strongly over- or underestimated, ΔDLR can exceed $\pm 5 \text{ W m}^{-2}$. Despite the coarse specific humidity profiles of the MWR, the ΔDLR are similarly small as for ERA5, which showed the smallest ΔDLR .

6.7 Conclusion

Humidity estimates in the Arctic Ocean are highly uncertain due to the lack of ground stations and challenges in satellite remote sensing, which also lead to uncertain IWV trends in that region. The humidity measurements from the MOSAiC expedition provide an excellent opportunity to assess the quality of water vapour products of models and satellite products in the Arctic Ocean. This study evaluated IWV, specific humidity and the characteristics of humidity inversions of four state-of-the-art models (ERA5, MERRA-2, ICON, CAFS) and satellite observations (IASI combined sounding product, and a new AMSR2-based

IWV retrieval) using reference data from radiosondes and the synergy of the MWRs HATPRO and MiRAC-P. Then, the impact of humidity inversions on the DLR was analyzed using the radiative transfer tool T-CARS.

The numerical models generally show excellent IWV agreement regarding biases and RMSD compared to the reference IWV given by the MWR synergy, except for negative biases of up to -10% (-5%) that were found for MERRA-2 (ICON and ERA5) in dry conditions. The IASI combined sounding product (henceforth, IASI) performs well in dry conditions but underestimates IWV by $10\text{--}15\%$ for IWV above 10 kg m^{-2} . The underestimation of IWV in humid conditions is likely related to clouds and rapidly changing surface characteristics during the melt season. In contrast to the findings of Roman et al. (2016), IASI shows no positive biases in dry conditions. AMSR2 overestimates IWV below 3 kg m^{-2} , which appears to be the sensitivity limit, and underestimates IWV in more humid conditions. Some of the IWV deviations of the tested data sets could be attributed to intraweek IWV variability (e.g., related to storms), whose timing and magnitude were not always well caught by the satellite products due to orbit limitations.

Concerning the RMSD of the specific humidity profiles, the MWR synergy performs similarly well as or even better than MERRA-2, ICON and CAFS at $0\text{--}4\text{ km}$ height in all seasons (RMSD of $10\text{--}30\%$). IASI shows higher RMSD at altitudes below 3 km but agrees slightly better with radiosondes than the MWRs above 3 km . The smallest RMSD was found for ERA5 at all heights and in all seasons ($< 25\%$). Specific humidity profiles in the numerical models are mostly negatively biased in the lowest 2 km during the cold seasons with respect to the radiosondes. IASI tends to overestimate the specific humidity in the lowest 1 km and underestimates it at higher altitudes in the cold seasons. MERRA-2 has the overall strongest and the MWR synergy the smallest bias in the lower troposphere in winter. Note that the assimilation of the MOSAiC radiosondes likely improved the performance of MERRA-2, ICON and especially ERA5, which uses an advanced assimilation scheme (4D-Var). Thus, these results may not reflect the actual performance over sea ice when there is no field campaign data for assimilation.

In this study, a new specific humidity inversion detection scheme was developed, focusing more on the major inversions than other studies (e.g., Devasthale et al., 2011; Nygård et al., 2014). Nygård et al. (2014) stressed that inversion characteristics, especially inversion base height, strength and depth, are sensitive to the chosen data (resolution) and methods (i.e., detection scheme, minimum strength and depth thresholds). Thus, the presented results are not directly comparable with those of previous studies. At first, the detectability of inversions

of the different data sets was tested. Due to the high occurrence of inversions (98 % of all radiosondes), the accuracy value calculated from the contingency table represents a reasonable estimate of the detectability of inversions. The MWR synergy detects the presence of inversions with an accuracy of 0.66, which is better than IASI but worse than the models. In addition, the synergy of HATPRO and MiRAC-P results in a significant improvement in the detectability of inversions compared to the HATPRO-only retrievals (accuracy of 0.66 vs. 0.03). All models detect the presence of inversions well with accuracy values above 0.92 but sometimes miss the absence of an inversion (especially ICON and CAFS).

Then, humidity inversion characteristics were compared. The number of inversions in a profile is mostly between 1 and 4 in the radiosonde data. All other data sets strongly underestimate the occurrence of cases with 2 or more inversions because elevated inversions are often missed. IASI and the MWR synergy rarely detect more than 1 inversion due to their coarse vertical resolution and generally miss inversions with bases above 600 m or predict them as surface inversions. In winter, the presence of near-surface inversions, which occurred 95 % of the time, is well caught by all models and the MWR. In other seasons, all data sets underestimate the occurrence of near-surface inversions. All data sets tend to slightly underestimate the inversion strength of elevated inversions and overestimate the extended depth compared to the radiosondes. The magnitude of the depth overestimation is mostly between 500 and 1000 m and is highest for CAFS and the MWR and smallest for ICON and ERA5. The results suggest that ERA5 does not represent humidity inversions better than the other models (regarding occurrence, strength and depth) despite having the highest vertical resolution. Model physics (e.g., turbulent fluxes) have also been related to discrepancies in humidity inversion characteristics (Naakka et al., 2018). More detailed research is needed to examine the origin of the discrepancies.

To estimate the sensitivity of the DLR to water vapour within humidity inversions, the humidity inversion strength was modified in clear sky scenes. The accordingly modified specific humidity profiles were used as input to the radiative transfer model T-CARS. No other parameter was altered to focus on the effect of water vapour. The sensitivity tests started with a single inversion, which was progressively shifted to higher altitudes and whose strength was changed at each step. The sensitivity of DLR to the water vapour within the inversion decreases nearly exponentially with height, as expected from radiative transfer. When the inversion is shifted to 2000 m or higher, specific humidity changes within the inversion (strength modifications) have a smaller effect on the DLR than humidity changes below the inversion.

The second part of the DLR simulations provides a statistical estimate of the

radiative effect of humidity inversions in clear sky conditions, which has been quantified for the first time, isolated from the effects of clouds and temperature. The radiative effect of humidity inversions was computed for 120 cases as the difference between the DLR based on the original specific humidity profile (with inversion) and the DLR based on the profile without inversion. For most of the cases, the radiative effect of humidity inversions is between 1 and 9 W m^{-2} but can also reach 16 W m^{-2} . In winter and early spring, the lowest (surface-based) inversion contributes more than 60% to the total radiative effect. In summer, when the occurrence of elevated inversions is higher, elevated inversions dominate the total radiative effect.

The additional DLR provided by humidity inversions is likely only relevant in clear sky conditions because the longwave radiative effect of clouds is usually much larger, obscuring the DLR effect of the humidity inversions ($\approx 50 \text{ W m}^{-2}$ Shupe and Intrieri, 2004). Additionally, the presence of clouds prevents the strong radiative cooling at the surface and therefore terminates one of the main formation mechanisms of surface-based humidity inversions. The isolated radiative effect of humidity inversions in cloudy conditions has yet to be analyzed in detail by modifying those humidity inversions that are not related to the cloud layer (for physical consistency).

In the final part of the DLR simulations, specific humidity profiles of each data set were used as input to T-CARS to estimate the uncertainty of DLR due to different specific humidity profiles and IWV. Lower tropospheric specific humidity and IWV deviations with respect to the radiosondes caused DLR differences of $\pm 2 \text{ W m}^{-2}$ in most of the 120 clear sky cases. The specific humidity profiles from ERA5 and MWR resulted in the smallest DLR deviations, where even the outliers lie within $\pm 3.5 \text{ W m}^{-2}$. The results demonstrate that the specific humidity profiles of the MWR synergy are similarly well suited for longwave radiation calculations as the ERA5 profiles. MERRA-2, ICON, CAFS and IASI denote DLR deviations of more than $\pm 5 \text{ W m}^{-2}$ because of greater IWV and specific humidity deviations. Both the near-surface specific humidity and the IWV must be well represented to avoid DLR biases. As the near-surface specific humidity is usually characterized by an inversion, the characteristics of the near-surface inversions (occurrence, strength and depth) should be captured accurately.

With the findings of this study, the research questions posed in the introduction (Sect. 6.1) can be answered:

1. How well do current models represent the IWV and the vertical water vapour distribution?

IWV estimates of the models agree well with the MWR reference data, except for the negative biases of MERRA-2 and, to a lower extent, of ICON

and ERA5 in dry conditions. Specific humidity RMSD with respect to radiosondes are similar for all models (10–30 %) except ERA5, which has the smallest RMSD. All models underestimate the specific humidity in the lowest 2 km in the cold seasons, except a shallow surface layer. It must be stressed that the evaluation of ERA5, MERRA-2 and ICON is likely strongly influenced by the assimilated radiosondes.

2. Are water vapour profiles and the typical Arctic humidity inversions sufficiently well captured by satellite and ground-based remote sensing to correctly assess the DLR?

In winter, the MWR errors are small in the lowest 4 km and surface inversions are well detected. As most inversions are surface-based, and surface-based inversions are radiatively most important, the humidity profiles derived from MWR synergy are suitable for accurate DLR calculations in winter. Radiative transfer simulations confirmed that DLR estimates from MWR profiles are accurate. Humidity profiles from IASI resulted in significantly more uncertain DLR estimates because of higher errors in the lower troposphere and missing inversions.

3. Is an accurate IWV estimate sufficient to calculate DLR? To what extent does the vertical water vapour distribution, especially with regard to inversions, affect the DLR?

Radiative transfer simulations showed that the radiative effect of humidity inversions can be up to 16 W m^{-2} if the water vapour within the inversions is removed. Sensitivity studies showed that IWV deviations could explain most but not all DLR deviations. In many cases, good representations of IWV and the near-surface specific humidity, and therefore the characteristics of surface-based inversions, are equally important.

As a next step, the DLR sensitivity tests could be extended to cloudy scenes to assess to what extent clouds conceal the direct DLR effect of humidity inversions. It would also be of interest to analyze the origin of misrepresentations of humidity inversions to understand why ICON and CAFS represent humidity inversions equally well as ERA5 despite their coarser vertical resolution. As humidity inversion characteristics vary across the Arctic (e.g., Devasthale et al., 2011), our inversion characteristics and DLR sensitivity analyses could be extended to other field campaigns and ground stations (e.g., Ny-Ålesund on Svalbard, Norway) to analyze regional differences. The long-term radiosonde observations at Ny-Ålesund are suitable to investigate trends in humidity inversion characteristics.

6.A Weighted average interpolation

Linear interpolation can lead to considerable errors in the presence of strong gradients, which can add up to an error in integrated quantities. Figure 6.16 demonstrates different vertical interpolation options for a specific humidity profile and the resulting IWV. Unweighted vertical averaging and weighted vertical averaging yield more accurate results because the vertical distribution is better represented. However, the unweighted vertical averages can also result in significant errors (Fig. 6.16).

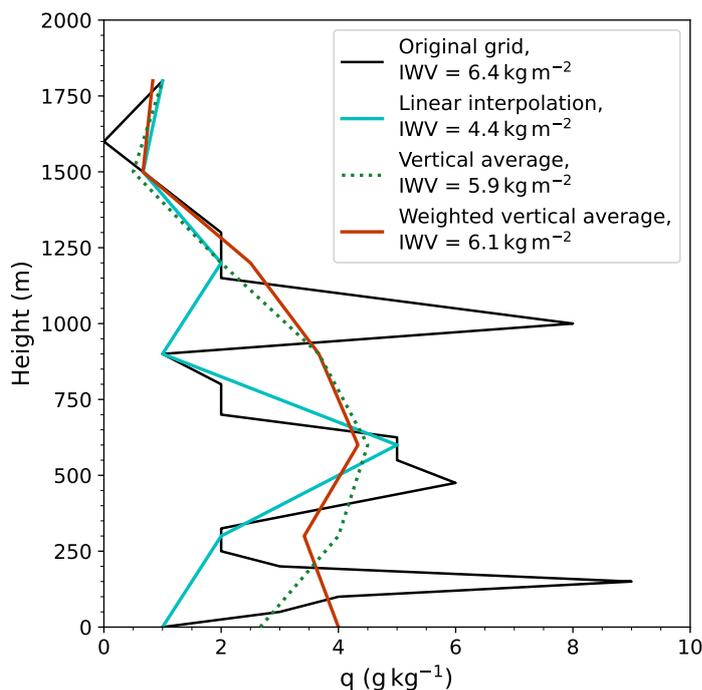


Figure 6.16: Example specific humidity profile shown for different height grids. The original profile (black) has been linearly interpolated (cyan), vertically averaged without (green dotted) and with weighting (orange) to a coarser height grid.

The goal of the weighted average interpolation is to keep the vertical distribution of data, which is on a base height grid z , as close to the original profile as possible while changing to a target height grid z_{ip} with a lower or higher vertical resolution. For the averaging, height layers of the base and target grid (z_{lay} and $z_{ip,lay}$, respectively) are considered. Height levels and layers at level i are generally related by $z_{lay,i} = 0.5(z_{i-1} + z_{i+1})$ if height level i is not at the top or bottom of the height grid, where grid boundaries must be respected.

The vertical averaging, visualized in Fig. 6.17, will be explained with an example. Initially, height levels of the base grid z , whose layers z_{lay} partly or fully lie within the current height layer of the target grid ($z_{ip,lay,4} - z_{ip,lay,5}$), are identified (here, z index 14, 15 and 16, coloured boxes in Fig. 6.17). Each base grid height

layer (coloured box) is then weighted according to its area (or fractional length) that lies within the current target grid layer. As only a small fraction of the base grid layer $z_{\text{lay},14}-z_{\text{lay},15}$ (yellow box) overlaps with the current target grid layer ($z_{\text{ip},\text{lay},4}-z_{\text{ip},\text{lay},5}$), its weight is small. The other two base grid layers, $z_{\text{lay},15}-z_{\text{lay},16}$ (green box) and $z_{\text{lay},16}-z_{\text{lay},17}$ (blue box), are fully or mostly within the target grid layer and therefore stronger weighted. In this example, an unweighted vertical average would only include z_{15} and z_{16} although the associated height layer of z_{16} ($z_{\text{lay},16}-z_{\text{lay},17}$) partly lies outside the target grid layer. The unweighted vertical average would completely ignore the data from z_{14} .

The weight w of each identified height level is computed by the fraction of its associated layer to the thickness of the current target grid layer, e.g., $w_{15} = (z_{\text{lay},16} - z_{\text{lay},15}) / (z_{\text{ip},\text{lay},5} - z_{\text{ip},\text{lay},4})$ for height level 15 (green box). For height levels whose layers do not fully lie within the current target grid layer, the boundaries of the target grid layer are respected. Once all weights are determined, the interpolation is performed with a weighted sum $x_{\text{ip}} = \sum_i w_i x_i$ where x and x_{ip} is the data on the base and target grid, respectively.

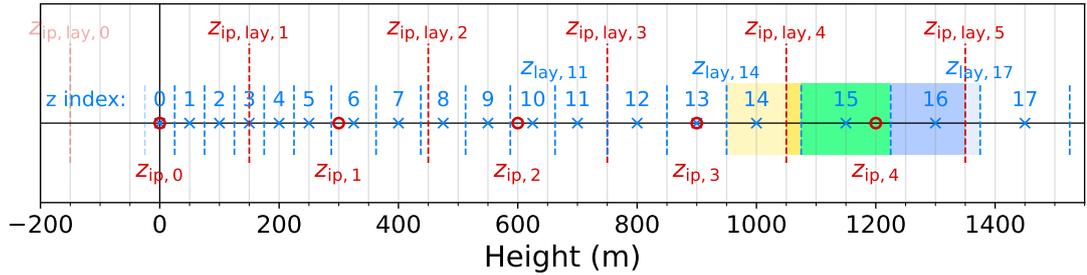


Figure 6.17: Visualization of a weighted average to interpolate data from a base height grid with levels z (blue crosses) and layers z_{lay} (blue dashed lines) to a target height grid with levels z_{ip} (red circles) and layers $z_{\text{ip},\text{lay}}$ (red dashed lines). Coloured boxes indicate which layers of the base grid are considered to compute data at $z_{\text{ip},4}$. Faded colours illustrate which height layer parts of z_{14} and z_{16} are not included in the weighed average (for details, see text).

6.B Integrated water vapour comparison with 7-day running mean

To evaluate the IWV without the sub-weekly IWV variability, which may be caused by storms or other synoptic events, a 7-day running mean has been applied to all data sets. The resulting deviations are shown in Fig. 6.18 and are generally smaller than in the evaluation without running mean. The running mean helps to identify the main deviations and eliminate errors due to uncertainties

regarding the exact timing of storms. Despite using the assimilated radiosondes, MERRA-2, ERA5 and ICON denote one of the strongest negative biases of -5 to -10% for IWV below 5 kg m^{-2} , and positive biases up to 4% for above 5 kg m^{-2} . The negative bias of IASI (here, about -10%) and its relatively strong RMSD (also about 10%) remain, although the RMSD has a slightly smaller magnitude compared to the unfiltered comparison. RMSD and biases of AMSR2 also seem to be strongly affected by sub-weekly IWV variability because filtering them out results in significantly smaller RMSD and biases, even smaller than for IASI for IWV above 13 kg m^{-2} .

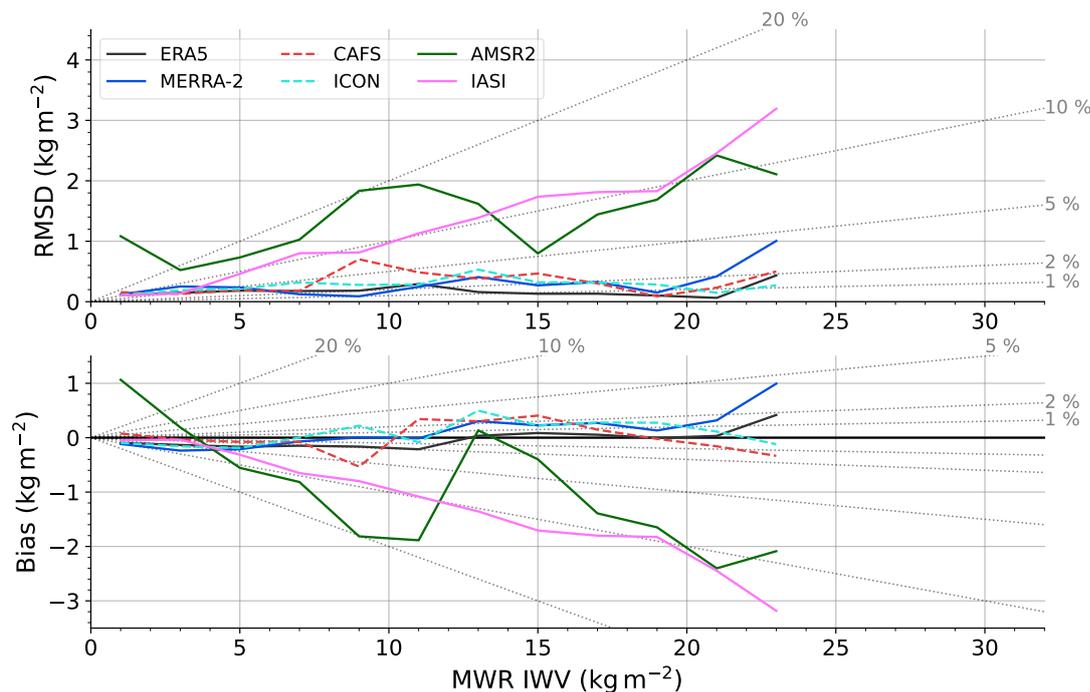


Figure 6.18: As Fig. 6.4 b) and c), but a 7-day running mean has been applied to all data sets.

6.C Vertical grid spacing of all data sets

The vertical grid spacing of each data set used in the specific humidity profile evaluation (Sect. 6.4.2) is illustrated in Fig. 6.19. The grid spacing may not be representative of the actual vertical resolution in some data sets (e.g., MWR and IASI) but shows the resolution to which radiosonde profiles were reduced for the comparison using the weighted vertical mean (Appendix 6.A). Before interpolating the radiosonde data to the other data sets' height grids, the radiosondes were brought to a height grid with 5 m spacing (Fig. 6.19). For the specific humidity profile evaluation, Polly^{XT} data were interpolated to a grid with 50 m spacing. All other data sets feature height grids whose spacing increases with height. For

example, ERA5 has a resolution (grid spacing) of about 25 m at the surface and about 250 m at 8 km height (Fig. 6.19). IASI has the coarsest grid in the lowest 1 km, while CAFS's height grid is coarsest at higher altitudes and similar to MERRA-2 at 8 km.

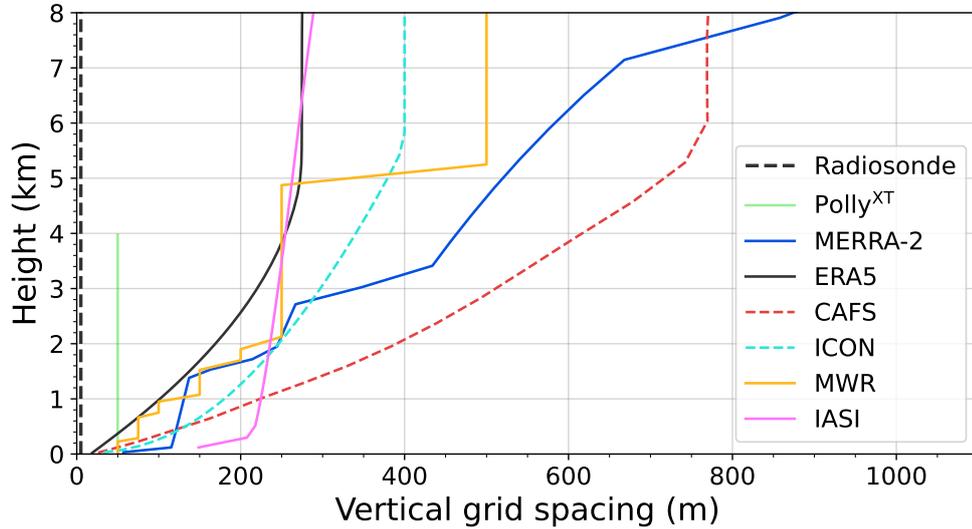


Figure 6.19: Vertical grid spacing of each data set (radiosonde: black dashed, Polly^{XT}: green solid, MERRA-2: blue solid, ERA5: black solid, CAFS: red dashed, ICON: blue dashed, MWR: yellow solid, IASI: pink solid).

6.D Benefit of the microwave radiometer synergy for humidity inversion detection

The improvement of the humidity inversion detectability can be assessed by comparing the contingency table statistics of the synergy of HATPRO and MiRAC-P to HATPRO-only. Table 6.4 shows the same statistics as Table 6.2 in Sect. 6.5.1. While the synergy correctly detected the presence and absence of humidity inversions in 66% of the time, HATPRO observed the correct situation only in 3% of the time. All cases when no inversion was observed in the radiosonde were correctly captured by both retrievals (probability of false detection is 0). As both the synergy and HATPRO did not have a false alarm (false positive) the success ratio is 1.00 for both. The Heidke skill score of the synergy is better than HATPRO's score (0.05 vs. 0.00) because the fraction of misses to the correct negatives is lower. The results are similar when using the entire MOSAiC year instead of the limited time range until 05 August 2020.

Table 6.4: As Table 6.2 but comparing only the MWR synergy and HATPRO.

	N	N_{inv}	Accuracy	Bias	POFD	SR	HSS
Synergy	1064	682	0.66	0.65	0.00	1.00	0.05
HATPRO	1087	12	0.03	0.01	0.00	1.00	0.00

6.E Downwelling longwave radiation sensitivity: Influence of temperature inversion

In addition to the sensitivity analysis of the downwelling longwave radiation (DLR) to specific humidity inversion modifications (Sect. 6.6.1), the effect of the temperature inversion is explained. Here, inversion modifications have also been performed with IWV conservation. Figure 6.20 shows the respective modified specific humidity profiles and the original temperature profile.

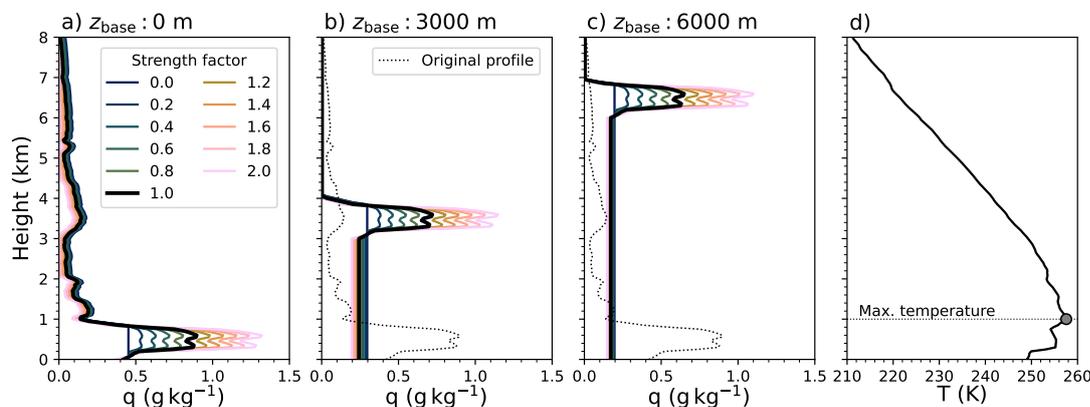


Figure 6.20: As Fig. 6.11 but showing different inversion base heights (z_{base}) in (b) and (c), and the temperature profile of the radiosounding in (d).

The simulated DLR of a modified specific humidity profile is compared against the DLR based on the original specific humidity profile by computing the difference $\Delta\text{DLR} = \text{DLR}_{\text{mod}} - \text{DLR}_{\text{orig}}$ as in Sect. 6.6. Note that the small ΔDLR jumps with inversion strength increase (e.g., at strength factor 1.2–1.4 for the base height of 5000 m, Fig. 6.21a) are partly caused by vertically varying temperature, and partly by additional humidity profile corrections performed to avoid negative specific humidity.

Positive ΔDLR (up to $+1.5 \text{ W m}^{-2}$) occurred only for inversions with base heights at or below 2500 m (Fig. 6.21a) and are related to the near-surface temperature inversion, whose maximum temperature is at 1000 m height (Fig. 6.20d). For example, when the humidity inversion is at the surface (Fig. 6.20a), positive ΔDLR are found for high relative strengths (Fig. 6.21a). Then, more water vapour is located at the altitudes of the highest temperatures. In contrast, for elevated inversions, e.g., with a base height of 2000 m, low inversion strength

is required to obtain positive ΔDLR (Fig. 6.21a). In this case, lower relative strengths resulted in a higher water vapour concentration *below* the inversion base, and thus at heights of the maximum temperature.

In contrast to Sect. 6.6.1, the temperature profile was modified in a second set of simulations to remove the effect of the vertical temperature distribution on DLR. Figure 6.21b shows ΔDLR for an isothermal temperature profile (273.15 K at all heights) where the same humidity inversion modifications have been performed as before. When using the isothermal temperature profile, positive ΔDLR are mostly eliminated. ΔDLR does not exceed $+0.1 \text{ W m}^{-2}$ and only occurs when the inversion is at the surface (Fig. 6.21). This finding supports that positive ΔDLR are mainly caused by the temperature inversion and less by the vertical water vapour distribution. However, similar to the original temperature profile, increasing the base height also led to a sign change of the relationship between inversion strength and ΔDLR from positive to negative (Fig. 6.21b). Thus, this behaviour is caused by the vertical water vapour distribution rather than the temperature profile.

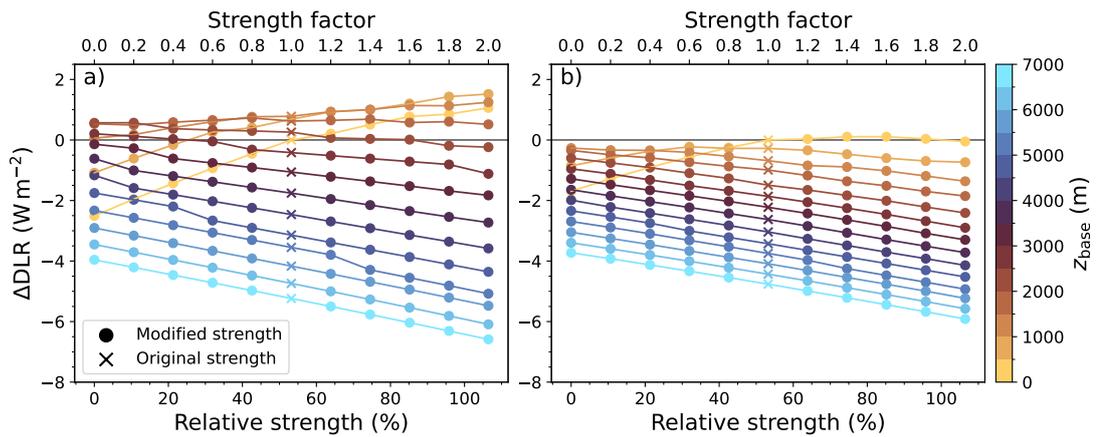


Figure 6.21: ΔDLR visualized against inversion strength modifications with IWV conservation using the original temperature profile (a) and an isothermal temperature profile with 273.15 K at all heights (b). The results shown in (a) are identical to Fig. 6.12a.

Chapter 7

Conclusions and Outlook

Accurate water vapour observations in the Arctic are challenging. Only a few ground stations exist at the coasts of the Arctic seas where high-quality water vapour observations are available. Arctic-wide coverage is only achieved by satellite remote sensing, which faces distinct challenges in the Arctic. Reanalyses, which assimilate parts of these observations, also disagree on water vapour estimates. As a result, the true amount of water vapour is often unclear in the data-sparse central Arctic due to significant uncertainties. Therefore, accurate and precise water vapour observations from field campaigns are crucial to identify deviations in satellite observations and numerical models in otherwise poorly sampled regions.

This thesis uses data from the largest and most comprehensive Arctic field campaign to date, the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition. We particularly focus on the measurements of two microwave radiometers (MWRs), the low-frequency HATPRO and the high-frequency MiRAC-P, installed on the German icebreaker RV *Polarstern*. A substantial part of this thesis was dedicated to creating high-quality water vapour products by synergizing the low- and high-frequency MWR measurements, exploiting their complementary moisture sensitivity.

With the highly temporally resolved integrated water vapour (IWV) retrieved from the MWRs and the radiosondes' vertically detailed specific humidity profiles, the performance of state-of-the-art satellite observations and numerical models could be assessed. The assessment also estimates the capability and limitations of ground-based MWRs and satellite-based remote sensing for monitoring Arctic humidity profiles, especially concerning the frequently occurring humidity inversions. Humidity inversions have important indirect effects on the downwelling longwave radiation (DLR) by influencing cloud formation and maintenance. However, the direct effect of humidity inversions on DLR, isolated from cloud or temperature influences, is unknown. We close this knowledge gap and thus im-

prove our understanding of the importance of humidity inversions using radiative transfer simulations.

The first challenge of this thesis was retrieving IWV and humidity profiles, as well as temperature profiles and liquid water path (LWP), from the MWR observations. Chapter 2 explained the physical basics of ground-based microwave remote sensing to understand how microwave radiances are sensitive to these parameters. Microwave observations only provide little information on vertical profiles but can be used in nearly all sky conditions. The derivation of atmospheric parameters is an inverse problem whose challenges are discussed in Chapter 3. In this thesis, the inverse problem is tackled using regression and Neural Network techniques. Statistical retrievals, such as regression and Neural Networks, are computationally much cheaper than physical approaches and are similarly accurate if carefully prepared. The main achievements are summarized in the following section. Afterwards, perspectives for future studies, which may benefit from the data sets generated with the MWR measurements, are presented.

7.1 Summary and conclusions

Study 1: Single microwave radiometer retrievals

The first study focused on the processing and quality control of the measurements of HATPRO and MiRAC-P during MOSAiC. Standardized regression retrievals were used to derive IWV, LWP, and temperature and humidity profiles from HATPRO. For the high-frequency MiRAC-P, regression approaches did not sufficiently capture the nonlinear relationship between the measurements and IWV. Thus, a Neural Network retrieval, specialized for the high-frequency observations of MiRAC-P, was developed. The retrieved LWP from HATPRO agrees well with the LWP from a two-channel MWR, which was also on board RV *Polarstern*. IWV and temperature and humidity profiles were evaluated with radiosonde observations from MOSAiC. The agreement of the temperature and humidity profiles is generally good and similar to previous studies, except for the lower troposphere. The frequent occurrence of temperature and humidity inversions increased the errors in the lower troposphere. Additionally, humidity profiling using only HATPRO observations proved to be challenging because of the dry conditions of the Arctic. The IWV comparison also showed that MiRAC-P's IWV retrieval is significantly more accurate and precise than HATPRO in dry conditions. In contrast, HATPRO excels in humid conditions where MiRAC-P's observations are partly saturated. The results showed that the large variability of humidity conditions, ranging from $\text{IWV} < 1$ to $> 30 \text{ kg m}^{-2}$, is challenging for the single-MWR

retrievals. Additionally, this study revealed the limitations of humidity profiling from HATPRO observations in the Arctic. The results motivated the synergy of the low-frequency HATPRO with the high-frequency MiRAC-P.

Study 2: Synergistic microwave radiometer retrievals

The goal of the second study was to create enhanced IWV and humidity profile retrievals by combining HATPRO and MiRAC-P observations and to quantify the benefit compared to the single-MWR retrievals. For each of the derived parameters, IWV and profiles of specific humidity temperature, a dedicated Neural Network retrieval was developed using the experience from the previous study. Tests with different input parameters in the Neural Network showed that including the 2 m temperature and the day of the year further improved the retrievals.

The Neural Network retrievals were found to be robust against random initialization by evaluating it for 20 runs with a synthetic evaluation data set based on the ERA5 reanalysis and simulated brightness temperatures. To quantify the benefit of the synergy compared to single-MWR retrievals, the errors of the IWV and specific humidity profiles with respect to radiosondes were compared. Additionally, the gain of the vertical information content was computed. The synergy of HATPRO and MiRAC-P achieves an excellent agreement with the radiosonde IWV, with root mean squared errors (RMSE) of only 2–3 % over almost the entire range of IWV conditions. In dry conditions, the synergy performs similarly well as MiRAC-P and even outperforms the HATPRO-only retrieval in humid conditions. The most significant improvement was found for specific humidity profiles. Low-altitude RMSE and biases were reduced by up to 50 and 75 %, respectively, and the vertical information content was strongly improved. The vertical information gain and the error reduction were most pronounced in dry conditions. The main achievements of the second study can be summarized by answering the first research question (RQ1):

RQ1: What is the benefit of combining low- and high-frequency MWRs with different moisture sensitivity for humidity profiling compared to more common single-MWR measurements?

The synergy provides highly accurate and precise IWV over a much larger range of IWV conditions than the single-MWR retrievals. Additionally, the RMSE was reduced by at least 15 %.

In the lower troposphere, the RMSE of specific humidity was improved by up to 50 % and the vertical resolution was enhanced by a factor of 2. The vertical information content was increased by 40 %.

Study 3: Evaluation of water vapour products and assessing the importance of humidity inversions

The third study assessed the quality of IWV and specific humidity profiles of state-of-the-art reanalyses (ERA5 and MERRA-2), weather forecast systems (ICON and CAFS) and satellite observations (combined sounding product using IASI, MHS and AMSU observations; hereafter referred to as IASI) in the Arctic during MOSAiC. The highly temporally resolved IWV from the MWR synergy and the radiosondes' vertically detailed specific humidity profiles were used as reference. All models generally show excellent agreement with the IWV from the MWR synergy, except for dry biases of 5–10 % in dry conditions for MERRA-2, ERA5 and ICON. IASI performs similarly well as the models in dry conditions but has a strong dry bias of 15 % in humid conditions. For the data sets with coarser temporal resolution, the errors could be partly attributed to intraweek IWV variability (e.g., due to storms).

Root mean squared deviations (RMSD) of specific humidity profiles with respect to radiosondes lie between 10 and 30 % in the lowest 4 km for ICON, CAFS, MERRA-2. In all seasons, the MWR synergy denotes similarly small or even smaller RMSD compared to these models. In winter, the MWR synergy has the smallest bias of all data sets. All models generally underestimate the specific humidity in the lower troposphere in the cold seasons. IASI overestimates the specific humidity near the surface and underestimates it at higher altitudes in all seasons. In summer, the negative bias is especially pronounced (consistent with the strong negative IWV bias). ERA5 has the smallest IWV and specific humidity profile errors but uses the most advanced assimilation technique (4D-var). Radiosondes launched during MOSAiC were assimilated by all models, except CAFS, and therefore influence the performance to an uncertain degree.

Most biases in the data sets can be related to humidity inversions, which occurred in 98 % of all radiosoundings during MOSAiC. A new humidity inversion detection method was developed to analyze the representation of humidity inversions, focusing more on the major inversions than previous methods. The presence of humidity inversions near or at the surface is well captured by the models in all seasons and by the MWR synergy in winter. Elevated inversions (base height > 1 km) are frequently missed by all data sets and not resolved at all by IASI and the MWR synergy. The comparison of the humidity inversion characteristics revealed that all data sets tend to underestimate the inversion strength and overestimate the inversion depth with respect to the radiosondes. The depth overestimation is slightly higher for data sets with a coarser vertical resolution (MWR, CAFS) than those with a finer vertical resolution (ICON, ERA5).

The results of the evaluation of water vapour products and the representation of humidity inversions answer RQ2 and RQ3:

RQ2: How well can ground-based MWRs capture the main features of the Arctic humidity profile, especially concerning humidity inversions? Can MWRs be used to correctly detect and monitor Arctic humidity inversions over long times in regions where radiosonde measurements are missing?

In winter, the ground-based MWRs excellently capture the presence and characteristics of surface-based and near-surface humidity inversions, leading to small specific humidity errors. In spring and autumn, the occurrence of inversions is slightly underestimated. In summer, the high occurrence of elevated inversions, which cannot be resolved, impairs the accurate monitoring of humidity inversions.

RQ3: Can the reference measurements from the MOSAiC expedition be used to identify biases in the humidity profile representation in satellite products and reanalyses? What are the capabilities of space-based remote sensing for humidity profiling as assessed by MOSAiC?

The comparison of the reanalysis data with the radiosonde observations from MOSAiC revealed overall small biases. The most notable biases are the dry biases of MERRA-2 and, to a lesser extent, ERA5 in the lower troposphere in the cold seasons. However, the results of the reanalyses are affected by the assimilation of the MOSAiC radiosondes and therefore represent an optimal performance. The errors are expected to be higher if the radiosondes were not assimilated. However, it is difficult to quantify the influence of the assimilation.

Humidity profiling proved particularly challenging from a satellite perspective (here, IASI) due to the near-permanent presence of humidity inversions. Surface-based inversions are frequently missed, leading to positive biases in the specific humidity profile near the surface and negative biases aloft. However, the specific humidity RMSD of IASI are mostly similar to the weather forecast models (ICON and CAFS) at low altitudes and smaller at high altitudes.

To assess the importance of the correct representation of humidity inversions, we first quantified the effect of humidity inversions on the DLR with radiative transfer simulations in clear sky conditions. The radiative effect of humidity inversions was found to range from 1 to 9 W m² in most cases but can also reach up to 16 W m². Surface-based inversions contribute most to the radiative effect

in winter, while elevated inversions dominate the radiative effect in the summer cases. However, the number of summer cases included in the analysis was small.

Subsequently, the specific humidity profiles of the different data sets were used as input in the radiative transfer model and compared to the DLR simulated with the radiosonde profiles. The largest DLR deviations were found for the specific humidity profiles of ICON, MERRA-2, CAFS and IASI, with several deviations of more than $\pm 5 \text{ W m}^2$. In contrast, the specific humidity profiles of ERA5 and the MWR synergy yield much lower DLR deviations where even outliers are within $\pm 3.5 \text{ W m}^2$. Thus, the specific humidity profiles from the MWR synergy are well suited to accurately compute DLR despite the coarse vertical resolution.

The results of the radiative sensitivity tests can be used to answer RQ4.

RQ4: How well do humidity inversions have to be represented to avoid biases in downwelling longwave radiation?

The radiative transfer simulations showed that it is most important to accurately capture the IWV and lower tropospheric specific humidity and, therefore, the near-surface humidity inversions. Elevated inversions are radiatively much less relevant in most cases, except for summer. The specific humidity profiles of the MWR were sufficiently accurate to avoid significant DLR biases.

Concluding remarks

This thesis has demonstrated that the synergy of low- and high-frequency MWRs is highly beneficial for Arctic water vapour assessments. The evaluation with radiosonde observations has shown that the IWV retrieval has a high accuracy over a wide range of conditions and that the retrieved specific humidity resolves the main humidity inversion well. Radiative transfer simulations enhanced the understanding of the radiative effect of humidity inversions in clear sky conditions. Near-surface humidity inversions have been found to dominate the longwave radiative effect in the cold season cases. Cloudy conditions have yet to be tested.

The evaluation of state-of-the-art models and satellite products with the reference observations from MOSAiC was essential to identify the quality of current water vapour products in a region where these data sets are known to be uncertain and where high-quality observations are missing. However, as the results are partly influenced by the assimilation of the MOSAiC radiosondes, they can be interpreted as representing the best possible performance of these data sets. An option to exclude campaign data in reanalyses would be beneficial for an independent and more representative evaluation. The identified biases reveal current limitations and help to improve the models and satellite observations. Future

reanalyses and improved satellite retrievals can be tested against the products analyzed in this thesis and evaluated with the MOSAiC observations.

As the ground-based MWR retrievals have been designed to be independent of measurements other than those from the two MWRs (the MWRs usually also include a weather station with air temperature measurements), and the retrievals are computationally cheap, they can be easily applied to other Arctic sites and field campaigns. Therefore, the retrievals are suitable to produce reference data beyond MOSAiC, which can be used to evaluate observations from recently launched and upcoming satellite missions (e.g., Arctic Weather Satellite) and reanalyses (e.g., ERA6). The ability of the MWRs to capture the surface-based humidity inversions in the cold seasons is promising for monitoring humidity profiles with a high temporal resolution at Arctic sites where no radiosondes are launched. Thus, the observations of the low- and high-frequency MWRs and the retrieval setup are valuable for future research on water vapour in the Arctic. The following section presents ideas to extend the current work.

7.2 Outlook

The highly temporally resolved MWR data can be used for a detailed analysis of the spatio-temporal water vapour variability in the Arctic. For example, during a strong warm air intrusion, the IWV can increase by a factor of 4 within a few hours (e.g., Crewell et al., 2021). The peak IWV, whose greenhouse effect can be critical for snow and ice melt onset, is often missed by observations and models with coarse temporal resolution (e.g., satellite observations, radiosondes). It may be quantified how much of the IWV variability is missed by different observations and models. The MOSAiC observations are valuable for investigating to what extent the missing IWV variability affects the surface energy budget through the greenhouse effect of water vapour.

Further, the water vapour products of the MWRs may be synergized with other observations (radiosondes, the Raman lidar Polly^{XT}, cloud radar, Met City) for detailed case studies on air mass transformation (e.g., during moist air intrusions). These combined observations with high temporal resolution are valuable to assess the impact of moist air intrusions on different atmospheric parameters (thermodynamics, cloud characteristics, precipitation, surface energy budget). The MWRs provide temporally detailed information on the atmospheric state in almost all sky conditions. The radar and the Raman lidar provide information on the cloud micro- and macrophysical properties and vertically detailed water vapour profiles below the cloud base, respectively. Such an observational analysis could be compared to the high-resolution model output created by Schnierstein

et al. (2024).

During the development of the synergistic retrieval, the idea came up to merge radiosonde and MWR observations for enhanced water vapour products. The IWV retrieved from the MWRs could be used to improve the linear interpolation of radiosonde observations between two launches. However, such an approach would not be able to capture the temporal evolution of humidity profiles correctly. Combining the radiosonde and MWR observations in a machine learning approach may be more appropriate for extracting information on the evolution of humidity profiles. This approach was only briefly tested and the results showed that the method generally worked. However, the information gain on the vertical profile between radiosonde launches seemed lower compared to the independent synergistic MWR retrievals. Additionally, such a synergistic approach would always require the presence of radiosonde observations. Nevertheless, the combination of radiosonde and MWR observations may yield an even better water vapour data set when carefully incorporated into machine learning frameworks.

To improve the knowledge of the spatial differences of the water vapour amount and trends at different Arctic sites during field campaigns or for continuous operation, low- and high-frequency MWRs (such as HATPRO and MiRAC-P) can provide useful insights. The MWRs can be deployed in regions where radiosonde observations are missing or only sparsely available. As Illingworth et al. (2019) states, ground-based MWRs improve weather forecasts (if assimilated) in regions where radiosonde observations are sparse by providing information on the thermodynamic state of the atmosphere. The operation of the MWRs can be automated, requires only little power supply, and is cheaper than launching radiosondes in the long-term perspective. To ensure high-quality observations, absolute calibrations are needed every 3–6 months (Küchler et al., 2016). As the retrievals developed in this thesis are computationally cheap, they are ideal for operational use.

Another advantage of the MWRs compared to the radiosondes is their high temporal resolution. Thus, low- and high-frequency MWR observations are also beneficial at Arctic sites where radiosondes are already launched regularly (e.g., Ny-Ålesund). The MWR observations complement the vertically detailed radiosondes by adding information on the temporal variability of the atmospheric state. MiRAC-P is expected to join HATPRO again at Ny-Ålesund in 2025 and will therefore improve the monitoring of humidity profiles and IWV, especially in the cold seasons. When HATPRO and MiRAC-P measure together at Ny-Ålesund, the performance of the retrievals developed in this thesis can be tested in the atmospheric conditions of an Arctic fjord.

While the MOSAiC expedition is the most extensive Arctic field campaign

so far, it cannot capture interannual variability. Thus, the long-term radiosonde measurements and observations from HATPRO and MiRAC-P (when available) at Ny-Ålesund or during further field campaigns are valuable to analyze the variability of water vapour over broader spatio-temporal scales. For example, measurements from radiosondes, as well as HATPRO and MiRAC-P during the *Polarstern* expedition PS131 (ATWAICE) provide further insights into the variability of water vapour in the marginal ice zone in summer. At Ny-Ålesund, the 30 years of radiosonde observations may be used to investigate interseasonal and interannual variability and trends of humidity inversion characteristics.

Recent and future satellite missions, such as the Arctic Weather Satellite, will further improve weather forecasts in the Arctic. The radiosonde observations and the retrieved products from HATPRO and MiRAC-P at Ny-Ålesund (in 2025) and onboard RV *Polarstern* during the currently ongoing cruise PS144 (ArcWatch 2) will be critical for the evaluation of the Arctic Weather Satellite observations. For the evaluation, the MWRs' high temporal resolution and good quality in almost all weather conditions are particularly important to maximize spatio-temporal overlaps with the satellite observations.

Bibliography

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <https://www.tensorflow.org/>, Accessed: 2024-08-19, 2015.
- Aggarwal, C. C.: Neural Networks and Deep Learning: A Textbook, Springer International Publishing, Cham, Switzerland, Second edn., ISBN 978-3-031-29642-0, doi:[10.1007/978-3-031-29642-0](https://doi.org/10.1007/978-3-031-29642-0), 2023.
- Aires, F., Prigent, C., Rossow, W. B., and Rothstein, M.: A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations, *J. Geophys. Res.: Atmos.*, 106, 14 887–14 907, doi:[10.1029/2001JD900085](https://doi.org/10.1029/2001JD900085), 2001.
- Alekseev, G., Kuzmina, S., Bobylev, L., Urazgildeeva, A., and Gnatiuk, N.: Impact of atmospheric heat and moisture transport on the Arctic warming, *Int. J. Climatol.*, 39, 3582–3592, doi:[10.1002/joc.6040](https://doi.org/10.1002/joc.6040), 2019.
- Alexeev, V. A., Langen, P. L., and Bates, J. R.: Polar amplification of surface warming on an aquaplanet in “ghost forcing” experiments without sea ice feedbacks, *Clim. Dyn.*, 24, 655–666, doi:[10.1007/s00382-005-0018-3](https://doi.org/10.1007/s00382-005-0018-3), 2005.
- Allan, R. P., Willett, K. M., John, V. O., and Trent, T.: Global Changes in Water Vapor 1979–2020, *J. Geophys. Res.: Atmos.*, 127, e2022JD036728, doi:[10.1029/2022JD036728](https://doi.org/10.1029/2022JD036728), 2022.
- Alraddawi, D., Sarkissian, A., Keckhut, P., Bock, O., Noël, S., Bekki, S., Irbah, A., Meftah, M., and Claud, C.: Comparison of total water vapour content

- in the Arctic derived from GNSS, AIRS, MODIS and SCIAMACHY, *Atmos. Meas. Tech.*, 11, 2949–2965, doi:[10.5194/amt-11-2949-2018](https://doi.org/10.5194/amt-11-2949-2018), 2018.
- Anderson, G. P., Clough, S. A., Kneizys, F. X., Chetwynd, J. H., and Shettle, E. P.: AFGL atmospheric constituent profiles (0 - 120 km), Tech. rep., Air Force Geophysics Laboratory (U.S.). Optical Physics Division, <https://apps.dtic.mil/sti/tr/pdf/ADA175173.pdf>, Accessed: 2024-07-13, 1986.
- Atmospheric Radiation Measurement (ARM) user facility: MWR Retrievals (MWRRET1LILJCLOU). 2019-10-11 to 2020-10-01, ARM Mobile Facility (MOS) MOSAIC (Drifting Obs - Study of Arctic Climate); AMF2 (M1), Compiled by D. Zhang. ARM Data Center, doi:[10.5439/1027369](https://doi.org/10.5439/1027369), 2019.
- August, T., Klaes, D., Schlüssel, P., Hultberg, T., Crapeau, M., Arriaga, A., O’Carroll, A., Coppens, D., Munro, R., and Calbet, X.: IASI on Metop-A: Operational Level 2 retrievals after five years in orbit, *J. Quant. Spectrosc. Radiat. Transfer*, 113, 1340–1371, doi:[10.1016/j.jqsrt.2012.02.028](https://doi.org/10.1016/j.jqsrt.2012.02.028), 2012.
- Avila-Diaz, A., Bromwich, D. H., Wilson, A. B., Justino, F., and Wang, S.-H.: Climate Extremes across the North American Arctic in Modern Reanalyses, *J. Climate*, 34, 2385–2410, doi:[10.1175/JCLI-D-20-0093.1](https://doi.org/10.1175/JCLI-D-20-0093.1), 2021.
- Barlakas, V., Deneke, H., and Macke, A.: The sub-adiabatic model as a concept for evaluating the representation and radiative effects of low-level clouds in a high-resolution atmospheric model, *Atmos. Chem. Phys.*, 20, 303–322, doi:[10.5194/acp-20-303-2020](https://doi.org/10.5194/acp-20-303-2020), 2020.
- Barrientos-Velasco, C., Deneke, H., Hünerbein, A., Griesche, H. J., Seifert, P., and Macke, A.: Radiative closure and cloud effects on the radiation budget based on satellite and shipborne observations during the Arctic summer research cruise, PS106, *Atmos. Chem. Phys.*, 22, 9313–9348, doi:[10.5194/acp-22-9313-2022](https://doi.org/10.5194/acp-22-9313-2022), 2022.
- Batrak, Y. and Müller, M.: On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice, *Nat. Commun.*, 10, 4170, doi:[10.1038/s41467-019-11975-3](https://doi.org/10.1038/s41467-019-11975-3), 2019.
- Bengtsson, L.: The global atmospheric water cycle, *Environ. Res. Lett.*, 5, 025 202, doi:[10.1088/1748-9326/5/2/025202](https://doi.org/10.1088/1748-9326/5/2/025202), 2010.
- Billault-Roux, A.-C. and Berne, A.: Integrated water vapor and liquid water path retrieval using a single-channel radiometer, *Atmos. Meas. Tech.*, 14, 2749–2769, doi:[10.5194/amt-14-2749-2021](https://doi.org/10.5194/amt-14-2749-2021), 2021.

- Bintanja, R., Graverson, R. G., and Hazeleger, W.: Arctic winter warming amplified by the thermal inversion and consequent low infrared cooling to space, *Nat. Geosci.*, 4, 758–761, doi:[10.1038/ngeo1285](https://doi.org/10.1038/ngeo1285), 2011.
- Blackport, R. and Fyfe, J. C.: Climate models fail to capture strengthening wintertime North Atlantic jet and impacts on Europe, *Sci. Adv.*, 8, eabn3112, doi:[10.1126/sciadv.abn3112](https://doi.org/10.1126/sciadv.abn3112), 2022.
- Blumstein, D., Chalon, G., Carlier, T., Buil, C., Hebert, P., Maciaszek, T., Ponce, G., Phulpin, T., Tournier, B., Simeoni, D., Astruc, P., Clauss, A., Kayal, G., and Jegou, R.: IASI instrument: technical overview and measured performances, in: *Infrared Spaceborne Remote Sensing XII*, edited by Strojnik, M., pp. 196 – 207, International Society for Optics and Photonics, SPIE, Denver, CO, doi:[10.1117/12.560907](https://doi.org/10.1117/12.560907), 2004.
- Bock, O., Bossler, P., Flamant, C., Doerflinger, E., Jansen, F., Fages, R., Bony, S., and Schnitt, S.: Integrated water vapour observations in the Caribbean arc from a network of ground-based GNSS receivers during EUREC4A, *Earth Syst. Sci. Data*, 13, 2407–2436, doi:[10.5194/essd-13-2407-2021](https://doi.org/10.5194/essd-13-2407-2021), 2021.
- Bresson, H., Rinke, A., Mech, M., Reinert, D., Schemann, V., Ebell, K., Maturilli, M., Viceto, C., Gorodetskaya, I., and Crewell, S.: Case study of a moisture intrusion over the Arctic with the ICOSahedral Non-hydrostatic (ICON) model: resolution dependence of its representation, *Atmos. Chem. Phys.*, 22, 173–196, doi:[10.5194/acp-22-173-2022](https://doi.org/10.5194/acp-22-173-2022), 2022.
- Brunke, M. A., Stegall, S. T., and Zeng, X.: A climatology of tropospheric humidity inversions in five reanalyses, *Atmos. Res.*, 153, 165–187, doi:[10.1016/j.atmosres.2014.08.005](https://doi.org/10.1016/j.atmosres.2014.08.005), 2015.
- Brümmer, B. and Pohlmann, S.: Wintertime roll and cell convection over Greenland and Barents Sea regions: A climatology, *J. Geophys. Res.: Atmos.*, 105, 15 559–15 566, doi:[10.1029/1999JD900841](https://doi.org/10.1029/1999JD900841), 2000.
- Cadeddu, M., Turner, D., and Liljegren, J.: A Neural Network for Real-Time Retrievals of PWV and LWP From Arctic Millimeter-Wave Ground-Based Observations, *IEEE Trans. Geosci. Remote Sens.*, 47, 1887–1900, doi:[10.1109/TGRS.2009.2013205](https://doi.org/10.1109/TGRS.2009.2013205), 2009.
- Cadeddu, M. P., Liljegren, J. C., and Pazmany, A. L.: Measurements and Retrievals From a New 183-GHz Water-Vapor Radiometer in the Arctic, *IEEE Trans. Geosci. Remote Sens.*, 45, 2207–2215, doi:[10.1109/TGRS.2006.888970](https://doi.org/10.1109/TGRS.2006.888970), 2007.

- Cadeddu, M. P., Cimini, D., Ghate, V., Lubin, D., Vogelmann, A. M., and Silber, I.: Examination of Humidity and Ice Supersaturation Profiles Over West Antarctica Using Ground-Based G-Band Radiometer Retrievals, *IEEE Trans. Geosci. Remote Sens.*, 60, 1–16, doi:[10.1109/TGRS.2021.3077088](https://doi.org/10.1109/TGRS.2021.3077088), 2022.
- Cattiaux, J., Peings, Y., Saint-Martin, D., Trou-Kechout, N., and Vavrus, S. J.: Sinuosity of midlatitude atmospheric flow in a warming world, *Geophys. Res. Lett.*, 43, 8259–8268, doi:[10.1002/2016GL070309](https://doi.org/10.1002/2016GL070309), 2016.
- Chandrasekhar, S.: Radiative Transfer, Dover Publications Inc., New York, USA, ISBN 978-0-486-60590-6, 1960.
- Chatterjee, D., Acquistapace, C., Deneke, H., and Crewell, S.: Understanding Cloud Systems’ Structure and Organization Using a Machine’s Self-Learning Approach, *Artif. Intell. Earth Syst.*, 2, e220 096, doi:[10.1175/AIES-D-22-0096.1](https://doi.org/10.1175/AIES-D-22-0096.1), 2023.
- Chellini, G. and Ebell, K.: Can state-of-the-art infrared satellite sounders and reanalyses detect moisture inversions in the Arctic?, *EGUsphere* [preprint], submitted to *Atmos. Meas. Tech.* on 2022-01-14, doi:[10.5194/amt-2022-22](https://doi.org/10.5194/amt-2022-22), 2022.
- Chen, B. and Liu, Z.: Global water vapor variability and trend from the latest 36 year (1979 to 2014) data of ECMWF and NCEP reanalyses, radiosonde, GPS, and microwave satellite, *J. Geophys. Res.: Atmos.*, 121, doi:[10.1002/2016JD024917](https://doi.org/10.1002/2016JD024917), 2016.
- Chollet, F. et al.: Keras, <https://keras.io>, Software, Accessed: 2024-08-24, 2015.
- Cimini, D., Westwater, E. R., and Gasiewski, A. J.: Temperature and Humidity Profiling in the Arctic Using Ground-Based Millimeter-Wave Radiometry and 1DVAR, *IEEE Trans. Geosci. Remote Sens.*, 48, 1381–1388, doi:[10.1109/TGRS.2009.2030500](https://doi.org/10.1109/TGRS.2009.2030500), 2010.
- Clark, J. P., Clothiaux, E. E., Feldstein, S. B., and Lee, S.: Drivers of Global Clear Sky Surface Downwelling Longwave Irradiance Trends From 1984 to 2017, *Geophys. Res. Lett.*, 48, e2021GL093 961, doi:[10.1029/2021GL093961](https://doi.org/10.1029/2021GL093961), 2021.
- Colman, R. A.: On the vertical extent of atmospheric feedbacks, *Clim. Dyn.*, 17, 391–405, doi:[10.1007/s003820000111](https://doi.org/10.1007/s003820000111), 2001.
- Connell, B. H. and Miller, D. R.: An Interpretation of Radiosonde Errors in the Atmospheric Boundary Layer, *J. Appl. Meteorol.*, 34, 1070–1081, doi:[10.1175/1520-0450\(1995\)034<1070:AIOREI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1995)034<1070:AIOREI>2.0.CO;2), 1995.

- Cox, C., Gallagher, M., Shupe, M., Persson, O., Blomquist, B., Grachev, A., Riihimäki, L., Kutchenreiter, M., Morris, V., Solomon, A., Brooks, I., Costa, D., Gattas, D., Hutchings, J., Osborn, J., Morris, S., Preusser, A., and Uttal, T.: Met City meteorological and surface flux measurements (Level 3 Final), Multi-disciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC), central Arctic, October 2019 - September 2020, NSF Arctic Data Center [data set], doi:[10.18739/A2PV6B83F](https://doi.org/10.18739/A2PV6B83F), 2023a.
- Cox, C. J., Morris, S. M., Uttal, T., Burgener, R., Hall, E., Kutchenreiter, M., McComiskey, A., Long, C. N., Thomas, B. D., and Wendell, J.: The De-Icing Comparison Experiment (D-ICE): a study of broadband radiometric measurements under icing conditions in the Arctic, *Atmos. Meas. Tech.*, 14, 1205–1224, doi:[10.5194/amt-14-1205-2021](https://doi.org/10.5194/amt-14-1205-2021), 2021.
- Cox, C. J., Gallagher, M. R., Shupe, M. D., Persson, P. O. G., Solomon, A., Fairall, C. W., Ayers, T., Blomquist, B., Brooks, I. M., Costa, D., Grachev, A., Gattas, D., Hutchings, J. K., Kutchenreiter, M., Leach, J., Morris, S. M., Morris, V., Osborn, J., Pezoa, S., Preußner, A., Riihimäki, L. D., and Uttal, T.: Continuous observations of the surface energy budget and meteorology over the Arctic sea ice during MOSAIC, *Sci. Data*, 10, 519, doi:[10.1038/s41597-023-02415-5](https://doi.org/10.1038/s41597-023-02415-5), 2023b.
- Crasemann, B., Handorf, D., Jaiser, R., Dethloff, K., Nakamura, T., Ukita, J., and Yamazaki, K.: Can preferred atmospheric circulation patterns over the North-Atlantic-Eurasian region be associated with arctic sea ice loss?, *Polar Sci.*, 14, 9–20, doi:[10.1016/j.polar.2017.09.002](https://doi.org/10.1016/j.polar.2017.09.002), 2017.
- Crewell, S. and Löhnert, U.: Accuracy of Boundary Layer Temperature Profiles Retrieved With Multifrequency Multiangle Microwave Radiometry, *IEEE Trans. Geosci. Remote Sens.*, 45, 2195–2201, doi:[10.1109/TGRS.2006.888434](https://doi.org/10.1109/TGRS.2006.888434), 2007.
- Crewell, S., Czekala, H., Löhnert, U., Simmer, C., Rose, T., Zimmermann, R., and Zimmermann, R.: Microwave Radiometer for Cloud Cartography: A 22-channel ground-based microwave radiometer for atmospheric research, *Radio Sci.*, 36, 621–638, doi:[10.1029/2000RS002396](https://doi.org/10.1029/2000RS002396), 2001.
- Crewell, S., Ebell, K., Konjari, P., Mech, M., Nomokonova, T., Radovan, A., Strack, D., Triana-Gómez, A. M., Noël, S., Scarlat, R., Spreen, G., Maturilli, M., Rinke, A., Gorodetskaya, I., Viceto, C., August, T., and Schröder, M.: A systematic assessment of water vapor products in the Arctic: from instant-

- neous measurements to monthly means, *Atmos. Meas. Tech.*, 14, 4829–4856, doi:10.5194/amt-14-4829-2021, 2021.
- Curry, J.: On the Formation of Continental Polar Air, *J. Atmos. Sci.*, 40, 2278–2292, doi:10.1175/1520-0469(1983)040<2278:OTFOCP>2.0.CO;2, 1983.
- Curry, J. A., Schramm, J. L., Serreze, M. C., and Ebert, E. E.: Water vapor feedback over the Arctic Ocean, *J. Geophys. Res.: Atmos.*, 100, 14 223–14 229, doi:10.1029/95JD00824, 1995.
- Dahlke, S. and Maturilli, M.: Contribution of Atmospheric Advection to the Amplified Winter Warming in the Arctic North Atlantic Region, *Adv. Meteorol.*, 2017, 1–8, doi:10.1155/2017/4928620, 2017.
- Dahlke, S., Shupe, M. D., Cox, C. J., Brooks, I. M., Blomquist, B., and Persson, P. O. G.: Extended radiosonde profiles 2019/09-2020/10 during MOSAiC Legs PS122/1 - PS122/5, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:10.1594/PANGAEA.961881, 2023.
- Dai, G., Althausen, D., Hofer, J., Engelmann, R., Seifert, P., Bühl, J., Mamouri, R.-E., Wu, S., and Ansmann, A.: Calibration of Raman lidar water vapor profiles by means of AERONET photometer observations and GDAS meteorological data, *Atmos. Meas. Tech.*, 11, 2735–2748, doi:10.5194/amt-11-2735-2018, 2018.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Devasthale, A., Willén, U., Karlsson, K.-G., and Jones, C. G.: Quantifying the clear-sky temperature inversion frequency and strength over the Arctic Ocean during summer and winter seasons from AIRS profiles, *Atmos. Chem. Phys.*, 10, 5565–5572, doi:10.5194/acp-10-5565-2010, 2010.
- Devasthale, A., Sedlar, J., and Tjernström, M.: Characteristics of water-vapour inversions observed over the Arctic by Atmospheric Infrared Sounder (AIRS)

- and radiosondes, *Atmos. Chem. Phys.*, 11, 9813–9823, doi:[10.5194/acp-11-9813-2011](https://doi.org/10.5194/acp-11-9813-2011), 2011.
- Devasthale, A., Sedlar, J., Kahn, B. H., Tjernström, M., Fetzer, E. J., Tian, B., Teixeira, J., and Pagano, T. S.: A Decade of Spaceborne Observations of the Arctic Atmosphere: Novel Insights from NASA’s AIRS Instrument, *Bull. Am. Meteorol. Soc.*, 97, 2163–2176, doi:[10.1175/BAMS-D-14-00202.1](https://doi.org/10.1175/BAMS-D-14-00202.1), 2016.
- Di Biagio, C., Pelon, J., Blanchard, Y., Loyer, L., Hudson, S. R., Walden, V. P., Raut, J. C., Kato, S., Mariage, V., and Granskog, M. A.: Toward a Better Surface Radiation Budget Analysis Over Sea Ice in the High Arctic Ocean: A Comparative Study Between Satellite, Reanalysis, and local-scale Observations, *J. Geophys. Res.: Atmos.*, 126, doi:[10.1029/2020JD032555](https://doi.org/10.1029/2020JD032555), 2021.
- Diaz, J. C. F., Carter, W. E., Shrestha, R. L., and Glennie, C. L.: LiDAR Remote Sensing, in: *Handbook of Satellite Applications*, edited by Pelton, J. N., Madry, S., and Camacho-Lara, S., pp. 929–980, Springer International Publishing, Cham, ISBN 978-3-319-23386-4, doi:[10.1007/978-3-319-23386-4_44](https://doi.org/10.1007/978-3-319-23386-4_44), 2017.
- Dicke, R. H., Beringer, R., Kyhl, R. L., and Vane, A. B.: Atmospheric Absorption Measurements with a Microwave Radiometer, *Phys. Rev.*, 70, 340–348, doi:[10.1103/PhysRev.70.340](https://doi.org/10.1103/PhysRev.70.340), 1946.
- Doyle, J. G., Lesins, G., Thackray, C. P., Perro, C., Nott, G. J., Duck, T. J., Damoah, R., and Drummond, J. R.: Water vapor intrusions into the High Arctic during winter, *Geophys. Res. Lett.*, 38, doi:[10.1029/2011GL047493](https://doi.org/10.1029/2011GL047493), 2011.
- Dutton, G., Hall, B., Dlugokencky, E., Lan, X., Nance, J., and Madronich, M.: Combined Atmospheric Nitrous Oxide Dry Air Mole Fractions from the NOAA GML Halocarbons Sampling Network, 1977-2023, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/GMZ7-2Q16](https://doi.org/10.15138/GMZ7-2Q16), Version: 2023-04-13, 2023a.
- Dutton, G., Hall, B., Montzka, S., and Nance, J.: Combined Atmospheric Carbon tetrachloride Dry Air Mole Fractions from the NOAA GML Halocarbons Sampling Network, 1995-2023, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/CV0A-J604](https://doi.org/10.15138/CV0A-J604), Version: 2023-04-13, 2023b.
- Dutton, G., Hall, B., Montzka, S., and Nance, J.: Combined Atmospheric Chlorofluorocarbon-11 Dry Air Mole Fractions from the NOAA GML Halocarbons Sampling Network, 1977-2023, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/BVQ6-2S69](https://doi.org/10.15138/BVQ6-2S69), Version: 2023-04-13, 2023c.

- Dutton, G., Hall, B., Montzka, S., and Nance, J.: Combined Atmospheric Chlorofluorocarbon-12 Dry Air Mole Fractions from the NOAA GML Halocarbon Sampling Network, 1977-2023, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/PJ63-H440](https://doi.org/10.15138/PJ63-H440), Version: 2023-04-13, 2023d.
- Ebell, K.: Retrieval coefficients for HATPRO observations during MOSAiC, Zenodo [data set], doi:[10.5281/zenodo.6673886](https://doi.org/10.5281/zenodo.6673886), 2022.
- Ebell, K. and Walbröl, A.: Training and test data for retrievals based on HATPRO observations during MOSAiC, Zenodo [data set], doi:[10.5281/zenodo.5741350](https://doi.org/10.5281/zenodo.5741350), 2021.
- Ebell, K., Orlandi, E., Hünerbein, A., Löhnert, U., and Crewell, S.: Combining ground-based with satellite-based measurements in the atmospheric state retrieval: Assessment of the information content, *J. Geophys. Res.: Atmos.*, 118, 6940–6956, doi:[10.1002/jgrd.50548](https://doi.org/10.1002/jgrd.50548), 2013.
- Ebell, K., Löhnert, U., Päschke, E., Orlandi, E., Schween, J. H., and Crewell, S.: A 1-D variational retrieval of temperature, humidity, and liquid cloud properties: Performance under idealized and real conditions, *J. Geophys. Res.: Atmos.*, 122, 1746–1766, doi:[10.1002/2016JD025945](https://doi.org/10.1002/2016JD025945), 2017.
- Ebell, K., Nomokonova, T., Maturilli, M., and Ritter, C.: Radiative Effect of Clouds at Ny-Ålesund, Svalbard, as Inferred from Ground-Based Remote Sensing Observations, *J. Appl. Meteorol. Climatol.*, 59, 3–22, doi:[10.1175/JAMC-D-19-0080.1](https://doi.org/10.1175/JAMC-D-19-0080.1), 2020.
- Ebell, K., Walbröl, A., Engelmann, R., Griesche, H., Radenz, M., Hofer, J., and Althausen, D.: Temperature and humidity profiles, integrated water vapour and liquid water path derived from the HATPRO microwave radiometer onboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.941389](https://doi.org/10.1594/PANGAEA.941389), 2022.
- Ellison, W. J.: Freshwater and sea water, in: *Thermal Microwave Radiation: Applications for Remote Sensing*, edited by Mätzler, C., vol. 52, pp. 431–455, The Institution of Engineering and Technology, London, UK, ISBN 978-086341-573-9, doi:[10.1049/PBEW052E](https://doi.org/10.1049/PBEW052E), 2006.
- Engelmann, R., Kanitz, T., Baars, H., Heese, B., Althausen, D., Skupin, A., Wandinger, U., Komppula, M., Stachlewska, I. S., Amiridis, V., Marinou, E.,

- Mattis, I., Linné, H., and Ansmann, A.: The automated multiwavelength Raman polarization and water-vapor lidar PollyXT: the neXT generation, *Atmos. Meas. Tech.*, 9, 1767–1784, doi:[10.5194/amt-9-1767-2016](https://doi.org/10.5194/amt-9-1767-2016), 2016.
- Engelmann, R., Ansmann, A., Ohneiser, K., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Dahlke, S., Maturilli, M., Veselovskii, I., Jimenez, C., Wiesen, R., Baars, H., Bühl, J., Gebauer, H., Haarig, M., Seifert, P., Wandinger, U., and Macke, A.: Wildfire smoke, Arctic haze, and aerosol effects on mixed-phase and cirrus clouds over the North Pole region during MOSAiC: an introduction, *Atmos. Chem. Phys.*, 21, 13 397–13 423, doi:[10.5194/acp-21-13397-2021](https://doi.org/10.5194/acp-21-13397-2021), 2021.
- Engelmann, R., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Walbröl, A., and Ebell, K.: Brightness temperatures of the HATPRO microwave radiometer onboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.941356](https://doi.org/10.1594/PANGAEA.941356), 2022.
- Engelmann, R., Althausen, D., Baars, H., Griesche, H., Hofer, J., Radenz, M., and Seifert, P.: Custom collection of classification data from RV Polarstern between 11 Oct 2019 and 30 Sep 2020, ACTRIS Cloudnet data portal [data set], doi:[10.60656/60EA0DD0A99746BA](https://doi.org/10.60656/60EA0DD0A99746BA), 2023.
- Engelmann, R., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Seidel, C., Baars, H., Klamt, A., and Hengst, R.: Water vapor profiles retrieved during the POLARSTERN cruises PS122/1, PS122/2 and PS122/3, 28 September 2019 – 29 February 2020, Arctic Ocean, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.957158](https://doi.org/10.1594/PANGAEA.957158), 2023.
- EUMETSAT: IASI Level 2: Product Generation Specification, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), https://www-cdn.eumetsat.int/files/2020-12/IASI%20Level%20_%20Product%20Generation%20Specification.pdf, EPS.SYS.SPE.990013, Accessed: 2024-08-28, 2017.
- EUMETSAT: IASI Combined Sounding Products - Metop, EUMETSAT Data Store [data set], <https://navigator.eumetsat.int/product/EO:EUM:DAT:METOP:IASSND02>, Accessed: 2024-05-25, 2024.
- Feldl, N., Po-Chedley, S., Singh, H. K. A., Hay, S., and Kushner, P. J.: Sea ice and atmospheric circulation shape the high-latitude lapse rate feedback, *npj Clim. Atmos. Sci.*, 3, 41, doi:[10.1038/s41612-020-00146-7](https://doi.org/10.1038/s41612-020-00146-7), 2020.

- Fionda, E., Cadeddu, M., Mattioli, V., and Pacione, R.: Intercomparison of Integrated Water Vapor Measurements at High Latitudes from Co-Located and Near-Located Instruments, *Remote Sens.*, 11, 2130, doi:[10.3390/rs11182130](https://doi.org/10.3390/rs11182130), 2019.
- Foken, T., ed.: Springer Handbook of Atmospheric Measurements, Springer Handbooks, Springer International Publishing, Cham, Switzerland, ISBN 978-3-030-52171-4, doi:[10.1007/978-3-030-52171-4](https://doi.org/10.1007/978-3-030-52171-4), 2021.
- Francis, J. A. and Vavrus, S. J.: Evidence for a wavier jet stream in response to rapid Arctic warming, *Environ. Res. Lett.*, 10, 014005, doi:[10.1088/1748-9326/10/1/014005](https://doi.org/10.1088/1748-9326/10/1/014005), 2015.
- Francis, J. A., White, D. M., Cassano, J. J., Gutowski, W. J., Hinzman, L. D., Holland, M. M., Steele, M. A., and Vörösmarty, C. J.: An arctic hydrologic system in transition: Feedbacks and impacts on terrestrial, marine, and human life, *J. Geophys. Res.*, 114, G04019, doi:[10.1029/2008JG000902](https://doi.org/10.1029/2008JG000902), 2009.
- Frank, H.: MMDFs for the DWD-ICON global forecast model for various Arctic sites, Norwegian Meteorological Institute [data set], doi:[10.21343/09KM-BJ07](https://doi.org/10.21343/09KM-BJ07), 2023.
- Ganeshan, M. and Yang, Y.: Evaluation of the Antarctic Boundary Layer Thermodynamic Structure in MERRA2 Using Dropsonde Observations from the Concordiasi Campaign, *Earth Space Sci.*, 6, 2397–2409, doi:[10.1029/2019EA000890](https://doi.org/10.1029/2019EA000890), 2019.
- Gao, B.-C. and Kaufman, Y. J.: Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels, *J. Geophys. Res.: Atmos.*, 108, 2002JD003023, doi:[10.1029/2002JD003023](https://doi.org/10.1029/2002JD003023), 2003.
- Gaustad, K., Turner, D., and McFarlane, S.: MWRRET Value-Added Product: The Retrieval of Liquid Water Path and Precipitable Water Vapor from Microwave Radiometer (MWR) Data Sets (Revision 2), Tech. Rep. DOE/SC-ARM/TR-081.2, 1019284, Department Of Energy - Office of Science Atmospheric Radiation Measurement (ARM) Program (United States), doi:[10.2172/1019284](https://doi.org/10.2172/1019284), 2011.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D.,

- Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *J. Climate*, 30, 5419–5454, doi:[10.1175/JCLI-D-16-0758.1](https://doi.org/10.1175/JCLI-D-16-0758.1), 2017.
- Ghatak, D. and Miller, J.: Implications for Arctic amplification of changes in the strength of the water vapor feedback, *J. Geophys. Res.: Atmos.*, 118, 7569–7578, doi:[10.1002/jgrd.50578](https://doi.org/10.1002/jgrd.50578), 2013.
- Gierens, R., Kneifel, S., Shupe, M. D., Ebell, K., Maturilli, M., and Löhnert, U.: Low-level mixed-phase clouds in a complex Arctic environment, *Atmos. Chem. Phys.*, 20, 3459–3481, doi:[10.5194/acp-20-3459-2020](https://doi.org/10.5194/acp-20-3459-2020), 2020.
- Graham, R. M., Cohen, L., Ritzhaupt, N., Segger, B., Graversen, R. G., Rinke, A., Walden, V. P., Granskog, M. A., and Hudson, S. R.: Evaluation of Six Atmospheric Reanalyses over Arctic Sea Ice from Winter to Early Summer, *J. Climate*, 32, 4121–4143, doi:[10.1175/JCLI-D-18-0643.1](https://doi.org/10.1175/JCLI-D-18-0643.1), 2019a.
- Graham, R. M., Hudson, S. R., and Maturilli, M.: Improved Performance of ERA5 in Arctic Gateway Relative to Four Global Atmospheric Reanalyses, *Geophys. Res. Lett.*, 46, 6138–6147, doi:[10.1029/2019GL082781](https://doi.org/10.1029/2019GL082781), 2019b.
- Graham, R. M., Itkin, P., Meyer, A., Sundfjord, A., Spreen, G., Smedsrud, L. H., Liston, G. E., Cheng, B., Cohen, L., Divine, D., Fer, I., Fransson, A., Gerland, S., Haapala, J., Hudson, S. R., Johansson, A. M., King, J., Merkouriadi, I., Peterson, A. K., Provost, C., Randelhoff, A., Rinke, A., Rösel, A., Sennéchaël, N., Walden, V. P., Duarte, P., Assmy, P., Steen, H., and Granskog, M. A.: Winter storms accelerate the demise of sea ice in the Atlantic sector of the Arctic Ocean, *Sci. Rep.*, 9, 9222, doi:[10.1038/s41598-019-45574-5](https://doi.org/10.1038/s41598-019-45574-5), 2019c.
- Graversen, R. G. and Wang, M.: Polar amplification in a coupled climate model with locked albedo, *Clim. Dyn.*, 33, 629–643, doi:[10.1007/s00382-009-0535-6](https://doi.org/10.1007/s00382-009-0535-6), 2009.
- Graversen, R. G., Mauritsen, T., Tjernström, M., Källén, E., and Svensson, G.: Vertical structure of recent Arctic warming, *Nature*, 451, 53–56, doi:[10.1038/nature06502](https://doi.org/10.1038/nature06502), 2008.
- Griesche, H., Seifert, P., Engelmann, R., Radenz, M., Hofer, J., and Althausen, D.: Low-level stratus mask from Polarstern during MOSAiC, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.961789](https://doi.org/10.1594/PANGAEA.961789), 2023.
- Griesche, H. J. and Seifert, P.: MOSAiC Cloudnet issue data set, Zenodo [data set], doi:[10.5281/ZENODO.7310858](https://doi.org/10.5281/ZENODO.7310858), 2023.

- Griesche, H. J., Seifert, P., Ansmann, A., Baars, H., Barrientos Velasco, C., Bühl, J., Engelmann, R., Radenz, M., Zhenping, Y., and Macke, A.: Application of the shipborne remote sensing supersite OCEANET for profiling of Arctic aerosols and clouds during Polarstern cruise PS106, *Atmos. Meas. Tech.*, 13, 5335–5358, doi:[10.5194/amt-13-5335-2020](https://doi.org/10.5194/amt-13-5335-2020), 2020.
- Griesche, H. J., Barrientos-Velasco, C., Deneke, H., Hünerbein, A., Seifert, P., and Macke, A.: Low-level Arctic clouds: a blind zone in our knowledge of the radiation budget, *Atmos. Chem. Phys.*, 24, 597–612, doi:[10.5194/acp-24-597-2024](https://doi.org/10.5194/acp-24-597-2024), 2024a.
- Griesche, H. J., Seifert, P., Engelmann, R., Radenz, M., Hofer, J., Althausen, D., Walbröl, A., Barrientos-Velasco, C., Baars, H., Dahlke, S., Tukiainen, S., and Macke, A.: Cloud micro- and macrophysical properties from ground-based remote sensing during the MOSAiC drift experiment, *Sci. Data*, 11, 505, doi:[10.1038/s41597-024-03325-w](https://doi.org/10.1038/s41597-024-03325-w), 2024b.
- Haas, C.: Links to master tracks in different resolutions of POLARSTERN cruise PS122/2, Arctic Ocean - Arctic Ocean, 2019-12-13 - 2020-02-24 (Version 2), PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.924674](https://doi.org/10.1594/PANGAEA.924674), 2020.
- Hahn, L. C., Armour, K. C., Zelinka, M. D., Bitz, C. M., and Donohoe, A.: Contributions to Polar Amplification in CMIP5 and CMIP6 Models, *Front. Earth Sci.*, 9, 710 036, doi:[10.3389/feart.2021.710036](https://doi.org/10.3389/feart.2021.710036), 2021.
- Haykin, S., ed.: *Neural Networks: A comprehensive foundation*, Pearson Education, Patparganj, Delhi, India, Second edn., ISBN 978-0139083853, 1998.
- Held, I. M. and Soden, B. J.: Water Vapor Feedback and Global Warming, *Annu. Rev. Energy Env.*, 25, 441–475, doi:[10.1146/annurev.energy.25.1.441](https://doi.org/10.1146/annurev.energy.25.1.441), 2000.
- Herrmannsdörfer, L., Müller, M., Shupe, M. D., and Rostosky, P.: Surface temperature comparison of the Arctic winter MOSAiC observations, ERA5 reanalysis, and MODIS satellite retrieval, *Elem. Sci. Anth.*, 11, 00 085, doi:[10.1525/elementa.2022.00085](https://doi.org/10.1525/elementa.2022.00085), 2023.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez,

- P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, *Q. J. R. Meteorolog. Soc.*, 146, 1999–2049, doi:[10.1002/qj.3803](https://doi.org/10.1002/qj.3803), 2020.
- Hong, G., Yang, P., Baum, B. A., Heymsfield, A. J., Weng, F., Liu, Q., Heygster, G., and Buehler, S. A.: Scattering database in the millimeter and submillimeter wave range of 100–1000 GHz for nonspherical ice particles, *J. Geophys. Res.*, 114, D06 201, doi:[10.1029/2008JD010451](https://doi.org/10.1029/2008JD010451), 2009.
- Hong, Y. and Liu, G.: The Characteristics of Ice Cloud Properties Derived from CloudSat and CALIPSO Measurements, *J. Climate*, 28, 3880–3901, doi:[10.1175/JCLI-D-14-00666.1](https://doi.org/10.1175/JCLI-D-14-00666.1), 2015.
- Hyland, R. W. and Wexler, A.: Formulations for the thermodynamic properties of the saturated phases of H₂O from 173.15 to 473.15 K, *Ashrae Transactions*, 89, 500–520, <https://api.semanticscholar.org/CorpusID:100743565>, Accessed: 2024-08-25, 1983.
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.: Atmos.*, 113, 2008JD009 944, doi:[10.1029/2008JD009944](https://doi.org/10.1029/2008JD009944), 2008.
- Illingworth, A. J., Hogan, R. J., O’Connor, E., Bouniol, D., Brooks, M. E., Delanoé, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W. F., Haeffelin, M., Baltink, H. K., Krasnov, O. A., Pelon, J., Piriou, J.-M., Protat, A., Russchenberg, H. W. J., Seifert, A., Tompkins, A. M., van Zadelhoff, G.-J., Vinit, F., Willén, U., Wilson, D. R., and Wrench, C. L.: Cloudnet, *Bull. Am. Meteorol. Soc.*, 88, 883–898, doi:[10.1175/BAMS-88-6-883](https://doi.org/10.1175/BAMS-88-6-883), 2007.
- Illingworth, A. J., Cimini, D., Haefele, A., Haeffelin, M., Hervo, M., Kotthaus, S., Löhnert, U., Martinet, P., Mattis, I., O’Connor, E. J., and Potthast, R.: How Can Existing Ground-Based Profiling Instruments Improve European Weather Forecasts?, *Bull. Am. Meteorol. Soc.*, 100, 605–619, doi:[10.1175/BAMS-D-17-0231.1](https://doi.org/10.1175/BAMS-D-17-0231.1), 2019.
- Inamdar, A. K., Ramanathan, V., and Loeb, N. G.: Satellite observations of the water vapor greenhouse effect and column longwave cooling rates: Relative roles of the continuum and vibration-rotation to pure rotation bands, *J. Geophys. Res.: Atmos.*, 109, 2003JD003 980, doi:[10.1029/2003JD003980](https://doi.org/10.1029/2003JD003980), 2004.
- Intergovernmental Panel on Climate Change: The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate

- Change, Cambridge University Press, 1 edn., ISBN 978-1-00-915796-4 978-1-00-915797-1, doi:[10.1017/9781009157964](https://doi.org/10.1017/9781009157964), 2022.
- Intergovernmental Panel on Climate Change: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1 edn., ISBN 978-1-00-915789-6, doi:[10.1017/9781009157896](https://doi.org/10.1017/9781009157896), 2023.
- Janssen, M. A., ed.: Atmospheric remote sensing by microwave radiometry, Wiley series in remote sensing, Wiley, New York, ISBN 978-0-471-62891-0, 1993.
- Johansson, E., Devasthale, A., Tjernström, M., Ekman, A. M. L., and L'Ecuyer, T.: Response of the lower troposphere to moisture intrusions into the Arctic, *Geophys. Res. Lett.*, 44, 2527–2536, doi:[10.1002/2017GL072687](https://doi.org/10.1002/2017GL072687), 2017.
- Jozef, G. C., Cassano, J. J., Dahlke, S., Dice, M., Cox, C. J., and De Boer, G.: An overview of the vertical structure of the atmospheric boundary layer in the central Arctic during MOSAiC, *Atmos. Chem. Phys.*, 24, 1429–1450, doi:[10.5194/acp-24-1429-2024](https://doi.org/10.5194/acp-24-1429-2024), 2024.
- Kanitz, T., Seifert, P., Ansmann, A., Engelmann, R., Althausen, D., Casiccia, C., and Rohwer, E. G.: Contrasting the impact of aerosols at northern and southern midlatitudes on heterogeneous ice formation, *Geophys. Res. Lett.*, 38, L17 802, doi:[10.1029/2011GL048532](https://doi.org/10.1029/2011GL048532), 2011.
- Kanitz, T., Ansmann, A., Engelmann, R., and Althausen, D.: North-south cross sections of the vertical aerosol distribution over the Atlantic Ocean from multiwavelength Raman/polarization lidar during Polarstern cruises, *J. Geophys. Res.: Atmos.*, 118, 2643–2655, doi:[10.1002/jgrd.50273](https://doi.org/10.1002/jgrd.50273), 2013.
- Kanzow, T.: Links to master tracks in different resolutions of POLARSTERN cruise PS122/3, Arctic Ocean - Longyearbyen, 2020-02-24 - 2020-06-04 (Version 2), PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.924681](https://doi.org/10.1594/PANGAEA.924681), 2020.
- Kapsch, M.-L., Graverson, R. G., and Tjernström, M.: Springtime atmospheric energy transport and the control of Arctic summer sea-ice extent, *Nat. Clim. Change*, 3, 744–748, doi:[10.1038/nclimate1884](https://doi.org/10.1038/nclimate1884), 2013.
- Karstens, U., Simmer, C., and Ruprecht, E.: Remote sensing of cloud liquid water, *Meteorol. Atmos. Phys.*, 54, 157–171, doi:[10.1007/BF01030057](https://doi.org/10.1007/BF01030057), 1994.
- Kerber, F., Querel, R. R., Rondanelli, R., Hanuschik, R., van den Ancker, M., Cuevas, O., Smette, A., Smoker, J., Rose, T., and Czekala, H.: An episode of

- extremely low precipitable water vapour over Paranal observatory, *Mon. Not. R. Astron. Soc.*, 439, 247–255, doi:[10.1093/mnras/stt2404](https://doi.org/10.1093/mnras/stt2404), 2014.
- Kiehl, J. T. and Trenberth, K. E.: Earth’s Annual Global Mean Energy Budget, *Bull. Am. Meteorol. Soc.*, 78, 197–208, doi:[10.1175/1520-0477\(1997\)078<0197:EAGMEB>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2), 1997.
- Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv, doi:[10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980), 2017.
- Kirbus, B., Tiedeck, S., Camplani, A., Chylik, J., Crewell, S., Dahlke, S., Ebell, K., Gorodetskaya, I., Griesche, H., Handorf, D., Höschel, I., Lauer, M., Neggers, R., Rückert, J., Shupe, M. D., Spreen, G., Walbröl, A., Wendisch, M., and Rinke, A.: Surface impacts and associated mechanisms of a moisture intrusion into the Arctic observed in mid-April 2020 during MOSAiC, *Front. Earth Sci.*, 11, 1147848, doi:[10.3389/feart.2023.1147848](https://doi.org/10.3389/feart.2023.1147848), 2023.
- Kirchhoff, G.: Ueber das Verhältniss zwischen dem Emissionsvermögen und dem Absorptionsvermögen der Körper für Wärme und Licht, *Annalen der Physik*, 185, 275–301, doi:[10.1002/andp.18601850205](https://doi.org/10.1002/andp.18601850205), 1860.
- Knust, R.: Polar Research and Supply Vessel POLARSTERN operated by the Alfred-Wegener-Institute, *Journal of large-scale research facilities*, 3, A119, doi:[10.17815/jlsrf-3-163](https://doi.org/10.17815/jlsrf-3-163), 2017.
- Komatsu, K. K., Alexeev, V. A., Repina, I. A., and Tachibana, Y.: Poleward upgliding Siberian atmospheric rivers over sea ice heat up Arctic upper air, *Sci. Rep.*, 8, 2872, doi:[10.1038/s41598-018-21159-6](https://doi.org/10.1038/s41598-018-21159-6), 2018.
- Kretzschmar, J., Salzmann, M., Mülmenstädt, J., and Quaas, J.: Arctic clouds in ECHAM6 and their sensitivity to cloud microphysics and surface fluxes, *Atmos. Chem. Phys.*, 19, 10 571–10 589, doi:[10.5194/acp-19-10571-2019](https://doi.org/10.5194/acp-19-10571-2019), 2019.
- Kruppen, T., von Albedyll, L., Goessling, H. F., Hendricks, S., Juhls, B., Spreen, G., Willmes, S., Belter, H. J., Dethloff, K., Haas, C., Kaleschke, L., Katlein, C., Tian-Kunze, X., Ricker, R., Rostosky, P., Rückert, J., Singha, S., and Sokolova, J.: MOSAiC drift expedition from October 2019 to July 2020: sea ice conditions from space and comparison with previous years, *The Cryosphere*, 15, 3897–3920, doi:[10.5194/tc-15-3897-2021](https://doi.org/10.5194/tc-15-3897-2021), 2021.
- Küchler, N., Turner, D. D., Löhnert, U., and Crewell, S.: Calibrating ground-based microwave radiometers: Uncertainty and drifts, *Radio Sci.*, 51, 311–327, doi:[10.1002/2015RS005826](https://doi.org/10.1002/2015RS005826), 2016.

- Lan, X., Dlugokencky, E., Mund, J., Crotwell, A., Crotwell, M., Moglia, E., Madronich, M., Neff, D., and Thoning, K.: Atmospheric Methane Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1983-2021, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/VNCZ-M766](https://doi.org/10.15138/VNCZ-M766), Version: 2022-11-21, 2022a.
- Lan, X., Dlugokencky, E., Mund, J., Crotwell, A., Crotwell, M., Moglia, E., Madronich, M., Neff, D., and Thoning, K.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1968-2021, NOAA Global Monitoring Laboratory Data Repository [data set], doi:[10.15138/wkgj-f215](https://doi.org/10.15138/wkgj-f215), Version: 2022-11-21, 2022b.
- Lawrence, H., Goddard, J., Sandu, I., Bormann, N., Bauer, P., and Magnusson, L.: An Assessment of the use of observations in the Arctic at ECMWF, European Centre for Medium-Range Weather Forecasts, doi:[10.21957/exlwb4o04](https://doi.org/10.21957/exlwb4o04), 2019.
- Lawrence, Z. D., Perlwitz, J., Butler, A. H., Manney, G. L., Newman, P. A., Lee, S. H., and Nash, E. R.: The Remarkably Strong Arctic Stratospheric Polar Vortex of Winter 2020: Links to Record-Breaking Arctic Oscillation and Ozone Loss, *J. Geophys. Res.: Atmos.*, 125, doi:[10.1029/2020JD033271](https://doi.org/10.1029/2020JD033271), 2020.
- Lee, S. H., Williams, P. D., and Frame, T. H. A.: Increased shear in the North Atlantic upper-level jet stream over the past four decades, *Nature*, 572, 639–642, doi:[10.1038/s41586-019-1465-z](https://doi.org/10.1038/s41586-019-1465-z), 2019.
- Lee, S.-M., Sohn, B.-J., and Kim, S.-J.: Differentiating between first-year and multiyear sea ice in the Arctic using microwave-retrieved ice emissivities, *J. Geophys. Res.: Atmos.*, 122, 5097–5112, doi:[10.1002/2016JD026275](https://doi.org/10.1002/2016JD026275), 2017.
- Liljegren, J., Boukabara, S.-A., Cady-Pereira, K., and Clough, S.: The effect of the half-width of the 22-GHz water vapor line on retrievals of temperature and water vapor profiles with a 12-channel microwave radiometer, *IEEE Trans. Geosci. Remote Sens.*, 43, 1102–1108, doi:[10.1109/TGRS.2004.839593](https://doi.org/10.1109/TGRS.2004.839593), 2005.
- Linke, O., Quaas, J., Baumer, F., Becker, S., Chylik, J., Dahlke, S., Ehrlich, A., Handorf, D., Jacobi, C., Kalesse-Los, H., Lelli, L., Mehrdad, S., Neggers, R. A. J., Riebold, J., Saavedra Garfias, P., Schnierstein, N., Shupe, M. D., Smith, C., Spreen, G., Verneuil, B., Vinjamuri, K. S., Vountas, M., and Wendisch, M.: Constraints on simulated past Arctic amplification and lapse rate feedback from observations, *Atmos. Chem. Phys.*, 23, 9963–9992, doi:[10.5194/acp-23-9963-2023](https://doi.org/10.5194/acp-23-9963-2023), 2023.

- Liou, K. N.: An Introduction to Atmospheric Radiation, vol. 84 of *International Geophysics Series*, Academic Press, 525 B Street, Suite 1900, San Diego, California 92101-4495, USA, Second edn., ISBN 978-0-12-451451-5, doi:[10.1016/s0074-6142\(02\)x8015-2](https://doi.org/10.1016/s0074-6142(02)x8015-2), 2002.
- Loeb, N. A., Crawford, A., Stroeve, J. C., and Hanesiak, J.: Extreme Precipitation in the Eastern Canadian Arctic and Greenland: An Evaluation of Atmospheric Reanalyses, *Front. Environ. Sci.*, 10, 866929, doi:[10.3389/fenvs.2022.866929](https://doi.org/10.3389/fenvs.2022.866929), 2022.
- Lohmann, U., Lüönd, F., and Mahrt, F.: An Introduction to Clouds: From the Microscale to Climate, Cambridge University Press, Cambridge, UK, First edn., ISBN 978-1-139-08751-3, doi:[10.1017/CBO9781139087513](https://doi.org/10.1017/CBO9781139087513), 2016.
- Löhnert, U.: Ground-based microwave radiometer reprocessing mwr_pro, Zenodo [software], doi:[10.5281/zenodo.7973553](https://doi.org/10.5281/zenodo.7973553), 2023.
- Löhnert, U. and Crewell, S.: Accuracy of cloud liquid water path from ground-based microwave radiometry 1. Dependency on cloud model statistics, *Radio Sci.*, 38, doi:[10.1029/2002RS002654](https://doi.org/10.1029/2002RS002654), 2003.
- Löhnert, U. and Maier, O.: Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges, *Atmos. Meas. Tech.*, 5, 1121–1134, doi:[10.5194/amt-5-1121-2012](https://doi.org/10.5194/amt-5-1121-2012), 2012.
- Löhnert, U., Crewell, S., and Simmer, C.: An Integrated Approach toward Retrieving Physically Consistent Profiles of Temperature, Humidity, and Cloud Liquid Water, *J. Appl. Meteorol. Climatol.*, 43, 1295–1307, doi:[10.1175/1520-0450\(2004\)043<1295:AIATRP>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<1295:AIATRP>2.0.CO;2), 2004.
- Löhnert, U., Turner, D. D., and Crewell, S.: Ground-Based Temperature and Humidity Profiling Using Spectral Infrared and Microwave Observations. Part I: Simulated Retrieval Performance in Clear-Sky Conditions, *J. Appl. Meteorol. Climatol.*, 48, 1017–1032, doi:[10.1175/2008JAMC2060.1](https://doi.org/10.1175/2008JAMC2060.1), 2009.
- L'Ecuyer, T. S., Drouin, B. J., Anheuser, J., Grames, M., Henderson, D. S., Huang, X., Kahn, B. H., Kay, J. E., Lim, B. H., Mateling, M., Merrelli, A., Miller, N. B., Padmanabhan, S., Peterson, C., Schlegel, N.-J., White, M. L., and Xie, Y.: The Polar Radiant Energy in the Far Infrared Experiment: A New Perspective on Polar Longwave Energy Exchanges, *Bull. Am. Meteorol. Soc.*, 102, E1431–E1449, doi:[10.1175/BAMS-D-20-0155.1](https://doi.org/10.1175/BAMS-D-20-0155.1), 2021.

- Macke, A., Kalisch, J., Zoll, Y., and Bumke, K.: Radiative effects of the cloudy atmosphere from ground and satellite based observations, *EPJ Web of Conferences*, 9, 83–94, doi:[10.1051/epjconf/201009006](https://doi.org/10.1051/epjconf/201009006), 2010.
- Marke, T., Löhnert, U., Tukiainen, S., Siipola, T., and Pospichal, B.: MWRpy: A Python package for processing microwave radiometer data, *Journal of Open Source Software*, 9, 6733, doi:[10.21105/joss.06733](https://doi.org/10.21105/joss.06733), 2024.
- Mas, J. F. and Flores, J. J.: The application of artificial neural networks to the analysis of remotely sensed data, *Int. J. Remote Sens.*, 29, 617–663, doi:[10.1080/01431160701352154](https://doi.org/10.1080/01431160701352154), 2008.
- Mathew, N., Heygster, G., Melsheimer, C., and Kaleschke, L.: Surface Emissivity of Arctic Sea Ice at AMSU Window Frequencies, *IEEE Trans. Geosci. Remote Sens.*, 46, 2298–2306, doi:[10.1109/TGRS.2008.916630](https://doi.org/10.1109/TGRS.2008.916630), 2008.
- Maturilli, M.: High resolution radiosonde measurements from station Ny-Ålesund (2017-04 et seq), PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.914973](https://doi.org/10.1594/PANGAEA.914973), 2020.
- Maturilli, M. and Kayser, M.: Homogenized radiosonde record at station Ny-Ålesund, Spitsbergen, 1993-2014, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.845373](https://doi.org/10.1594/PANGAEA.845373), 2016.
- Maturilli, M. and Kayser, M.: Arctic warming, moisture increase and circulation changes observed in the Ny-Ålesund homogenized radiosonde record, *Theor. Appl. Climatol.*, 130, 1–17, doi:[10.1007/s00704-016-1864-0](https://doi.org/10.1007/s00704-016-1864-0), 2017a.
- Maturilli, M. and Kayser, M.: Homogenized radiosonde record at station Ny-Ålesund, Spitsbergen, 2015-2016, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.875196](https://doi.org/10.1594/PANGAEA.875196), 2017b.
- Maturilli, M., Herber, A., and König-Langlo, G.: Climatology and time series of surface meteorology in Ny-Ålesund, Svalbard, *Earth Syst. Sci. Data*, 5, 155–163, doi:[10.5194/essd-5-155-2013](https://doi.org/10.5194/essd-5-155-2013), 2013.
- Maturilli, M., Herber, A., and König-Langlo, G.: Surface radiation climatology for Ny-Ålesund, Svalbard (78.9° N), basic observations for trend detection, *Theor. Appl. Climatol.*, 120, 331–339, doi:[10.1007/s00704-014-1173-4](https://doi.org/10.1007/s00704-014-1173-4), 2015.
- Maturilli, M., Holdridge, D. J., Dahlke, S., Graeser, J., Sommerfeld, A., Jaiser, R., Deckelmann, H., and Schulz, A.: Initial radiosonde data from 2019-10 to 2020-09 during project MOSAiC, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.928656](https://doi.org/10.1594/PANGAEA.928656), 2021.

- Mech, M., Kliesch, L.-L., Anhäuser, A., Rose, T., Kollias, P., and Crewell, S.: Microwave Radar/radiometer for Arctic Clouds (MiRAC): first insights from the ACLOUD campaign, *Atmos. Meas. Tech.*, 12, 5019–5037, doi:[10.5194/amt-12-5019-2019](https://doi.org/10.5194/amt-12-5019-2019), 2019a.
- Mech, M., Maahn, M., Ori, D., and Orlandi, E.: PAMTRA: Passive and Active Microwave TRAnsfer tool v1.0, Zenodo [software], doi:[10.5281/zenodo.3582992](https://doi.org/10.5281/zenodo.3582992), 2019b.
- Mech, M., Maahn, M., Kneifel, S., Ori, D., Orlandi, E., Kollias, P., Schemann, V., and Crewell, S.: PAMTRA 1.0: the Passive and Active Microwave radiative TRAnsfer tool for simulating radiometer and radar measurements of the cloudy atmosphere, *Geosci. Model Dev.*, 13, 4229–4251, doi:[10.5194/gmd-13-4229-2020](https://doi.org/10.5194/gmd-13-4229-2020), 2020.
- Mewes, D. and Jacobi, C.: Heat transport pathways into the Arctic and their connections to surface air temperatures, *Atmos. Chem. Phys.*, 19, 3927–3937, doi:[10.5194/acp-19-3927-2019](https://doi.org/10.5194/acp-19-3927-2019), 2019.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.: Atmos.*, 102, 16 663–16 682, doi:[10.1029/97JD00237](https://doi.org/10.1029/97JD00237), 1997.
- Männel, B., Zus, F., Dick, G., Glaser, S., Semmling, M., Balidakis, K., Wickert, J., Maturilli, M., Dahlke, S., and Schuh, H.: GNSS-based water vapor estimation and validation during the MOSAiC expedition, *Atmos. Meas. Tech.*, 14, 5127–5138, doi:[10.5194/amt-14-5127-2021](https://doi.org/10.5194/amt-14-5127-2021), 2021.
- Naakka, T., Nygård, T., and Vihma, T.: Arctic Humidity Inversions: Climatology and Processes, *J. Climate*, 31, 3765–3787, doi:[10.1175/JCLI-D-17-0497.1](https://doi.org/10.1175/JCLI-D-17-0497.1), 2018.
- Newell, R. E., Newell, N. E., Zhu, Y., and Scott, C.: Tropospheric rivers? - A pilot study, *Geophys. Res. Lett.*, 19, 2401–2404, doi:[10.1029/92GL02916](https://doi.org/10.1029/92GL02916), 1992.
- Nie, Y., Chen, G., Lu, J., Zhou, W., and Zhang, Y.: Constraining the Varied Response of Northern Hemisphere Winter Circulation Waviness to Climate Change, *Geophys. Res. Lett.*, 50, e2022GL102 150, doi:[10.1029/2022GL102150](https://doi.org/10.1029/2022GL102150), 2023.
- Ning, T. and Elgered, G.: Trends in the Atmospheric Water Vapor Content From Ground-Based GPS: The Impact of the Elevation Cutoff Angle, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 5, 744–751, doi:[10.1109/JSTARS.2012.2191392](https://doi.org/10.1109/JSTARS.2012.2191392), 2012.

- Ning, T., Wang, J., Elgered, G., Dick, G., Wickert, J., Bradke, M., Sommer, M., Querel, R., and Smale, D.: The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations, *Atmos. Meas. Tech.*, 9, 79–92, doi:[10.5194/amt-9-79-2016](https://doi.org/10.5194/amt-9-79-2016), 2016.
- Nixdorf, U., Dethloff, K., Rex, M., Shupe, M., Sommerfeld, A., Perovich, D. K., Nicolaus, M., Heuzé, C., Rabe, B., Loose, B., Damm, E., Gradinger, R., Fong, A., Maslowski, W., Rinke, A., Kwok, R., Spreen, G., Wendisch, M., Herber, A., Hirsekorn, M., Mohaupt, V., Frickenhaus, S., Immerz, A., Weiss-Tuider, K., König, B., Mengedoht, D., Regnery, J., Gerchow, P., Ransby, D., Krumpen, T., Morgenstern, A., Haas, C., Kanzow, T., Rack, F. R., Saitzev, V., Sokolov, V., Makarov, A., Schwarze, S., Wunderlich, T., Wurr, K., and Boetius, A.: MOSAiC Extended Acknowledgement, Zenodo, doi:[10.5281/zenodo.5541624](https://doi.org/10.5281/zenodo.5541624), 2021.
- NOAA, NASA, and USFS: U.S. Standard Atmosphere, 1976, Tech. rep., National Oceanic and Atmospheric Administration (NOAA), National Aeronautics and Space Administration (NASA) and United States Air Force (USAF), <https://ntrs.nasa.gov/api/citations/19770009539/downloads/19770009539.pdf>, Accessed: 2024-07-14, 1976.
- Nomokonova, T., Ebell, K., Löhnert, U., Maturilli, M., Ritter, C., and O’Connor, E.: Statistics on clouds and their relation to thermodynamic conditions at Ny-Ålesund using ground-based sensor synergy, *Atmos. Chem. Phys.*, 19, 4105–4126, doi:[10.5194/acp-19-4105-2019](https://doi.org/10.5194/acp-19-4105-2019), 2019.
- Nomokonova, T., Ebell, K., Löhnert, U., Maturilli, M., and Ritter, C.: The influence of water vapor anomalies on clouds and their radiative effect at Ny-Ålesund, *Atmos. Chem. Phys.*, 20, 5157–5173, doi:[10.5194/acp-20-5157-2020](https://doi.org/10.5194/acp-20-5157-2020), 2020.
- Nygård, T., Valkonen, T., and Vihma, T.: Antarctic Low-Tropospheric Humidity Inversions: 10-Yr Climatology, *J. Climate*, 26, 5205–5219, doi:[10.1175/JCLI-D-12-00446.1](https://doi.org/10.1175/JCLI-D-12-00446.1), 2013.
- Nygård, T., Valkonen, T., and Vihma, T.: Characteristics of Arctic low-tropospheric humidity inversions based on radio soundings, *Atmos. Chem. Phys.*, 14, 1959–1971, doi:[10.5194/acp-14-1959-2014](https://doi.org/10.5194/acp-14-1959-2014), 2014.
- Nygård, T., Naakka, T., and Vihma, T.: Horizontal Moisture Transport Dominates the Regional Moistening Patterns in the Arctic, *J. Climate*, 33, 6793–6807, doi:[10.1175/JCLI-D-19-0891.1](https://doi.org/10.1175/JCLI-D-19-0891.1), 2020.

- Ohmura, A.: Physical Basis for the Temperature-Based Melt-Index Method, *J. Appl. Meteorol. Climatol.*, 40, 753–761, doi:[10.1175/1520-0450\(2001\)040<0753:PBFTTB>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0753:PBFTTB>2.0.CO;2), 2001.
- Oliver, D. S., Reynolds, A. C., and Liu, N.: Inverse theory for petroleum reservoir characterization and history matching, Cambridge University Press, Cambridge, UK, ISBN 978-0-521-88151-7, 2008.
- Orlandi, E. and Walbröl, A.: Training and test data for retrievals based on MiRAC-P observations during MOSAiC, Zenodo [data set], doi:[10.5281/zenodo.5846394](https://doi.org/10.5281/zenodo.5846394), 2022.
- Parracho, A. C., Bock, O., and Bastin, S.: Global IWV trends and variability in atmospheric reanalyses and GPS observations, *Atmos. Chem. Phys.*, 18, 16 213–16 237, doi:[10.5194/acp-18-16213-2018](https://doi.org/10.5194/acp-18-16213-2018), 2018.
- Peixoto, J. P. and Oort, A. H.: The Climatology of Relative Humidity in the Atmosphere, *J. Climate*, 9, 3443–3463, doi:[10.1175/1520-0442\(1996\)009<3443:TCORHI>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<3443:TCORHI>2.0.CO;2), 1996.
- Perro, C., Lesins, G., Duck, T. J., and Cadeddu, M.: A microwave satellite water vapour column retrieval for polar winter conditions, *Atmos. Meas. Tech.*, 9, 2241–2252, doi:[10.5194/amt-9-2241-2016](https://doi.org/10.5194/amt-9-2241-2016), 2016.
- Persson, P. O. G., Fairall, C. W., Andreas, E. L., Guest, P. S., and Perovich, D. K.: Measurements near the Atmospheric Surface Flux Group tower at SHEBA: Near-surface conditions and surface energy budget, *J. Geophys. Res.: Oceans*, 107, doi:[10.1029/2000JC000705](https://doi.org/10.1029/2000JC000705), 2002.
- Petty, G. W.: A first course in atmospheric radiation, Sundog Publishing, Madison, Wisconsin, USA, Second edn., ISBN 978-0-9729033-1-8, 2006.
- Pithan, F., Medeiros, B., and Mauritsen, T.: Mixed-phase clouds cause climate model biases in Arctic wintertime temperature inversions, *Clim. Dyn.*, 43, 289–303, doi:[10.1007/s00382-013-1964-9](https://doi.org/10.1007/s00382-013-1964-9), 2014.
- Priestley, M. D. K. and Catto, J. L.: Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure, *Weather Clim. Dyn.*, 3, 337–360, doi:[10.5194/wcd-3-337-2022](https://doi.org/10.5194/wcd-3-337-2022), 2022.
- Prigent, C., Aires, F., Wang, D., Fox, S., and Harlow, C.: Sea-surface emissivity parametrization from microwaves to millimetre waves: Sea-Surface Emissivity up to 700 GHz, *Q. J. R. Meteorolog. Soc.*, 143, 596–605, doi:[10.1002/qj.2953](https://doi.org/10.1002/qj.2953), 2017.

- Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, *Commun. Earth Environ.*, 3, 168, doi:[10.1038/s43247-022-00498-3](https://doi.org/10.1038/s43247-022-00498-3), 2022.
- Raval, A. and Ramanathan, V.: Observational determination of the greenhouse effect, *Nature*, 342, 758–761, doi:[10.1038/342758a0](https://doi.org/10.1038/342758a0), 1989.
- Reinert, D., Prill, F., Frank, H., Denard, M., Baldauf, M., Schraff, C., Gebhardt, C., Marsigli, C., and Zängl, G.: DWD Database Reference for the Global and Regional ICON and ICON-EPS Forecasting System, Deutscher Wetterdienst (DWD), https://www.dwd.de/SharedDocs/downloads/DE/modelldokumentationen/nwv/icon/icon_dbbeschr_aktuell.html, Accessed: 2024-03-14, 2024.
- Rex, M.: Links to master tracks in different resolutions of POLARSTERN cruise PS122/1, Tromsø - Arctic Ocean, 2019-09-20 - 2019-12-13 (Version 2), PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.924668](https://doi.org/10.1594/PANGAEA.924668), 2020.
- Rex, M.: Master tracks in different resolutions of POLARSTERN cruise PS122/4, Longyearbyen - Arctic Ocean, 2020-06-04 - 2020-08-12, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.926829](https://doi.org/10.1594/PANGAEA.926829), 2021a.
- Rex, M.: Master tracks in different resolutions of POLARSTERN cruise PS122/5, Arctic Ocean - Bremerhaven, 2020-08-12 - 2020-10-12, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.926910](https://doi.org/10.1594/PANGAEA.926910), 2021b.
- Rhodes, B.: Skyfield: High precision research-grade positions for planets and Earth satellites generator, Astrophysics Source Code Library, record ascl:1907.024 [software], <https://ui.adsabs.harvard.edu/abs/2019ascl.soft07024R>, Provided by the SAO/NASA Astrophysics Data System, Accessed: 2024-06-14, 2019.
- Rigor, I. G., Colony, R. L., and Martin, S.: Variations in Surface Air Temperature Observations in the Arctic, 1979–97, *J. Climate*, 13, 896–914, doi:[10.1175/1520-0442\(2000\)013<0896:VISATO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0896:VISATO>2.0.CO;2), 2000.
- Rinke, A., Maturilli, M., Graham, R. M., Matthes, H., Handorf, D., Cohen, L., Hudson, S. R., and Moore, J. C.: Extreme cyclone events in the Arctic: Winter-

- time variability and trends, *Environ. Res. Lett.*, 12, 094 006, doi:[10.1088/1748-9326/aa7def](https://doi.org/10.1088/1748-9326/aa7def), 2017.
- Rinke, A., Segger, B., Crewell, S., Maturilli, M., Naakka, T., Nygård, T., Vihma, T., Alshawaf, F., Dick, G., Wickert, J., and Keller, J.: Trends of Vertically Integrated Water Vapor over the Arctic during 1979–2016: Consistent Moistening All Over?, *J. Climate*, 32, 6097–6116, doi:[10.1175/JCLI-D-19-0092.1](https://doi.org/10.1175/JCLI-D-19-0092.1), 2019.
- Rinke, A., Cassano, J. J., Cassano, E. N., Jaiser, R., and Handorf, D.: Meteorological conditions during the MOSAiC expedition, *Elem. Sci. Anth.*, 9, 00 023, doi:[10.1525/elementa.2021.00023](https://doi.org/10.1525/elementa.2021.00023), 2021.
- Risse, N., Mech, M., Prigent, C., Spreen, G., and Crewell, S.: Assessing the sea ice microwave emissivity up to submillimeter waves from airborne and satellite observations, *EGUsphere* [preprint], submitted to *The Cryosphere*, in discussion since 2024-03-06, doi:[10.5194/egusphere-2024-179](https://doi.org/10.5194/egusphere-2024-179), 2024.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, no. 2 in *Series on atmospheric, oceanic and planetary physics*, World Scientific, Singapore, repr edn., ISBN 978-981-02-2740-1, 2000.
- Roman, J., Knuteson, R., August, T., Hultberg, T., Ackerman, S., and Revercomb, H.: A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations, *J. Geophys. Res.: Atmos.*, 121, 8925–8948, doi:[10.1002/2016JD024806](https://doi.org/10.1002/2016JD024806), 2016.
- Rose, T., Crewell, S., Löhnert, U., and Simmer, C.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, *Atmos. Res.*, 75, 183–200, doi:[10.1016/j.atmosres.2004.12.005](https://doi.org/10.1016/j.atmosres.2004.12.005), 2005.
- Rosenkranz, P. W.: Water vapor microwave continuum absorption: A comparison of measurements and models, *Radio Sci.*, 33, 919–928, doi:[10.1029/98RS01182](https://doi.org/10.1029/98RS01182), 1998.
- Roy, R. J., Lebsock, M., and Kurowski, M. J.: Spaceborne differential absorption radar water vapor retrieval capabilities in tropical and subtropical boundary layer cloud regimes, *Atmos. Meas. Tech.*, 14, 6443–6468, doi:[10.5194/amt-14-6443-2021](https://doi.org/10.5194/amt-14-6443-2021), 2021.
- Rückert, J. E., Huntemann, M., Tonboe, R. T., and Spreen, G.: Modeling Snow and Ice Microwave Emissions in the Arctic for a Multi-Parameter Retrieval of Surface and Atmospheric Variables From Microwave Radiometer Satellite Data, *Earth Space Sci.*, 10, e2023EA003 177, doi:[10.1029/2023EA003177](https://doi.org/10.1029/2023EA003177), 2023a.

- Rückert, J. E., Rostosky, P., Huntemann, M., Clemens-Sewall, D., Ebell, K., Kaleschke, L., Lemmetyinen, J., Macfarlane, A. R., Naderpour, R., Stroeve, J., Walbröl, A., and Spreen, G.: Sea ice concentration satellite retrievals influenced by surface changes due to warm air intrusions: A case study from the MOSAiC expedition, *Elem. Sci. Anth.*, 11, 00 039, doi:[10.1525/elementa.2023.00039](https://doi.org/10.1525/elementa.2023.00039), 2023b.
- Rüeger, J. M.: Refractive index formulae for electronic distance measurement with radio and millimetre waves, in: *Refractive indices of light, infrared and radio waves in the atmosphere*, edited by Rüeger, J. M., pp. 1–52, School of Surveying and Spatial Information Systems, University of New South Wales, Sydney, Australia, ISBN 0-7334-1866 X, <https://www.unsw.edu.au/content/dam/pdfs/engineering/civil-environmental/historical-documents/unisurv-special-series/s68.pdf>, Accessed: 2024-09-01, 2002.
- Saouabe, T., El Khalki, E. M., Saidi, M. E. M., Najmi, A., Hadri, A., Rachidi, S., Jadoud, M., and Trambly, Y.: Evaluation of the GPM-IMERG Precipitation Product for Flood Modeling in a Semi-Arid Mountainous Basin in Morocco, *Water*, 12, 2516, doi:[10.3390/w12092516](https://doi.org/10.3390/w12092516), 2020.
- Scarlat, R. C., Heygster, G., and Pedersen, L. T.: Experiences With an Optimal Estimation Algorithm for Surface and Atmospheric Parameter Retrieval From Passive Microwave Data in the Arctic, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 10, 3934–3947, doi:[10.1109/JSTARS.2017.2739858](https://doi.org/10.1109/JSTARS.2017.2739858), 2017.
- Schemann, V. and Ebell, K.: Simulation of mixed-phase clouds with the ICON large-eddy model in the complex Arctic environment around Ny-Ålesund, *Atmos. Chem. Phys.*, 20, 475–485, doi:[10.5194/acp-20-475-2020](https://doi.org/10.5194/acp-20-475-2020), 2020.
- Schnierstein, N., Chylik, J., Shupe, M. D., and Neggers, R.: Standardized daily high-resolution large-eddy simulations of the Arctic boundary layer and clouds during the complete MOSAiC drift, submitted to JAMES on 2024-03-01, doi:[10.22541/essoar.171033230.00246941/v1](https://doi.org/10.22541/essoar.171033230.00246941/v1), 2024.
- Schnitt, S., Löhnert, U., and Preusker, R.: Potential of Dual-Frequency Radar and Microwave Radiometer Synergy for Water Vapor Profiling in the Cloudy Trade Wind Environment, *J. Atmos. Oceanic Technol.*, 37, 1973–1986, doi:[10.1175/JTECH-D-19-0110.1](https://doi.org/10.1175/JTECH-D-19-0110.1), 2020.
- Schröder, M., Lockhoff, M., Fell, F., Forsythe, J., Trent, T., Bennartz, R., Borbas, E., Bosilovich, M. G., Castelli, E., Hersbach, H., Kachi, M., Kobayashi, S.,

- Kursinski, E. R., Loyola, D., Mears, C., Preusker, R., Rossow, W. B., and Saha, S.: The GEWEX Water Vapor Assessment archive of water vapour products from satellite observations and reanalyses, *Earth Syst. Sci. Data*, 10, 1093–1117, doi:[10.5194/essd-10-1093-2018](https://doi.org/10.5194/essd-10-1093-2018), 2018.
- Screen, J. A. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, *Nature*, 464, 1334–1337, doi:[10.1038/nature09051](https://doi.org/10.1038/nature09051), 2010.
- Screen, J. A., Deser, C., and Simmonds, I.: Local and remote controls on observed Arctic warming, *Geophys. Res. Lett.*, 39, doi:[10.1029/2012GL051598](https://doi.org/10.1029/2012GL051598), 2012.
- Sedlar, J., Shupe, M. D., and Tjernström, M.: On the Relationship between Thermodynamic Structure and Cloud Top, and Its Climate Significance in the Arctic, *J. Climate*, 25, 2374–2393, doi:[10.1175/JCLI-D-11-00186.1](https://doi.org/10.1175/JCLI-D-11-00186.1), 2012.
- Sedlar, J., Tjernström, M., Rinke, A., Orr, A., Cassano, J., Fettweis, X., Heinemann, G., Seefeldt, M., Solomon, A., Matthes, H., Phillips, T., and Webster, S.: Confronting Arctic Troposphere, Clouds, and Surface Energy Budget Representations in Regional Climate Models With Observations, *J. Geophys. Res.: Atmos.*, 125, e2019JD031783, doi:[10.1029/2019JD031783](https://doi.org/10.1029/2019JD031783), 2020.
- Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, *Global Planet. Change*, 77, 85–96, doi:[10.1016/j.gloplacha.2011.03.004](https://doi.org/10.1016/j.gloplacha.2011.03.004), 2011.
- Serreze, M. C., Rehder, M. C., Barry, R. G., Kahl, J. D., and Zaitseva, N. A.: The distribution and transport of atmospheric water vapour over the Arctic Basin, *Int. J. Climatol.*, 15, 709–727, doi:[10.1002/joc.3370150702](https://doi.org/10.1002/joc.3370150702), 1995.
- Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., and Holland, M. M.: The emergence of surface-based Arctic amplification, *The Cryosphere*, 3, 11–19, doi:[10.5194/tc-3-11-2009](https://doi.org/10.5194/tc-3-11-2009), 2009.
- Serreze, M. C., Barrett, A. P., and Stroeve, J.: Recent changes in tropospheric water vapor over the Arctic as assessed from radiosondes and atmospheric reanalyses, *J. Geophys. Res.: Atmos.*, 117, D10104, doi:[10.1029/2011JD017421](https://doi.org/10.1029/2011JD017421), 2012.
- Shine, K. P. and Sinha, A.: Sensitivity of the Earth’s climate to height-dependent changes in the water vapour mixing ratio, *Nature*, 354, 382–384, doi:[10.1038/354382a0](https://doi.org/10.1038/354382a0), 1991.

- Shupe, M. D. and Intrieri, J. M.: Cloud Radiative Forcing of the Arctic Surface: The Influence of Cloud Properties, Surface Albedo, and Solar Zenith Angle, *J. Climate*, 17, 616–628, doi:[10.1175/1520-0442\(2004\)017<0616:CRFOTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0616:CRFOTA>2.0.CO;2), 2004.
- Shupe, M. D., Rex, M., Dethloff, K., Damm, E., Fong, A. A., Gradinger, R., Heuzé, C., B, L., Makarov, A., Maslowski, W., Nicolaus, M., Perovich, D., Rabe, B., Rinke, A., Sokolov, V., and Sommerfeld, A.: Arctic Report Card 2020: The MOSAiC Expedition: A Year Drifting with the Arctic Sea Ice, Arctic Report Card, doi:[10.25923/9G3V-XH92](https://doi.org/10.25923/9G3V-XH92), 2020.
- Shupe, M. D., Rex, M., Blomquist, B., Persson, P. O. G., Schmale, J., Uttal, T., Althausen, D., Angot, H., Archer, S., Bariteau, L., Beck, I., Bilberry, J., Bucci, S., Buck, C., Boyer, M., Brasseur, Z., Brooks, I. M., Calmer, R., Cassano, J., Castro, V., Chu, D., Costa, D., Cox, C. J., Creamean, J., Crewell, S., Dahlke, S., Damm, E., de Boer, G., Deckelmann, H., Dethloff, K., Dütsch, M., Ebell, K., Ehrlich, A., Ellis, J., Engelmann, R., Fong, A. A., Frey, M. M., Gallagher, M. R., Ganzeveld, L., Gradinger, R., Graeser, J., Greenamyre, V., Griesche, H., Griffiths, S., Hamilton, J., Heinemann, G., Helmig, D., Herber, A., Heuzé, C., Hofer, J., Houchens, T., Howard, D., Inoue, J., Jacobi, H.-W., Jaiser, R., Jokinen, T., Jourdan, O., Jozef, G., King, W., Kirchgaessner, A., Klingebiel, M., Krassovski, M., Krumpfen, T., Lampert, A., Landing, W., Laurila, T., Lawrence, D., Lonardi, M., Loose, B., Lüpkes, C., Maahn, M., Macke, A., Maslowski, W., Marsay, C., Maturilli, M., Mech, M., Morris, S., Moser, M., Nicolaus, M., Ortega, P., Osborn, J., Pätzold, F., Perovich, D. K., Petäjä, T., Pilz, C., Pirazzini, R., Posman, K., Powers, H., Pratt, K. A., Preußner, A., Quéléver, L., Radenz, M., Rabe, B., Rinke, A., Sachs, T., Schulz, A., Siebert, H., Silva, T., Solomon, A., Sommerfeld, A., Spreen, G., Stephens, M., Stohl, A., Svensson, G., Uin, J., Viegas, J., Voigt, C., von der Gathen, P., Wehner, B., Welker, J. M., Wendisch, M., Werner, M., Xie, Z., and Yue, F.: Overview of the MOSAiC expedition—Atmosphere, *Elem. Sci. Anth.*, 10, 00060, doi:[10.1525/elementa.2021.00060](https://doi.org/10.1525/elementa.2021.00060), 2022.
- Solheim, F., Godwin, J. R., Westwater, E. R., Han, Y., Keihm, S. J., Marsh, K., and Ware, R.: Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods, *Radio Sci.*, 33, 393–404, doi:[10.1029/97RS03656](https://doi.org/10.1029/97RS03656), 1998.
- Solomon, A. and Jozef, G.: Coupled Arctic Forecast System (CAFS) forecasts used in study evaluating atmospheric stability with MOSAiC soundings,

- PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.964776](https://doi.org/10.1594/PANGAEA.964776), 2024.
- Solomon, A., Shupe, M. D., Persson, P. O. G., and Morrison, H.: Moisture and dynamical interactions maintaining decoupled Arctic mixed-phase stratocumulus in the presence of a humidity inversion, *Atmos. Chem. Phys.*, 11, 10 127–10 148, doi:[10.5194/acp-11-10127-2011](https://doi.org/10.5194/acp-11-10127-2011), 2011.
- Solomon, A., Shupe, M. D., Svensson, G., Barton, N. P., Batrak, Y., Bazile, E., Day, J. J., Doyle, J. D., Frank, H. P., Keeley, S., Remes, T., and Tolstykh, M.: The winter central Arctic surface energy budget: A model evaluation using observations from the MOSAiC campaign, *Elem. Sci. Anth.*, 11, 00 104, doi:[10.1525/elementa.2022.00104](https://doi.org/10.1525/elementa.2022.00104), 2023.
- Solomon, A., Cox, C. J., Intrieri, J. M., Persson, P. O. G., de Boer, G., Shupe, M. D., Hughes, M., and Capotondi, A.: Description and Evaluation of the NOAA Experimental Coupled Arctic Forecast System (CAFS) Model, NOAA Physical Sciences Laboratory, doi:[10.25923/V7VJ-Z744](https://doi.org/10.25923/V7VJ-Z744), 2024.
- Sommer, M., Dirksen, R., and Immler, F.: RS92 GRUAN Data Product Version 2 (RS92-GDP.2), GRUAN Lead Centre (at Deutscher Wetterdienst), doi:[10.5676/GRUAN/RS92-GDP.2](https://doi.org/10.5676/GRUAN/RS92-GDP.2), 2012.
- Sommer, M., von Rohden, C., Simeonov, T., Oelsner, P., Naebert, T., Romanens, G., Jauhianinen, H., Survo, P., and Dirksen, R.: GRUAN characterization and data processing of the Vaisala RS41 radiosonde, Tech. rep., GRUAN Lead Centre (at Deutscher Wetterdienst), https://www.gruan.org/gruan/editor/documents/gruan/GRUAN-TD-8_RS41_v1.0.0_20230628_final.pdf, GRUAN-TD-8, 1.0.0, Accessed: 2024-08-27, 2023.
- Spreen, G., Kaleschke, L., and Heygster, G.: Sea ice remote sensing using AMSR-E 89-GHz channels, *J. Geophys. Res.*, 113, C02S03, doi:[10.1029/2005JC003384](https://doi.org/10.1029/2005JC003384), 2008.
- Steinke, S., Eikenberg, S., Löhnert, U., Dick, G., Klocke, D., Di Girolamo, P., and Crewell, S.: Assessment of small-scale integrated water vapour variability during HOPE, *Atmos. Chem. Phys.*, 15, 2675–2692, doi:[10.5194/acp-15-2675-2015](https://doi.org/10.5194/acp-15-2675-2015), 2015.
- Stendel, M., Francis, J., White, R., Williams, P. D., and Woollings, T.: The jet stream and climate change, in: *Climate Change*, edited by Letcher, T. M.,

- pp. 327–357, Elsevier, ISBN 978-0-12-821575-3, doi:[10.1016/B978-0-12-821575-3.00015-3](https://doi.org/10.1016/B978-0-12-821575-3.00015-3), 2021.
- Tang, S. and Yang, Y.: Why neural networks apply to scientific computing?, *Theor. Appl. Mech. Lett.*, 11, 100242, doi:[10.1016/j.taml.2021.100242](https://doi.org/10.1016/j.taml.2021.100242), 2021.
- Tjernström, M. and Graversen, R. G.: The vertical structure of the lower Arctic troposphere analysed from observations and the ERA-40 reanalysis, *Q. J. R. Meteorolog. Soc.*, 135, 431–443, doi:[10.1002/qj.380](https://doi.org/10.1002/qj.380), 2009.
- Tjernström, M., Leck, C., Persson, P. O. G., Jensen, M. L., Oncley, S. P., and Targino, A.: Experimental Equipment: A Supplement to The Summertime Arctic Atmosphere: Meteorological Measurements during the Arctic Ocean Experiment 2001, *Bull. Am. Meteorol. Soc.*, 85, ES14–ES18, doi:[10.1175/BAMS-85-9-Tjernstrom](https://doi.org/10.1175/BAMS-85-9-Tjernstrom), 2004.
- Tjernström, M., Shupe, M. D., Brooks, I. M., Persson, P. O. G., Prytherch, J., Salisbury, D. J., Sedlar, J., Achtert, P., Brooks, B. J., Johnston, P. E., Sotiropoulou, G., and Wolfe, D.: Warm-air advection, air mass transformation and fog causes rapid ice melt, *Geophys. Res. Lett.*, 42, 5594–5602, doi:[10.1002/2015GL064373](https://doi.org/10.1002/2015GL064373), 2015.
- Tjernström, M., Shupe, M. D., Brooks, I. M., Achtert, P., Prytherch, J., and Sedlar, J.: Arctic Summer Airmass Transformation, Surface Inversions, and the Surface Energy Budget, *J. Climate*, 32, 769–789, doi:[10.1175/JCLI-D-18-0216.1](https://doi.org/10.1175/JCLI-D-18-0216.1), 2019.
- Toporov, M. and Löhnert, U.: Synergy of Satellite- and Ground-Based Observations for Continuous Monitoring of Atmospheric Stability, Liquid Water Path, and Integrated Water Vapor: Theoretical Evaluations Using Reanalysis and Neural Networks, *J. Appl. Meteorol. Climatol.*, 59, 1153–1170, doi:[10.1175/JAMC-D-19-0169.1](https://doi.org/10.1175/JAMC-D-19-0169.1), 2020.
- Town, M. S., Walden, V. P., and Warren, S. G.: Spectral and Broadband Longwave Downwelling Radiative Fluxes, Cloud Radiative Forcing, and Fractional Cloud Cover over the South Pole, *J. Climate*, 18, 4235–4252, doi:[10.1175/JCLI3525.1](https://doi.org/10.1175/JCLI3525.1), 2005.
- Trenberth, K. E.: Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change, *Clim. Change*, 39, 667–694, doi:[10.1023/A:1005319109110](https://doi.org/10.1023/A:1005319109110), 1998.

- Tukiainen, S., O'Connor, E., and Korpinen, A.: CloudnetPy: A Python package for processing cloud remote sensing data, *Journal of Open Source Software*, 5, 2123, doi:[10.21105/joss.02123](https://doi.org/10.21105/joss.02123), 2020.
- Turner, D., Cadeddu, M., Lohnert, U., Crewell, S., and Vogelmann, A.: Modifications to the Water Vapor Continuum in the Microwave Suggested by Ground-Based 150-GHz Observations, *IEEE Trans. Geosci. Remote Sens.*, 47, 3326–3337, doi:[10.1109/TGRS.2009.2022262](https://doi.org/10.1109/TGRS.2009.2022262), 2009.
- Tyndall, J.: I. The Bakerian Lecture.—On the absorption and radiation of heat by gases and vapours, and on the physical connexion of radiation, absorption, and conduction, *Philosophical Transactions of the Royal Society of London*, 151, 1–36, doi:[10.1098/rstl.1861.0001](https://doi.org/10.1098/rstl.1861.0001), 1861.
- Uttal, T., Starkweather, S., Drummond, J. R., Vihma, T., Makshtas, A. P., Darby, L. S., Burkhart, J. F., Cox, C. J., Schmeisser, L. N., Haiden, T., Maturilli, M., Shupe, M. D., De Boer, G., Saha, A., Grachev, A. A., Crepinsek, S. M., Bruhwiler, L., Goodison, B., McArthur, B., Walden, V. P., Dlugokencky, E. J., Persson, P. O. G., Lesins, G., Laurila, T., Ogren, J. A., Stone, R., Long, C. N., Sharma, S., Massling, A., Turner, D. D., Stanitski, D. M., Asmi, E., Aurela, M., Skov, H., Eleftheriadis, K., Virkkula, A., Platt, A., Førland, E. J., Iijima, Y., Nielsen, I. E., Bergin, M. H., Candlish, L., Zimov, N. S., Zimov, S. A., O'Neill, N. T., Fogal, P. F., Kivi, R., Konopleva-Akish, E. A., Verlinde, J., Kustov, V. Y., Vasel, B., Ivakhov, V. M., Viisanen, Y., and Intrieri, J. M.: International Arctic Systems for Observing the Atmosphere: An International Polar Year Legacy Consortium, *Bull. Am. Meteorol. Soc.*, 97, 1033–1056, doi:[10.1175/BAMS-D-14-00145.1](https://doi.org/10.1175/BAMS-D-14-00145.1), 2016.
- Vaisala: HMT330 Series Humidity and Temperature Transmitters, <https://docs.vaisala.com/v/u/B210951EN-U/en-US>, Accessed: 2024-07-13, 2023a.
- Vaisala: PTU300 Combined Pressure, Humidity and Temperature Transmitter, <https://docs.vaisala.com/v/u/B210954EN-J/en-US>, Accessed: 2024-07-05, 2023b.
- Van Malderen, R., Brenot, H., Pottiaux, E., Beirle, S., Hermans, C., De Mazière, M., Wagner, T., De Backer, H., and Bruyninx, C.: A multi-site intercomparison of integrated water vapour observations for climate change analysis, *Atmos. Meas. Tech.*, 7, 2487–2512, doi:[10.5194/amt-7-2487-2014](https://doi.org/10.5194/amt-7-2487-2014), 2014.
- Vaquero-Martínez, J., Antón, M., Román, R., Cachorro, V. E., Wang, H., González Abad, G., and Ritter, C.: Water vapor satellite products in the Eu-

- ropean Arctic: An inter-comparison against GNSS data, *Sci. Total Environ.*, 741, 140 335, doi:[10.1016/j.scitotenv.2020.140335](https://doi.org/10.1016/j.scitotenv.2020.140335), 2020.
- Vavrus, S. J.: The Influence of Arctic Amplification on Mid-latitude Weather and Climate, *Curr. Clim. Change Rep.*, 4, 238–249, doi:[10.1007/s40641-018-0105-2](https://doi.org/10.1007/s40641-018-0105-2), 2018.
- Viceto, C., Gorodetskaya, I. V., Rinke, A., Maturilli, M., Rocha, A., and Crewell, S.: Atmospheric rivers and associated precipitation patterns during the ACLOUD and PASCAL campaigns near Svalbard (May–June 2017): case studies using observations, reanalyses, and a regional climate model, *Atmos. Chem. Phys.*, 22, 441–463, doi:[10.5194/acp-22-441-2022](https://doi.org/10.5194/acp-22-441-2022), 2022.
- Vihma, T., Kilpeläinen, T., Manninen, M., Sjöblom, A., Jakobson, E., Palo, T., Jaagus, J., and Maturilli, M.: Characteristics of Temperature and Humidity Inversions and Low-Level Jets over Svalbard Fjords in Spring, *Adv. Meteorol.*, 2011, 1–14, doi:[10.1155/2011/486807](https://doi.org/10.1155/2011/486807), 2011.
- Vihma, T., Screen, J., Tjernström, M., Newton, B., Zhang, X., Popova, V., Deser, C., Holland, M., and Prowse, T.: The atmospheric role in the Arctic water cycle: A review on processes, past and future changes, and their impacts, *J. Geophys. Res.: Biogeosci.*, 121, 586–620, doi:[10.1002/2015JG003132](https://doi.org/10.1002/2015JG003132), 2016.
- Walbröl, A.: Codes for: Atmospheric temperature, water vapour and liquid water path from two microwave radiometers during MOSAiC, Zenodo [software], doi:[10.5281/zenodo.6673957](https://doi.org/10.5281/zenodo.6673957), 2022.
- Walbröl, A.: Neural Network predictions and ERA5 reference of integrated water vapour, and temperature and specific humidity profiles based on simulated microwave radiometer observations, Zenodo [data set], doi:[10.5281/zenodo.10998146](https://doi.org/10.5281/zenodo.10998146), 2024a.
- Walbröl, A.: Information content estimation output for specific humidity profiles, Zenodo [data set], doi:[10.5281/zenodo.10997692](https://doi.org/10.5281/zenodo.10997692), 2024b.
- Walbröl, A.: Simulated microwave brightness temperatures based on two radiosoundings performed during the MOSAiC expedition, Zenodo [data set], doi:[10.5281/zenodo.11092210](https://doi.org/10.5281/zenodo.11092210), 2024c.
- Walbröl, A.: Codes for: Combining low and high frequency microwave radiometer measurements from the MOSAiC expedition for enhanced water vapour products, Zenodo [software], doi:[10.5281/zenodo.13750797](https://doi.org/10.5281/zenodo.13750797), 2024d.

- Walbröl, A. and Mech, M.: ERA5 based training, validation and evaluation data for retrievals combining 22-58 GHz with 175-340 GHz microwave radiometer measurements during MOSAiC, Zenodo [data set], doi:[10.5281/zenodo.10997365](https://doi.org/10.5281/zenodo.10997365), 2024.
- Walbröl, A., Crewell, S., Engelmann, R., Orlandi, E., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Maturilli, M., and Ebell, K.: Atmospheric temperature, water vapour and liquid water path from two microwave radiometers during MOSAiC, *Sci. Data*, 9, 534, doi:[10.1038/s41597-022-01504-1](https://doi.org/10.1038/s41597-022-01504-1), 2022.
- Walbröl, A., Ebell, K., Engelmann, R., Griesche, H., Radenz, M., Hofer, J., Althausen, D., and Crewell, S.: Brightness temperatures of the MiRAC-P microwave radiometer onboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.941407](https://doi.org/10.1594/PANGAEA.941407), 2022a.
- Walbröl, A., Orlandi, E., Crewell, S., and Ebell, K.: Integrated water vapour derived from the MiRAC-P microwave radiometer onboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.941470](https://doi.org/10.1594/PANGAEA.941470), 2022b.
- Walbröl, A., Engelmann, R., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Ebell, K., Mech, M., and Crewell, S.: Improved integrated water vapour derived from the combination of the microwave radiometers HAT-PRO and MiRAC-P aboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.968778](https://doi.org/10.1594/PANGAEA.968778), 2024a.
- Walbröl, A., Engelmann, R., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Ebell, K., Mech, M., and Crewell, S.: Improved temperature and humidity profiles derived from the combination of the microwave radiometers HAT-PRO and MiRAC-P aboard the Polarstern during the MOSAiC expedition, PANGAEA - Data Publisher for Earth & Environmental Science [data set], doi:[10.1594/PANGAEA.968760](https://doi.org/10.1594/PANGAEA.968760), 2024b.
- Walbröl, A., Griesche, H. J., Mech, M., Crewell, S., and Ebell, K.: Combining low- and high-frequency microwave radiometer measurements from the MOSAiC expedition for enhanced water vapour products, *Atmospheric Measurement Techniques*, 17, 6223–6245, doi:[10.5194/amt-17-6223-2024](https://doi.org/10.5194/amt-17-6223-2024), 2024.
- Wang, C., Graham, R. M., Wang, K., Gerland, S., and Granskog, M. A.: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and

- precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution, *The Cryosphere*, 13, 1661–1679, doi:[10.5194/tc-13-1661-2019](https://doi.org/10.5194/tc-13-1661-2019), 2019.
- Wang, D., Prigent, C., Kilic, L., Fox, S., Harlow, C., Jimenez, C., Aires, F., Grassotti, C., and Karbou, F.: Surface Emissivity at Microwaves to Millimeter Waves over Polar Regions: Parameterization and Evaluation with Aircraft Experiments, *J. Atmos. Oceanic Technol.*, 34, 1039–1059, doi:[10.1175/JTECH-D-16-0188.1](https://doi.org/10.1175/JTECH-D-16-0188.1), 2017.
- Wang, Z., Ding, Q., Wu, R., Ballinger, T. J., Guan, B., Bozkurt, D., Nash, D., Baxter, I., Topál, D., Li, Z., Huang, G., Chen, W., Chen, S., Cao, X., and Chen, Z.: Role of atmospheric rivers in shaping long term Arctic moisture variability, *Nat. Commun.*, 15, 5505, doi:[10.1038/s41467-024-49857-y](https://doi.org/10.1038/s41467-024-49857-y), 2024.
- Ware, R., Carpenter, R., Güldner, J., Liljegren, J., Nehr Korn, T., Solheim, F., and Vandenberghe, F.: A multichannel radiometric profiler of temperature, humidity, and cloud liquid, *Radio Sci.*, 38, 2002RS002856, doi:[10.1029/2002RS002856](https://doi.org/10.1029/2002RS002856), 2003.
- Wendisch, M., Brückner, M., Burrows, J., Crewell, S., Dethloff, K., Ebell, K., Lüpkes, C., Macke, A., Notholt, J., Quaas, J., Rinke, A., and Tegen, I.: Understanding Causes and Effects of Rapid Warming in the Arctic, *Eos*, doi:[10.1029/2017EO064803](https://doi.org/10.1029/2017EO064803), 2017.
- Wendisch, M., Brückner, M., Crewell, S., Ehrlich, A., Notholt, J., Lüpkes, C., Macke, A., Burrows, J. P., Rinke, A., Quaas, J., Maturilli, M., Schemann, V., Shupe, M. D., Akansu, E. F., Barrientos-Velasco, C., Bärfuss, K., Blechschmidt, A.-M., Block, K., Bougoudis, I., Bozem, H., Böckmann, C., Bracher, A., Bresson, H., Bretschneider, L., Buschmann, M., Chechin, D. G., Chylik, J., Dahlke, S., Deneke, H., Dethloff, K., Donth, T., Dorn, W., Dupuy, R., Ebell, K., Egerer, U., Engelmann, R., Eppers, O., Gerdes, R., Gierens, R., Gorodetskaya, I. V., Gottschalk, M., Griesche, H., Gryanik, V. M., Handorf, D., Harm-Altstädter, B., Hartmann, J., Hartmann, M., Heinold, B., Herber, A., Herrmann, H., Heygster, G., Höschel, I., Hofmann, Z., Hölemann, J., Hünerbein, A., Jafariserajehlou, S., Jäkel, E., Jacobi, C., Janout, M., Jansen, F., Jourdan, O., Jurányi, Z., Kalesse-Los, H., Kanzow, T., Käthner, R., Kliesch, L. L., Klingebiel, M., Knudsen, E. M., Kovács, T., Körtke, W., Krampe, D., Kretschmar, J., Kreyling, D., Kulla, B., Kunkel, D., Lampert, A., Lauer, M., Lelli, L., von Lerber, A., Linke, O., Löhnert, U., Lonardi, M., Losa, S. N., Losch, M., Maahn, M., Mech, M., Mei, L., Mertes, S., Metzner, E., Mewes, D., Michaelis, J., Mioche, G., Moser, M., Nakoudi, K., Neggers, R., Neuber,

- R., Nomokonova, T., Oelker, J., Papakonstantinou-Presvelou, I., Pätzold, F., Pefanis, V., Pohl, C., van Pinxteren, M., Radovan, A., Rhein, M., Rex, M., Richter, A., Risse, N., Ritter, C., Rostosky, P., Rozanov, V. V., Donoso, E. R., Saavedra Garfias, P., Salzmann, M., Schacht, J., Schäfer, M., Schneider, J., Schnierstein, N., Seifert, P., Seo, S., Siebert, H., Soppa, M. A., Spreen, G., Stachlewska, I. S., Stapf, J., Stratmann, F., Tegen, I., Viceto, C., Voigt, C., Vountas, M., Walbröl, A., Walter, M., Wehner, B., Wex, H., Willmes, S., Zanatta, M., and Zeppenfeld, S.: Atmospheric and Surface Processes, and Feedback Mechanisms Determining Arctic Amplification: A Review of First Results and Prospects of the (AC)3 Project, *Bull. Am. Meteorol. Soc.*, 104, E208–E242, doi:[10.1175/BAMS-D-21-0218.1](https://doi.org/10.1175/BAMS-D-21-0218.1), 2023.
- Willmes, S., Nicolaus, M., and Haas, C.: The microwave emissivity variability of snow covered first-year sea ice from late winter to early summer: a model study, *The Cryosphere*, 8, 891–904, doi:[10.5194/tc-8-891-2014](https://doi.org/10.5194/tc-8-891-2014), 2014.
- Winton, M.: Amplified Arctic climate change: What does surface albedo feedback have to do with it?, *Geophys. Res. Lett.*, 33, L03701, doi:[10.1029/2005GL025244](https://doi.org/10.1029/2005GL025244), 2006.
- Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., and Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance, *Appl. Phys. B*, 96, 201–213, doi:[10.1007/s00340-009-3365-7](https://doi.org/10.1007/s00340-009-3365-7), 2009.
- Woods, C. and Caballero, R.: The Role of Moist Intrusions in Winter Arctic Warming and Sea Ice Decline, *J. Climate*, 29, 4473–4485, doi:[10.1175/JCLI-D-15-0773.1](https://doi.org/10.1175/JCLI-D-15-0773.1), 2016.
- You, C., Tjernström, M., and Devasthale, A.: Warm-Air Advection Over Melting Sea-Ice: A Lagrangian Case Study, *Boundary Layer Meteorol.*, 179, 99–116, doi:[10.1007/s10546-020-00590-1](https://doi.org/10.1007/s10546-020-00590-1), 2021.
- You, C., Tjernström, M., and Devasthale, A.: Warm and moist air intrusions into the winter Arctic: a Lagrangian view on the near-surface energy budgets, *Atmos. Chem. Phys.*, 22, 8037–8057, doi:[10.5194/acp-22-8037-2022](https://doi.org/10.5194/acp-22-8037-2022), 2022a.
- You, C., Tjernström, M., Devasthale, A., and Steinfeld, D.: The Role of Atmospheric Blocking in Regulating Arctic Warming, *Geophys. Res. Lett.*, 49, doi:[10.1029/2022GL097899](https://doi.org/10.1029/2022GL097899), 2022b.
- Yu, Y., Xiao, W., Zhang, Z., Cheng, X., Hui, F., and Zhao, J.: Evaluation of 2-m Air Temperature and Surface Temperature from ERA5 and ERA-I Using

- Buoy Observations in the Arctic during 2010–2020, *Remote Sens.*, 13, 2813, doi:[10.3390/rs13142813](https://doi.org/10.3390/rs13142813), 2021.
- Zhang, T., Stamnes, K., and Bowling, S. A.: Impact of the Atmospheric Thickness on the Atmospheric Downwelling Longwave Radiation and Snowmelt under Clear-Sky Conditions in the Arctic and Subarctic, *J. Climate*, 14, 920–939, doi:[10.1175/1520-0442\(2001\)014<0920:IOTATO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0920:IOTATO>2.0.CO;2), 2001.
- Zhou, W., Leung, L. R., and Lu, J.: Steady threefold Arctic amplification of externally forced warming masked by natural variability, *Nat. Geosci.*, 17, 508–515, doi:[10.1038/s41561-024-01441-1](https://doi.org/10.1038/s41561-024-01441-1), 2024.
- Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, *Q. J. R. Meteorolog. Soc.*, 141, 563–579, doi:[10.1002/qj.2378](https://doi.org/10.1002/qj.2378), 2015.
- Zängl, G., Reinert, D., and Prill, F.: Grid refinement in ICON v2.6.4, *Geosci. Model Dev.*, 15, 7153–7176, doi:[10.5194/gmd-15-7153-2022](https://doi.org/10.5194/gmd-15-7153-2022), 2022.

Acknowledgements

After all, a dissertation is teamwork. First and foremost, I am deeply grateful for the excellent supervision and support of Prof. Dr. Susanne Crewell and Dr. Kerstin Ebell throughout the years. They provided new ideas and perspectives on the current work during countless discussions and helped me connect findings. I would like to extend my gratitude to Dr. Gunnar Spreen for joining my thesis advisory committee and providing new insights.

I highly appreciate that I could join two field campaigns: HALO-($\mathcal{A}\mathcal{C}$)³ and the RV *Polarstern* cruise ATWAICE. I express my gratitude to Dr. Davide Ori with whom I've shared a hut for six weeks. You are a great cook and the discussions about work and life were always enjoyable. I would also like to thank Prof. Dr. Manfred Wendisch and Susanne for the opportunity to plan and execute the Polar Low research flight during HALO-($\mathcal{A}\mathcal{C}$)³. It was a great experience and I learned a lot! Special thanks go to Susanne, Henning Dorff, Benjamin Kirbus, Sebastian Becker and Dr. Janosch Michaelis whose help was essential to complete the synoptic overview paper for the HALO-($\mathcal{A}\mathcal{C}$)³ campaign. I could count on you, even when having discussions after common working hours and on weekends. Joining the ATWAICE cruise was only possible thanks to Kerstin and Dr. Mario Mech. You know how much I love fieldwork and the Arctic environment. I would like to thank Dr. Janna Rückert, Dr. Hannah Niehaus and Gunnar for the great time on RV *Polarstern* during ATWAICE. It was an awesome and refreshing experience, and I learned a lot about snow and sea ice thanks to Janna and Gunnar.

I also had a great time in Boulder with Benjamin, Dr. Jan Chylik, Niklas Schnierstein and Prof. Dr. Roel Neggers. The unconventional hike to Green Mountain was quite a memorable experience. Next time, we might even consider using a path on our way down.

I appreciate the assistance of Dr. Carola Barrientos-Velasco for setting up the T-CARS simulations. I also thank Dr. Hannes Griesche for providing the experimental Cloudnet and low-level stratus data set. I extend my gratitude to the countless developers of open-source Python tools. Without your dedication, daily work would sometimes be much less efficient. I would like to thank my office mates Nils Risse and Dr. Dwaipayan Chatterjee (DC). The discussions about results were always helpful. I also appreciate DC's words on work and life. I want to express my gratitude to Kerstin, Susanne, DC, Christian Buhren, my brother and my parents for proofreading parts of this thesis. Your input and comments helped a lot.

A final and special thank you is reserved for my family and friends. I cannot express how grateful I am for the unconditional support of my parents and my grandmother in all those years of my life! Also my brother's patience during the past months is highly appreciated. A special thanks to my friends for the great laughs during the long nights we spent together (online and in person). To the active members of the Speläologische Arbeitsgemeinschaft Hessen e.V.: Big thanks for the unforgettable times below and above the surface, for the family-like atmosphere and great friendship!

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen

- Walbröl, A. et al., 2022: Atmospheric temperature, water vapour and liquid water path from two microwave radiometers during MOSAiC, *Scientific Data*, 9, 534, [10.1038/s41597-022-01504-1](https://doi.org/10.1038/s41597-022-01504-1).
- Walbröl, A. et al. 2024: Combining low- and high-frequency microwave radiometer measurements from the MOSAiC expedition for enhanced water vapour products, *Atmospheric Measurement Techniques*, 17, 6223-6245, [10.5194/amt-17-6223-2024](https://doi.org/10.5194/amt-17-6223-2024).

Andreas Walbröl

Köln, den 13. Februar 2025