

Benchmark Quality Processing of Radio Occultation Data with Integrated Uncertainty Propagation

Jakob Schwarz

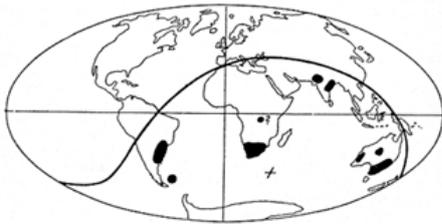
July 2018

Financially supported by



The **Wegener Center for Climate and Global Change** combines as an interdisciplinary, internationally oriented research institute the competences of the University of Graz in the research area “Climate, Environmental and Global Change”. It brings together, in a dedicated building close to the University central campus, research teams and scientists from fields such as geo- and climate physics, meteorology, economics, geography, and regional sciences. At the same time close links exist and are further developed with many cooperation partners, both nationally and internationally. The research interests extend from monitoring, analysis, modeling and prediction of climate and environmental change via climate impact research to the analysis of the human dimensions of these changes, i.e., the role of humans in causing and being effected by climate and environmental change as well as in adaptation and mitigation. (more information at www.wegcenter.at)

The present report is the result of a PhD thesis work completed in June 2017 (updated by a published version of a peer-reviewed article becoming available May 2018). The work was funded by the Austrian Aeronautics and Space Agency of the Austrian Research Promotion Agency (FFG-ALR) under the ASAP projects OPSCLIMPROP and OPSCLIMTRACE (ASAP–Austrian Space Applications Programme sponsor is the Austrian Ministry for Transport, Innovation, and Technology BMVIT) as well as by the European Space Agency (ESA) under project MMVaIRO-E.



Alfred Wegener (1880–1930), after whom the Wegener Center is named, was founding holder of the University of Graz Geophysics Chair (1924–1930). In his work in the fields of geophysics, meteorology, and climatology he was a brilliant scientist and scholar, thinking and acting in an interdisciplinary way, far ahead of his time with this style. The way of his ground-breaking research on continental drift is a shining role model—his sketch on the relations of continents based on traces of an ice age about 300 million years ago (left) as basis for the Wegener Center Logo is thus a continuous encouragement to explore equally innovative ways: *paths emerge in that we walk them* (Motto of the Wegener Center).

Wegener Center Verlag • Graz, Austria

© 2018 All Rights Reserved.

Selected use of individual figures, tables or parts of text is permitted for non-commercial purposes, provided this report is correctly and clearly cited as the source. Publisher contact for any interests beyond such use: wegcenter@uni-graz.at.

ISBN 978-3-9504501-5-6

July 2018

Contact: Dr. Jakob Schwarz
schwarz.jakob@gmx.at

Wegener Center for Climate and Global Change
University of Graz
Brandhofgasse 5
A-8010 Graz, Austria
www.wegcenter.at

Dissertation
zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften

Benchmark Quality Processing of Radio Occultation Data with Integrated Uncertainty Propagation

MAG. JAKOB SCHWARZ, B.A.

Betreuer:

Univ.-Prof. Mag. Dr. Gottfried Kirchengast

Mitbetreuer:

Mag. Dr. Marc Schwärz

This thesis was submitted in June 2017 and modified in May 2018. Chapter 3 was updated to the final published paper version of May 2018 and minor editorial updates and corrections were implemented at other places. This report does not describe later advancements or additions to the reference occultation processing system that was at the core of this work.

Abstract

GNSS radio occultation (RO) observations have the potential to provide climate data records of benchmark quality, due to the unique properties of the RO technique. RO records are highly accurate, long-term stable, globally available and provide Essential Climate Variables (ECVs) for the thermodynamic state of the free atmosphere. These ECVs, pressure, temperature and tropospheric water vapor profiles, can be derived from the raw RO observations through an atmospheric profiles retrieval chain. To realize the climate benchmark potential, these RO retrievals need to be very accurate and the remaining uncertainties quantified and traced throughout the retrieval chain from raw observations to the ECVs. The new Reference Occultation Processing System (rOPS) at the Wegener Center aims to deliver such an accurate chain with integrated uncertainty propagation.

This thesis contributed to the development and implementation of algorithms in the rOPS to propagate uncertainty, first, from RO excess phase profiles to atmospheric bending angle profiles, then further through refractivity to dry-air profiles, and finally, using also background profiles, to moist-air thermodynamic profiles (the ECVs). In the context of this thesis, propagation of estimated systematic and random uncertainties, vertical error correlations and resolution estimates, and also observation-to-background weighting ratio profiles was implemented into the rOPS retrieval chain.

Results from the implemented covariance propagation were validated using Monte Carlo ensemble methods. The algorithm performance was demonstrated by test-day ensembles of simulated and real RO event data from the satellite missions CHAMP, COSMIC, and MetOp, which show that the new uncertainty estimation chain performs robustly and delivers reliable results. This thesis thus contributes to establishing the benchmark capability of the rOPS for the benefit of climate change monitoring and other applications.

Zusammenfassung

AUFGRUND der spezifischen Eigenschaften der Methode der GNSS Radio-Okkultation (RO) haben RO Beobachtungen das Potenzial Klimadaten in Referenz-Qualität zur Verfügung zu stellen. RO Messungen zeichnen sich durch hohe Genauigkeit, Langzeitstabilität und globale Abdeckung aus. Wesentliche Klimavariablen (ECVs) bezüglich des thermodynamischen Zustandes der freien Atmosphäre können davon abgeleitet werden. Diese ECVs, nämlich vertikale Druck-, Temperatur- und Feuchteprofile, werden mithilfe eines Prozessierungssystems aus den RO Rohdaten errechnet. Um Referenz-Klimabeobachtungen bereitstellen zu können, muss die Genauigkeit der Prozessierung sehr hoch sein. Verbleibende Unsicherheiten müssen quantifiziert und von den Rohdaten durch die gesamte Prozessierungskette bis zu den ECVs fortgepflanzt werden. Das neue Referenz-Okkultations-Prozessierungssystem (rOPS) am Wegener Center hat zum Ziel, eine solche hoch genaue Prozessierung mit integrierter Unsicherheitsfortpflanzung zu realisieren.

Diese Dissertation hat zur Entwicklung und rOPS Implementierung von Algorithmen beigetragen, um Unsicherheitsprofile zuerst von Exzess-Phasen zu Brechungswinkeln, dann über Refraktivitäten zu trockenen Parametern, und letztlich mithilfe externer Zusatzinformationen zu den thermodynamischen ECVs fortzupflanzen. Im Rahmen der Dissertation wurden die Algorithmen zur Fortpflanzung von Profilen zufälliger und systematischer Unsicherheiten, vertikaler Korrelationen, vertikaler Auflösung, und dem Gewichtungsverhältnis zwischen Messung und Zusatzinformation entwickelt und implementiert.

Die korrekte Berechnung der Resultate der Fortpflanzung wurden mittels Monte-Carlo Validierung unabhängig geprüft. Die Robustheit des Algorithmus wurde durch Prozessierung von realen Messdaten für verschiedene RO Satellitenmissionen demonstriert. Diese Dissertation trägt damit zur Etablierung der Referenzfähigkeit des rOPS bei, zum Nutzen der Klimabeobachtung und anderer Anwendungen.

Acknowledgments

“Keep your eyes on the stars and
your feet on the ground.”

(Theodore Roosevelt)

FOLLOWING President Roosevelt’s advice has occasionally presented quite a challenge during the course of this PhD. The stars seemed particularly far away and the ground was rocky and the path forward strenuous. Now that the stars have come close, I want to use the opportunity to express my gratitude to some of the people who’s support was indispensable for the progress over the last years.

First and foremost, I would like to thank my supervisor, co-founder and director of the Wegener Center for Climate and Global Change, Univ.-Prof. Mag. Dr.rer.nat. Gottfried Kirchengast for the opportunity to take part in the endeavour of solving the global climate monitoring problem, but also more generally for the privilege of being able to engage in an interesting and meaningful (and paid) activity over multiple years. I also want to thank for the immense amount of time invested in getting this thesis to fly, and for the invaluable advice and support provided, not just with respect to the thesis, but also major life decisions along the way.

I am particularly grateful for the comprehensive support provided by Dr. Marc Schwärz, be it with regard to software and programming issues, or just prioritizing my time based on the 80/20 rule. It has been a joy working with you on the rOPS. I am sure, with your expertise behind it, a great (traceable) future of the rOPS lies ahead.

I would like to express my gratitude to Dipl.-Ing. Dr. Veronika Proschek, for her advice and some real life (beyond svn and git) teamwork, helping to reduce the monotony of a programmers daily routine significantly. I also want to thank my roommates, particularly Dipl.-Ing. Jürgen Fuchsberger for his invaluable advice and assistance regarding IT issues, but also for personal advice when it was really needed.

I learned a lot about GNU Emacs, L^AT_EX, Python and other useful software tools from Dr. Florian Ladstädter, who was also willing to help out when others would have probably said “rtfm”. Thank you!

I would like to thank all of my past and present colleagues at the Wegener Center for Climate and Global Change, for the privilege of working with you in this incredibly enjoyable atmosphere. Some of you have become close friends over the course of the last seven years (and then typically moved to Norway).

This thesis received appreciated financial support from the Austrian Aeronautics and

Space Agency of the Austrian Research Promotion Agency (Österreichische Forschungsförderungsgesellschaft) (FFG-ALR) under projects OPSCLIMPROP and OPSCLIMTRACE and by the European Space Agency (ESA) under project MMValRO-E. Data used in the thesis were provided generously by UCAR/CDAAC, Boulder, whom I thank for access to their RO excess phase and orbit data (available at <http://cdaac-www.cosmic.ucar.edu/>) as well as ECMWF, Reading, for access to their analysis and forecast data (available at <http://www.ecmwf.int/en/forecasts/datasets>) This document is built on a ready-made L^AT_EX template created by Mag. Michael Pock.

Finally I would like to thank my family, i.e., my parents Christine and Friedrich Schwarz, and my sister Judith Schwarz, for the immense amount of love and support I received over the last 30+ years, making this PhD, but also other ambitious endeavours of mine possible.

Preface

THIS PhD thesis is based on the following publications which I worked on as leading author over the course of the past years:

J. C. Schwarz et al. (2018a). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles’. In: *Atmos. Meas. Tech.*, 11, pp. 2601–2631. DOI: 10.5194/amt-11-2601-2018

J. C. Schwarz et al. (2017). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles’. In: *Earth Space Sci.*, 4.4, pp. 200–228. DOI: 10.1002/2016EA000234

J. C. Schwarz et al. (2018b). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis’. For: *Rem. Sens. Environ.*, manuscript in preparation

Contents

Abstract	iii
Zusammenfassung	v
Acknowledgments	vii
Preface	ix
Introduction	1
1 Climate Benchmark Data and the Role of Uncertainty Propagation	5
1.1 The Global Climate Observing System (GCOS)	5
1.1.1 Development and Status of GCOS	5
1.1.2 Essential Climate Variables (ECVs)	7
1.2 Climate Benchmark Observations for the Free Atmosphere	8
1.2.1 The Climate Monitoring Problem	8
1.2.2 Properties of Climate Benchmark Observations	9
1.3 The Global Navigation Satellite System (GNSS) Radio Occultation Technique	11
1.3.1 Historic Development of Radio Occultation (RO)	11
1.3.2 Deriving ECVs through GNSS RO	11
1.4 GNSS RO for Climate Benchmark Observations	14
1.4.1 Properties of Benchmark Observations from RO	14
1.4.2 ECVs from Fundamental Climate Data Records	16
1.5 The Reference Occultation Processing System (rOPS) - Ensuring Traceability	17
1.5.1 Motivation for the Development of the rOPS	17
1.5.2 Development History of the rOPS	17
1.5.3 Basic Structure of the rOPS	18
1.5.4 rOPS Innovations	22
1.5.5 Ensuring Traceability	26
1.6 The rOPS Development Context	27
2 Estimation, Propagation and Validation of Measurement Uncertainty	29
2.1 The Guide to the Expression of Uncertainty in Measurement (GUM) . . .	29
2.1.1 Motivation for the Establishment of the GUM	30
2.1.2 Terminology for the Measurement Process	30

2.1.3	Evaluating Uncertainty with Type A and Type B Methods	35
2.2	Measurement Models and Uncertainty Propagation Algorithms	37
2.2.1	Univariate, Linear Measurement Model	37
2.2.2	Multivariate, Linear Measurement Model	39
2.2.3	Non-Linear Measurement Models	40
2.2.4	Implicit Measurement Models	40
2.2.5	Validation of Propagated Uncertainties	40
2.3	Uncertainty Propagation in the rOPS	41
2.3.1	Sources of Uncertainties in the rOPS	41
2.3.2	Estimated Random and Systematic Uncertainties	42
2.3.3	Operator Types and Corresponding Propagation Rules	44
3	Integrating uncertainty propagation in GNSS radio occultation re-	
	trieval: from excess phase to atmospheric bending angle profiles	47
3.1	Abstract	47
3.2	Introduction	48
3.3	Methods and data	55
3.3.1	Methods	55
3.3.2	Data sources and preparation	57
3.4	Algorithm sequence and example results	62
3.4.1	Doppler shift retrieval	62
3.4.2	Bending angle retrieval	65
3.4.3	Atmospheric bending angle derivation	67
3.5	Algorithm validation	70
3.6	Performance demonstration	74
3.7	Conclusions	78
3.8	Acknowledgements	78
4	Integrating uncertainty propagation in GNSS radio occultation re-	
	trieval: from bending angle to dry-air atmospheric profiles	81
4.1	Abstract	81
4.2	Introduction	82
4.2.1	The RO Method	82
4.2.2	Essential Climate Variables from RO	83
4.2.3	Reference Occultation Processing System (rOPS)	84
4.2.4	Study Approach and Outline	86
4.3	Methods and Data	88
4.3.1	Methods	88
4.3.2	Data Sources	92
4.3.3	Data Preparation	96

4.4	Algorithm Sequence and Example Results	98
4.4.1	High Altitude Initialization	98
4.4.2	Refractivity Retrieval	99
4.4.3	Dry-Air Retrieval	99
4.5	Performance Demonstration	101
4.5.1	Algorithm Validation Using a Monte Carlo Method	102
4.5.2	Algorithm Performance for Individual Profiles	104
4.5.3	Statistical Performance Evaluation with Real Data	106
4.6	Conclusions	112
4.7	Acknowledgments	112
5	Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis	113
5.1	Abstract	113
5.2	Introduction	114
5.3	Methods and Data	117
5.3.1	Uncertainty Propagation Methods	117
5.3.2	Data Sources and Preparation	122
5.4	Algorithm Sequence and Example Results	124
5.4.1	Direct Retrievals of Temperature and Specific Humidity	125
5.4.2	Optimal Estimation of Temperature and Specific Humidity	128
5.4.3	Optimal Estimation of other Thermodynamic Variables	129
5.5	Algorithm Validation	130
5.6	Performance Demonstration	131
5.7	Conclusions	132
5.8	Acknowledgements	133
	Summary and Conclusions	135
	Appendix A Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles	137
A.1	Algorithm Description	137
A.1.1	Doppler shift retrieval	138
A.1.2	Bending angle retrieval	143
A.1.3	Atmospheric bending angle derivation	150
A.2	Variance propagation for comparison	154

Appendix B Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles	155
B.1 Algorithm Description	155
B.1.1 High Altitude Initialization	155
B.1.2 Refractivity Retrieval	157
B.1.3 Dry-Air Retrieval	160
Appendix C Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis	165
C.1 Algorithm Description	165
C.1.1 Direct Retrievals of Temperature and Specific Humidity	166
C.1.2 Optimal Estimation of Temperature and Specific Humidity	171
C.1.3 Derivation of Further Thermodynamic Variables	173
C.2 Proof of Simplification of Equation 5.2	176
List of Figures	179
List of Tables	181
Bibliography	183

Introduction

According to the Intergovernmental Panel on Climate Change (IPCC) the warming of the Earth's climate system is unequivocal. The Earth's atmosphere and oceans have warmed, the amounts of snow and ice decreased, sea levels have risen. The increase in uptake of energy in the climate system is due to additional radiative forcing caused mainly by an increase in the concentration of atmospheric greenhouse gases (GHGs), particularly carbon dioxide (CO₂) emitted by anthropogenic combustion processes (IPCC 2013).

In light of these challenges and in order for policy makers to take the adequate policy responses to adapt to and mitigate these changes in the climate system, it is the responsibility of the climate sciences to provide high-quality observations of the current and historic state of the climate system, and create reliable projections of the future state of the climate system, based on these observations. Timeliness of the data-provision is crucial as the window for action to prevent dangerous climate change is closing quickly (Rogelj et al. 2016).

To address these demands the World Meteorological Organization (WMO) , together with the Intergovernmental Oceanographic Commission (IOC) of the United Nations Educational Scientific and Cultural Organization (UNESCO), the United Nations Environment Programme (UNEP), and the International Council for Science (ICSU) established the Global Climate Observing System (GCOS). The objective of GCOS is to provide comprehensive information on the total climate system, including a range of physical, chemical and biological properties, and atmospheric, oceanic, hydrological, cryospheric and terrestrial processes and to ensure that the observations and information needed to address climate-related issues are obtained and made available to all potential users (GCOS 2015).

The GCOS consists of an in-situ and a space-based remote sensing component. For observations of the free atmosphere and to ensure global coverage, the space-based component is of crucial importance and although the generation and supply of products derived from space-based observations have progressed, and increasing attention has been paid to product quality and uncertainty and their documentation, some deficiencies remain (Bojinski et al. 2014; GCOS 2015).

Observations of the free atmosphere were historically aimed at aviation, weather research and weather forecasting, and thus the observing systems have shortcomings from

a climate perspective. For climate purposes the accuracy, long-term stability, and data homogeneity were not sufficient in the respective observation records (Karl et al. 1995). Thus, *climate benchmark data* records, i.e., data records which are of global coverage, of high accuracy and long-term stability, which are tied to irrefutable standards, and measure Essential Climate Variables (ECVs), are of crucial importance (GCOS 2015; NRC 2007).

Thanks to its unique properties, the Radio Occultation (RO) technique has the potential to deliver these ECV data records for thermodynamic parameters of the free atmosphere, if the uncertainties of the retrieved ECVs can be traced to the uncertainties of the RO raw data. This requires that the RO processing system, or *retrieval*, deriving the ECVs from the raw data, is highly accurate and that uncertainties are propagated through the processing, such that the claimed accuracy of the ECV data, expressed by the uncertainties provided, can be traced back to the raw data (Kirchengast et al. 2016a).

The Reference Occultation Processing System (rOPS) (Kirchengast et al. 2017b), developed at the Wegener Center for Climate and Global Change (WEGC), aims to establish such a fully traceable processing (Kirchengast et al. 2016a; Kirchengast et al. 2016b). Among other aspects, the rOPS provides fully integrated uncertainty propagation from raw GNSS orbits and RO excess phase measurements to atmospheric ECVs along the processing chain.

In this thesis the rOPS uncertainty propagation algorithm from excess phase profiles to atmospheric ECVs is introduced, and its performance demonstrated showing the first (validated) results from the implemented propagation algorithm in the rOPS.

The thesis is structured into the following chapters:

Chapter 1 provides an introduction to the GCOS, the nature and importance of and the criteria for climate benchmark data, why the GNSS RO technique has the potential to provide benchmark data, if RO retrievals fulfill the traceability criteria. It also introduces the rOPS, with its aim to be traceable through an integrated uncertainty propagation.

Chapter 2 gives an introduction to the uncertainty framework of the Guide to the expression of Uncertainty in Measurement (GUM), the uncertainty propagation rules depending on the characteristics of the retrieval operators and examples for specific operators and propagation equations applied in the rOPS retrieval, to propagate uncertainties from the accurate raw data and excess phase measurements to thermodynamic profiles (ECVs).

Chapter 3 presents work submitted for publication about the actual propagation of uncertainty in the rOPS retrieval, from excess phase measurements to atmospheric bending angle profiles (L1b processing).

Chapter 4 presents published work about the rOPS uncertainty propagation from atmospheric bending angle profiles to dry air variables (L2a processing).

Chapter 5 presents work to be published about the propagation of uncertainty from dry air variables to thermodynamic profiles (ECVs) through the rOPS moist air retrieval (L2b processing).

A summary provides conclusions regarding uncertainty propagation in the rOPS and an outlook concerning remaining and potential future developments on the way to achieving climate benchmark capability with the rOPS.

Appendix A to Appendix C provide detailed step-by-step descriptions of the algorithms developed for and implemented in the rOPS from excess phase to moist air variables (i.e., L1b, L2a and L2b processing).

Climate Benchmark Data and the Role of Uncertainty Propagation

1.1 The Global Climate Observing System (GCOS)

1.1.1 Development and Status of GCOS

When the Second World Climate Conference in 1990 in Geneva, hosted by the World Meteorological Organization (WMO), reviewed the findings in the first assessment report of the freshly founded Intergovernmental Panel on Climate Change (IPCC), the need for a greatly improved global climate observing system was evident. Contemporary observation systems for monitoring the climate system were deemed inadequate for operational and research purposes and were at the time deteriorating in industrialized and in developing countries. The World Climate Research Programme (WCRP), founded at the First World Climate Conference in 1979, had come to the same conclusion earlier and provided compelling evidence that an integrated cross-domain (between atmosphere, ocean and land) observing system as a basis for monitoring, understanding and predicting variability and anthropogenic influence on the climate system was required. The Second World Climate Conference followed these conclusions and established the Global Climate Observing System (GCOS) (Houghton et al. 2012).

GCOS was sponsored by the WMO, the Intergovernmental Oceanographic Commission (IOC), the United Nations Environment Programme (UNEP) and the International Council for Science (ICSU) and officially founded in 1992 by a memorandum of understanding among these organizations. Investments into GCOS remained low and the quality of the observing systems did not improve until the late 90s, when the United Nations Framework Convention on Climate Change (UNFCCC) — also established at the Second World Climate Conference — became the *de facto* sponsor of GCOS and also pushed for its *climate monitoring principles* (Houghton et al. 2012). Particularly with

the 'GCOS Implementation Plan' (GCOS 2010b) the focus shifted to meeting the needs of the UNFCCC and the commitments of the Parties under Article 5 of the convention. Figure 1.1 shows the relationship between Global Climate Observing System (GCOS), the WCRP, the IPCC and the UNFCCC.

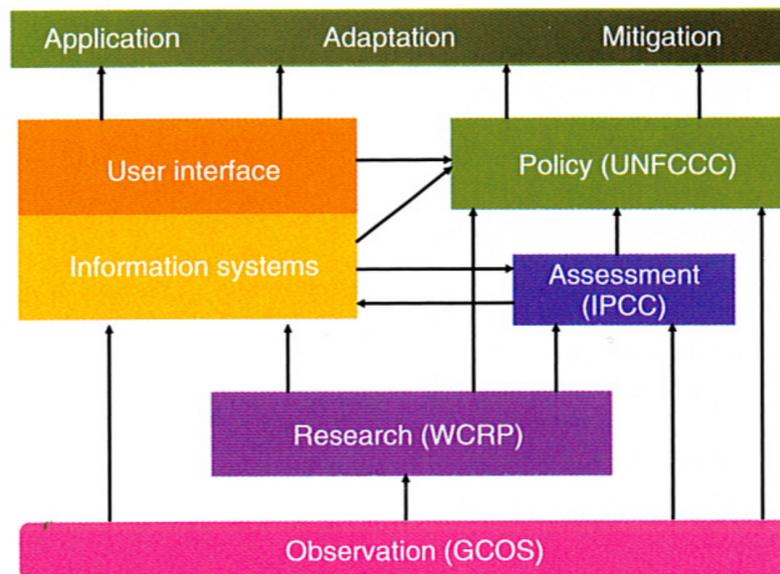


Figure 1.1: The relation of GCOS to the WCRP the Intergovernmental Panel on Climate Change (IPCC) and the United Nations Framework Convention on Climate Change (UNFCCC) (source: Houghton et al. 2012).

GCOS is now aiming to be a long-term, user-driven operational system for providing the observations needed for monitoring the climate system, detecting and attributing climate change, assessing impacts of, and supporting adaptation to, climate variability and change. GCOS aims to provide climate system information on physical, chemical and biological properties, in the atmospheric, oceanic, terrestrial, hydrologic, and cryospheric domains (GCOS 2015).

It is not a single, centrally managed observing system, but rather a *system of systems* comprising multiple climate-relevant observing, data-management, product-generation and data-distribution systems. It is built on the WMO Integrated Global Observing System, the IOC-led Global Ocean Observing System and various land-surface observing systems. It has in-situ and remote sensing (space-based) components. Also climate monitoring done by other programmes concerned with particular components of the climate system are incorporated (GCOS 2015).

1.1.2 Essential Climate Variables (ECVs)

To achieve a more coordinated approach to observing the climate on a global scale, the concept of the Essential Climate Variables (ECVs) was introduced in 2003 (GCOS 2003), i.e., principal physical variables or combinations of variables that would be most suitable for long-term climate monitoring. According to the concept, Essential Climate Variables (ECVs) are identified based on the following criteria (Bojinski et al. 2014; GCOS 2010b):

Relevance: The variable is critical for characterizing the climate system and its changes.

Feasibility: Observing or deriving the variable on a global scale is technically feasible and scientifically understood.

Cost effectiveness: Generating and archiving data on the variable is affordable, mainly relying on coordinated observing systems using proven technology.

The GCOS ECVs are listed in Figure 1.2, organized by domain.

OCEANIC	ATMOSPHERIC	TERRESTRIAL
Surface (10) Sea-surface temperature Sea-surface salinity Sea level Sea state Sea ice Surface current Ocean colour Carbon dioxide partial pressure Ocean acidity Phytoplankton	Composition (3) Carbon dioxide Methane and other long-lived greenhouse gases Ozone and Aerosol supported by their precursors	Biological/Ecological (6) Land cover FAPAR Leaf area index Above ground biomass Soil carbon Fire disturbance
Sub-surface (8) Temperature Salinity Current Nutrients Carbon dioxide partial pressure Ocean acidity Oxygen Tracers	Upper-air (5) Temperature Wind speed and direction Water vapour Cloud properties Earth radiation budget (including solar irradiance)	Hydrological (5) River discharge Water use Ground water Lakes Soil moisture
	Surface (6) Air temperature Wind speed and direction Water vapour Pressure Precipitation Surface radiation budget	Cryospheric (4) Snow cover Glaciers and ice caps Ice sheets Permafrost
		Other (1) Albedo

Figure 1.2: The GCOS Essential Climate Variables in the three domains atmosphere, ocean, land (source: Houghton et al. 2012).

Another contribution to systematize climate observations were the GCOS climate monitoring principles, adopted in 2003, that addressed problems with accuracy, long-term

stability, data-homogeneity and continuity of records, sampling and calibration, to ensure effective climate monitoring systems (GCOS 2010b).

1.2 Climate Benchmark Observations for the Free Atmosphere

1.2.1 The Climate Monitoring Problem

GCOS generally aims to establish tiered observing networks and constellations. For example, in-situ networks should be built up by three layers

Comprehensive networks that provide data of general quality with the highest spatial and temporal resolution, and the shortest latency.

Baseline networks with a limited number of globally distributed selected locations that provide long-term high-quality data records and with a greater degree of monitoring and management than comprehensive networks.

Reference networks with sparse coverage, but highest-quality observations, metrologically traceable to international standards, with well-quantified uncertainties (throughout this thesis the term *benchmark* and the WMO term *reference* are used interchangeably).

For the space-based component the same concepts apply, but no reference mission traceable to international standards could as of yet be established, with little progress in recent years (GCOS 2015).

This problem, also termed the *climate monitoring problem*, stems from the fact that historically, observations of the free atmosphere (i.e. starting at about 2 km above the ground) have been performed for aviation and for weather research and forecasting, and were consequentially not adequate for climate monitoring purposes. Most space-based remote sensing observing systems had inadequate vertical resolution and difficulties to deliver continuous data records due to drifting orbits and short lifespans of their individual satellites. In situ measurements like from the global radiosonde network had poor spatial and temporal coverage and an insufficiently characterized accuracy, because the measurements were not tied to absolute Systeme International (SI) standards. The long-term stability of the historical radiosonde data record was compromised by changes in instrumentation and observing methods, undertaken without sufficient overlap, intercomparison and documentation. These shortcomings severely limit the utility of these data records for detecting and quantifying climate trends in the troposphere and stratosphere (Bodeker et al. 2016; GCOS 2007).

So in 2007, the growing need for high quality upper-air climate observations lead to the establishment of a reference measurement network for the free atmosphere. The GCOS

Reference Upper Air Network (GRUAN) is designed to provide accurate climate data records not only at the surface of the Earth but throughout the atmospheric column into the stratosphere. The backbone of GCOS Reference Upper Air Network (GRUAN) are radiosondes, launched from sites worldwide (GCOS 2007), but to date the number of certified GRUAN launch sites remains small (Bodeker et al. 2016).

To achieve global coverage and sufficient sounding density to capture unbiased averages of the climate system a reference mission would thus have to be satellite-based (unlike GRUAN) (CEOS 2012).

1.2.2 Properties of Climate Benchmark Observations

To be considered a climate benchmark observation, such a satellite-based climate monitoring system would need to ensure the establishment of *global, highly accurate, long-term stable* climate records, which are tied to irrefutable standards (*traceability*) and measure significant indicators to describe the climate system, such as the *ECVs* (GCOS 2013; Leroy et al. 2006; NRC 2007).

Global coverage is important, because geographically insufficient coverage of the global climate state can lead to sampling errors in the observations. The varying accessibility of different regions on the planet gives rise to a very inhomogeneous global distribution of in-situ measurements. For space-based observations, the geographic coverage depends mostly on the inclination and altitude of the satellites orbits. A satellite in a polar orbit is able to cover every part of the Earth as it progresses along its orbit (Borowitz 2014).

Accuracy is crucial, because changes in the climate occur over long time periods of time and are often very small compared to temperatures at a single location, that may fluctuate by multiple degrees during just one day. According to GCOS (2010a) accuracy is

measured by the bias or systematic error of the data, i.e. the difference between the short-term average measured value of a variable and the true value. The short-term average is the average of a sufficient number of successive measurements of the variable under identical conditions, such that the random error is negligible relative to the systematic error. The latter can be introduced by instrument biases or through the choice of remote sensing retrieval schemes.

Stability on the other hand is an important issue, because climate measurements require data records on time scales much longer than the typical lifespan of one individual satellite. To allow for cross-calibration, overlaps between old and new sensors are

necessary. This helps to ensure that the level of uncertainty does not increase over time. GCOS (2010a) defines stability as

the extent to which accuracy remains constant with time. Stability is measured by the maximum excursion of the short-term average (e.g., daily, monthly, seasonal) measured value of a variable under identical conditions over the long term, e.g. a decade. The smaller the maximum excursion, the greater the stability of the dataset.

Traceability is a metrological concept ensuring that the result of a measurement by an instrument is related to a known standard through an unbroken chain of comparisons. This can also be used to determine an instrument's precision, accuracy and stability (GCOS 2010a). According to the International Vocabulary of Metrology (VIM) (JCGM 2012) traceability is the

property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.

Essential Climate Variables (ECVs) are relevant physical variables describing the state of the climate system, and the observation of which is feasible and cost effective, as described in Section 1.1.

In practice, satellite-based remote sensing techniques rarely measure ECVs directly, but ECVs have to be derived from a so-called Fundamental Climate Data Record (FCDR) using an often elaborate *retrieval* (GCOS 2010a). According to GCOS (2010a),

the term 'Fundamental Climate Data Record' (FCDR) is used to denote a long-term data record, generally involving a series of instruments (all platforms), with potentially changing measurement approaches, but with overlaps, calibration and quality control sufficient to allow the generation of homogeneous products providing a measure of the intended variable that is accurate and stable enough for climate monitoring.

It is recommended that providers of FCDRs document all steps taken to generate their record, including algorithms. Expected accuracy, stability, resolution and homogeneity should be documented and evaluated, access to the Fundamental Climate Data Record (FCDR), products and documentation provided, information on the scientific review process related to the FCDR construction disclosed (GCOS 2010a). ECV records retrieved from these FCDRs can serve as climate benchmark data sets.

Further, climate benchmark observations require that the uncertainty of the measurement (including corrections) has been determined, and that the entire measurement procedure and set of processing algorithms are properly documented and accessible also part of the climate monitoring principles (GCOS 2010b; GCOS 2013).

Currently no satellite-based atmospheric ECV record can serve as authoritative benchmark over months to decades (Kirchengast et al. 2017b). Besides GRUAN one other proposal to achieve the standard is the Climate Absolute Radiance and Refractivity Observatory (CLARREO) instrument. CLARREO would have the on-orbit ability to trace measurements to SI units and to detect instrument changes for the life of the mission, but a launch is not expected before 2023, due to National Aeronautics and Space Administration (NASA) budget cuts.

Global Navigation Satellite System (GNSS) Radio Occultation (RO)'s potential to deliver climate benchmark data is subject of the following two sections.

1.3 The Global Navigation Satellite System (GNSS) Radio Occultation Technique

1.3.1 Historic Development of Radio Occultation (RO)

Occultations have been used for the remote sensing of planetary atmospheres for decades (Baum and Code 1953). During a stellar occultation, rays transmitted by a star pass through the atmosphere of e.g. another planet and are then received on Earth. As the rays probe deeper into the planet's atmosphere, due to the increase of density, the rays are refracted and the intensity of the detected rays drops characteristically, allowing the derivation of e.g. the scale height of the sensed atmosphere (Elliot 1996).

When interplanetary probes, such as the Mariner, Pioneer and Voyager missions were launched in the 1960s and 1970s, refraction of radio-signals rather than star light was used to determine the properties of the planets under investigation (Fjeldbo and Eshleman 1968; Fjeldbo et al. 1971).

Remote sensing the Earth's atmosphere through the occultation of Global Positioning System (GPS) signals for was first proposed for the NASA Earth Observing System (Yunck et al. 1988), and demonstrated successfully in 1995, when measurements by a GPS receiver onboard GPS/MET confirmed the potential of the technique (Kursinski et al. 1997; Ware et al. 1996).

GNSS RO has since then emerged as remote sensing technique for the Earth's atmosphere that is relatively inexpensive, has high precision, accuracy, long-term stability, vertical resolution and all-weather-capability. These properties also indicate the general suitability of RO for climate applications, which will be in the focus in this section (Anthes et al. 2008; Steiner et al. 2011).

1.3.2 Deriving ECVs through GNSS RO

GNSS RO is a limb-sounding remote sensing technique, using GNSS signals with wavelengths of multiple centimeters. As of yet, only GPS signals (with carrier frequencies of $f_{T1} = 1.57542$ GHz and $f_{T2} = 1.22760$ GHz) have been employed for RO. In

the event of an occultation, the GNSS signals pass through the Earth’s atmosphere and are received by a Low Earth Orbit (LEO) satellite. Due to the atmospheric refractivity gradients—a function of the thermodynamic state of the atmosphere—the signal is bend. Depending on the motion of GNSS and LEO satellite relative to the Earth, during one occultation the atmosphere is scanned from top downwards (setting event) or from bottom upwards (rising event). The geometry of an occultation event is schematically shown by Figure 1.3.

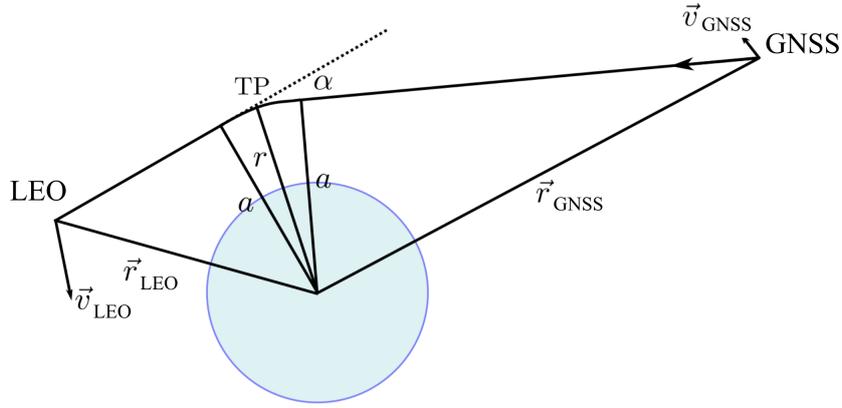


Figure 1.3: Scheme of the occultation geometry. Including the tangent point (TP), radius from the center of refraction to the TP (r), impact parameter (a), bending angle (α), orbit position (\vec{r}_{LEO}) and velocity (\vec{v}_{LEO}) vectors of the receiver, and of the transmitter satellite (\vec{r}_{GNSS} and \vec{v}_{GNSS}) (source: Pirscher 2010).

At the minimum between residual neutral gas and residual ionospheric refractive effects at about 80 km (‘top of atmosphere’), the signal passes almost unattenuated and on a straight line from the transmitter to the receiver. As the GNSS and LEO satellites move downward (upward), relative to the Earth, the signals pass through thicker (thinner) layers of the atmosphere and changes in optical path length are measured on a dense time grid of typically 50 Hz sampling rate, yielding near vertical profiles of changes in optical path relative to the top of the atmosphere (so-called *excess phase* profiles) (Hajj et al. 2002; Kursinski et al. 1997; Schreiner et al. 2010).

The refraction of the GNSS transmitted radio-waves depends on the thermodynamic state of the Earth’s atmosphere along their path. This dependence can be formalized by the Smith-Weintraub equation (Smith and Weintraub 1953), linking refractivity $N(z)$ and the thermodynamic state, expressed by the density and water vapor content of the neutral gas atmosphere, the concentration of free electrons in the ionosphere, and the

concentration of particulates,

$$N(z) = k_1 \frac{p(z)}{T(z)} + k_2 \frac{e(z)}{T(z)^2} + k_3 \frac{n_e(z)}{f_i^2} + k_4 W(z). \quad (1.1)$$

where $p(z)$ is the atmospheric pressure in hPa, $T(z)$ the atmospheric temperature in K, $e(z)$ the partial pressure of water vapor in hPa, $n_e(z)$ the electron density in m^{-3} , f_i the transmitter frequencies in Hz, and $W(z)$ the mass of condensed water in the atmosphere in g m^{-3} . The constants are $k_1 = 77.60 \text{ K hPa}^{-1}$, $k_2 = 3.73 \times 10^5 \text{ K}^2 \text{ hPa}^{-1}$, $k_3 = 4.03 \times 10^7 \text{ m}^3 \text{ s}^{-2}$, and $k_4 = 1.4 \text{ m}^3 \text{ g}^{-1}$. More advanced refractivity equations, as introduced by Thayer (1974), or Aparicio and Laroche (2011) consider additional variables, as e.g. compressibility of air, and are thus more accurate than Equation 1.1, but also the Smith Weintraub equation is still widely in use.

When focussing on the neutral gas atmosphere, the dispersive effect of the ionosphere onto the two GNSS channels' carrier frequencies can be exploited, and ionospheric effects largely corrected. Thus the third term of Equation 1.1 becomes negligible (Sokolovskiy et al. 2009; Vorob'ev and Krasil'nikova 1994). The impact of particulates (mostly liquid water) on the radio waves, is expressed in the fourth term and is also very small compared to the other terms for realistic suspension of water or ice (Kursinski et al. 1997).

The effect of water vapour (the second term) can be disregarded in regions above $\sim 8 \text{ km}$ at the pole of the winterhemisphere and above $\sim 14 \text{ km}$ in the tropics (Scherllin-Pirscher et al. 2011b).

Hence, with the help of the hydrostatic equation and the equation of state, Equation 1.1 can then be solved for $T(z)$ and $p(z)$.

In the troposphere, where the contributions of $e(z)$ can be significant, background information is needed to constrain both temperature and humidity (Engeln et al. 2005; Li et al. 2018; Palmer et al. 2000, e.g.). [TODO: best moist-air quotes]

According to Fermat's principle, electromagnetic rays, and thus also the GNSS transmitter signals, minimize the length of the optical path depending on the refractive index n of the medium through which they propagate, and thus (Syndergaard 1999),

$$L = \int_{\text{GNSS}}^{\text{LEO}} n(f_T, s) ds = \min \quad (1.2)$$

where L is the excess phase (essentially the change in length of the optical path between the transmitter and receiver satellites in meter). It is equal to the refractive index, $n(f_T, s)$, of the Earth's atmosphere along the ray path s , integrated along the ray-path. $n(f_T, s)$ is a function of the transmitter frequency f_T (in Hertz) and the path s (in meter), and is assumed to be spherically symmetric along the ray path. The refractive index n is related to refractivity N through

$$n(z) = 1 + 10^{-6} \cdot N(z). \quad (1.3)$$

Equations 1.1 to 1.3 show how the thermodynamic state of the atmosphere affects the refractivity, and how refractivity in turn affects the length of the paths of the GNSS rays through the atmosphere. In climate applications of RO, this relation is inverted, i.e. the thermodynamic state of the neutral gas atmosphere is calculated based on measurements of the length of the optical path relative to the top of the atmosphere.

The integral in Equation 1.2 can not be inverted directly, because $n(f_T, s)$ is not sufficiently constrained. In practice, a bending angle profile (see Figure 1.3) relative to the straight-line connection between GNSS and LEO satellite is calculated from the excess phase profile and the orbit position and velocity of the two satellites under the assumption that the refractivity is spherically symmetric in the area around the signal's path through the atmosphere. Then this bending angle profile can be inverted, using an Abel integral transform to get refractive index $n(z)$ (Kursinski et al. 1997).

To visualize the geometric relation of bending angle α , orbit position \vec{r}_{GNSS} and velocity \vec{v}_{GNSS} of the transmitter, and orbit position \vec{r}_{LEO} and velocity \vec{v}_{LEO} of the receiver, Figure 1.3 shows these quantities for one instant during the occultation event. Profiles of each of these quantities at the 50 Hz sampling rate describe the course of the occultation over time.

A detailed step-by-step description of an RO inversion algorithm, or *retrieval*, as implemented in the Reference Occultation Processing System (rOPS), from excess phase profile to physical pressure, temperature and humidity profiles, is provided in the Appendices A to C.

1.4 GNSS RO for Climate Benchmark Observations

Climate benchmark observation must be *global, highly accurate, long-term stable* climate records. These records also need to be tied to irrefutable standards (*traceability*), have sufficient *resolution* and measure significant indicators of climate change, such as the *ECVs*. Uncertainties should be provided, and special documentation and transparency requirements, and the GCOS climate monitoring principles apply.

Due to their unique properties, RO data records have the potential to serve as climate benchmark observations (Kirchengast et al. 2016a):

1.4.1 Properties of Benchmark Observations from RO

Global coverage: The dual-frequency GNSS signals in the microwave range are unaffected by clouds and aerosols, and ensure all-weather capability. Due to the abundance of signal sources, GNSS RO observations are globally available (Leroy et al. 2006). With the GPS constellation of at least 24 satellites, over 20 BeiDou Navigation Satellite System (BDS) satellites, and potentially Galileo satellites as available transmitters, and several presently working GPS RO missions and potential future missions, up to 20000 occultation events per day can be expected to be processed

by 2020. Between 2006 and 2016, roughly 2000 occultations per day, mostly from the mission Challenging Mini-Satellite Payload (CHAMP) (Wickert et al. 2006), the six-satellite Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) (Anthes et al. 2008), the twin satellite Gravity Recovery and Climate Experiment (GRACE) mission (Beyerle et al. 2005), and the Meteorological Operational (MetOp) (Bonnedal et al. 2010) have been processed at the Wegener Center for Climate and Global Change (WEGC) data center. A combination of polar orbiting and low-inclination receiving LEO satellites ensures coverage of all latitudes.

Accuracy: The GNSS receiver on the LEO satellite replicates the code of the GNSS signals and by comparing the receiver-generated phase to the received GNSS phase, the receiver tracks the shift in phase (in meter) from one measurement point to the next by means of counting the difference in cycles (full cycles and cycle fraction) multiplied by the wavelength. At the beginning (end) of each setting (rising) event, at the *top of the atmosphere* the signal passes almost unrefracted through the atmosphere and the instrument is thus essentially naturally re-calibrated for each individual RO event. The accuracy of the RO phase shift measurements and the Precise Orbit Determination (POD) of the LEO depends on the accuracy of the clock of the GNSS satellite, commonly highly stable to fractions of a nanosecond. Each GNSS satellite carries a precise atomic oscillator producing the reference frequency by stimulated radiation of rubidium or caesium. From this reference frequency the carrier frequency is derived. Potential clock errors of the LEO satellites can be removed by differencing methods, but generally also LEO satellite clocks are ultra-stable and do not require differencing (Dow et al. 2009; Griffiths and Ray 2009; Schreiner et al. 2010). Due to these properties, the RO excess phase measurement (the FCDR) is effectively tied to the accuracy of the SI second (Leroy et al. 2006).

Long-term stability: The highly accurate GNSS and LEO clocks ensure measurement stability during the occultation event, but also between different occultation events. Given the same atmospheric and ionospheric structure, due to the self-calibrating property, different models of receivers at different times and places would measure the the same phase shift, as long as the accuracy of transmitter and receiver clocks does not deteriorate (Steiner et al. 2011). Consequently RO FCDR are long-term stable, and reproducibility and homogeneity are assured. Thus an FCDR can be combined with others from different RO missions without inter-calibration and overlap (Foelsche et al. 2009; Hajj et al. 2004).

Resolution: Due to the limb-sounding geometric setup, the horizontal resolution of RO FCDRs is relatively low (~ 300 km), compared to nadir-looking microwave sounding techniques or radiosondes, but particularly for climatologies with spatially averaged

data, the resolution is sufficient (Foelsche et al. 2008). The vertical resolution on the other hand is high and ranges from near 200 m in the lower troposphere, where it is diffraction- and horizontal variation-limited (Gorbunov et al. 2004) to ~ 500 m in the lower stratosphere, to a half-Fresnel-scale resolution of near 800 m in the upper stratosphere (Kursinski et al. 1997).

Traceability The RO excess phase measurements and orbit information are based on the timing of the GNSS clocks and traceability of the FC DR to the international time standard is thus possible (Leroy et al. 2006), but so far not ensured for RO based ECVs, because the retrievals need to fulfill particular requirements (as discussed below and in Section 1.5).

1.4.2 ECVs from Fundamental Climate Data Records

From the RO FC DR various atmospheric ECVs can be retrieved, particularly pressure, temperature in the (dry) Upper Troposphere–Lower Stratosphere (UTLS) region and using background information also pressure, temperature, and specific humidity in the (lower) troposphere. RO has the potential to reach the *breakthrough* criteria of some of the GCOS observation requirements (GCOS 2010a) for upper-air temperature and specific humidity. These are e.g., 2.5 km vertical resolution in the stratosphere, with an accuracy of smaller 1 K error for temperature and 5% for specific humidity in the UTLS. The horizontal resolution of the temperature observations pass the *minimum* criteria (Kursinski et al. 1997; Rocken et al. 1997; Steiner et al. 1999).

RO observations also obey the GCOS climate monitoring principles (GCOS 2010a). To give a few examples, due to the self-calibrating nature of RO observations, Principle 15 (on-board calibration) is automatically fulfilled, and Principles 2, 12 and 13 (overlap between missions) are obsolete (although generally fulfilled). Principle 7 (global coverage) and Principle 11 (constant sampling) depend on the orbit inclinations of the RO missions, but have been reached by the realized missions.

RO observations have the potential to deliver climate benchmark data, because they fulfill requirements like global availability, accuracy, long-term stability, high resolution, and provision of ECVs, but the establishment of traceability of the ECV records accuracy to the Systeme International (SI) second has not yet been realized. Traceability would require that the ECVs can be related to the SI reference through a documented unbroken chain of calibrations, along the retrieval chain (JCGM 2012).

Thus the accuracy of the FC DR must be sustained also for the quantities retrieved from the FC DR and traceability be ensured by propagating the remaining (small) uncertainties through the retrieval (Kirchengast et al. 2015).

1.5 The Reference Occultation Processing System (rOPS) - Ensuring Traceability

1.5.1 Motivation for the Development of the rOPS

Existing retrievals at the main RO processing centers already reach high accuracy for atmospheric profiles of bending angle, refractivity, pressure, geopotential height, and temperature (< 1 K), and structural uncertainties between RO records from different processing centers are low (Ho et al. 2012). However, despite substantial advances in the processing of RO observations since the proof-of-concept RO mission Global Positioning System/Meteorology (GPS/MET) in the mid-1990ties, no rigorous trace from fundamental time (including uncertainties of side information) to the ECVs exists so far.

The new Reference Occultation Processing System (rOPS) at the Wegener Center for Climate and Global Change (WEGC) (Kirchengast et al. 2015), developed at the WEGC at the University of Graz, Austria, aims to realize the benchmark quality potential of RO by establishing (1) an even more accurate retrieval with residual relative numerical uncertainties below 10^{-4} , and (2) integrating a complete uncertainty propagation chain from the fundamental-time uncertainty and relevant side influences from RO raw tracking data and high-accuracy GNSS orbit data through four processors to atmospheric variables such that remaining uncertainties can be traced from the ECVs back to the FCDR and the raw data, tied to the SI standard.

1.5.2 Development History of the rOPS

Building on experience gained with operating the previous occultation processing system at the WEGC, the End-to-End Generic Occultation Performance Simulation and Processing System (EGOPS) Version 5.6 (Fritzer et al. 2009), the development of the new rOPS began in 2011 with the basic L1b, L2a and L2b components, funded by the OPSGRAS project of the European Space Agency (ESA)).

Most of the remaining rOPS components were completed over the last years, and received support from the projects OPSCLIMPROP, OPSCLIMTRACE, OPSCLIMVALUE from the Austrian Research Promotion Agency (Österreichische Forschungsförderungsgesellschaft) (FFG)-Aeronautics and Space Agency (ALR) and from the MMValRO-E project from ESA/ESRIN.

The development has received support from experts at various institutions, as e.g., from EUMETSAT in Darmstadt, ROM-SAF in Copenhagen-Reading, the Institute of Atmospheric Physics (IAP) in Moscow, the Astronomical Institute at the University of Bern (AIUB) in Bern, the DLR in Oberpfaffenhofen, from UCAR in Boulder, from JPL in Pasadena, from IGG/CAS in Wuhan, and from the Royal Melbourne Institute of Technology (RMIT) in Melbourne.

Based on the new rOPS, the WEGC is scheduled to serve as global distribution center for RO data in benchmark quality from multiple future satellite missions such as the Constellation Observing System for Meteorology, Ionosphere, and Climate-2 (COSMIC-2) and the Feng Yun series-3 satellite C (FY-3C)/GNSS radio-occultation sounder (GNOS), but also past and current RO missions, like Global Navigation Satellite Systems Receiver for Atmospheric Sounding (GRAS) on MetOp and the FORMOSAT-3/COSMIC (F3C) mission. The aim is to broadly provide a new reference standard for thermodynamic ECVs for atmosphere and climate research and applications (Kirchengast et al. 2015).

A concise overview of the rOPS goal, processing tracks, and the processor structure, is provided in Table 1.1.

Table 1.1: Overall objective and processing tracks and processors (source: Kirchengast et al. 2016a).

Goal:
Provide benchmark-quality reference RO data for calibration/validation and for climate monitoring, research, and services, complementary to NRT.
Three processing tracks:
<ul style="list-style-type: none">• Fast-track reference (FTR) data: daily on follow-on day of observations• Postprocess-track reference (PTR) data: daily within one month latency• Re-processing reference (RPR) data: occasionally (kicked or manually), as highest-fidelity climate records over entire multi-satellite periods
Four processors:
<ul style="list-style-type: none">• Level 1a processor: Raw ($L0_{pf}$) data to excess phase level data (L1a)• Level 1b processor: Excess phase/SNR to atmospheric bending angle (L1b) ($t \rightarrow a$ space; observation only, no background info)• Level 2a processor: Bending angle to refractivity and dry-air variables (L2a) ($a \rightarrow z$ space; background info from high-altitude initialization only)• Level 2b processor: Dry-air variables to moist-air variables (L2b) (tropospheric moist-air retrieval in z space: background info on T, q)

1.5.3 Basic Structure of the rOPS

Figures 1.4 to 1.7 show the workflow of the rOPS retrieval chain through its four main processing steps. The processors include a first step to excess phase profiles (L1a, Figure 1.4), a second to atmospheric bending angle profiles (L1b, Figure 1.5), a third to refractivity and dry-air profiles (L2a Figure 1.6), and a fourth to final thermodynamic profiles (ECVs) (L2b, Figure 1.7).

The rOPS retrieval chain starts with the POD (Item L1a (1) in Figure 1.4), provision

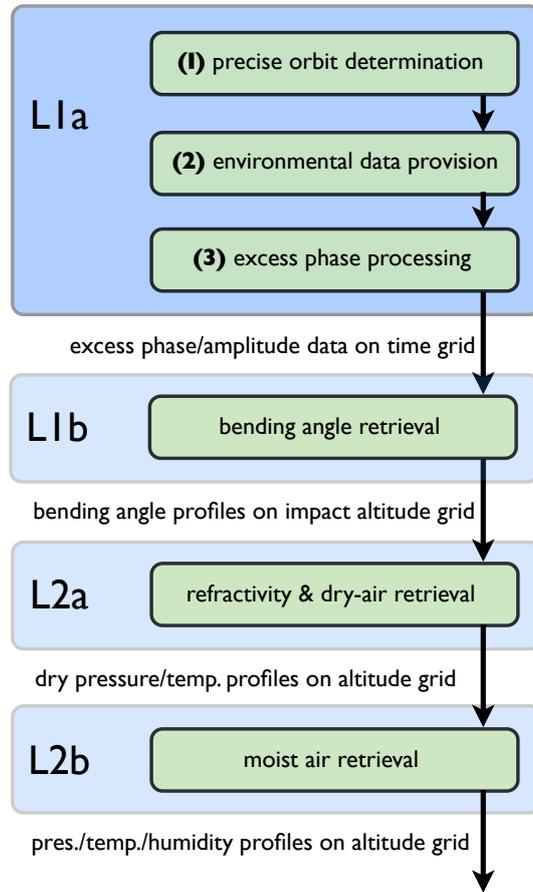


Figure 1.4: Schematic view of the main processors of the retrieval chain in the rOPS (L1a highlighted, L1b, L2a, L2b) and the main operators of the L1a processor, the precise orbit determination (1), environmental data provision (2) and the excess phase processing (3).

of environmental data (mostly on magneto- and ionosphere) (Item L1a (2) in Figure 1.4), and excess phase processing (Item L1a (3) in Figure 1.4) where raw orbit and occultation data are taken as input to retrieve the excess phase profiles on a 50 Hz time grid $L_r(t)$.

In the bending angle retrieval, the excess phase profiles from both GNSS transmitter channels are processed to retrieve separate geometric-optics (GO) and wave-optics (WO) bending angle profiles which are then merged (Item L1b (2) in Figure 1.5). Then the two GNSS channels' bending angle profiles are used to correct for ionospheric effects (Item L1b (3) in Figure 1.5), and the neutral gas atmospheric bending angle profile as a function of impact altitude is derived $\alpha_r(z_a)$. A detailed step-by-step description

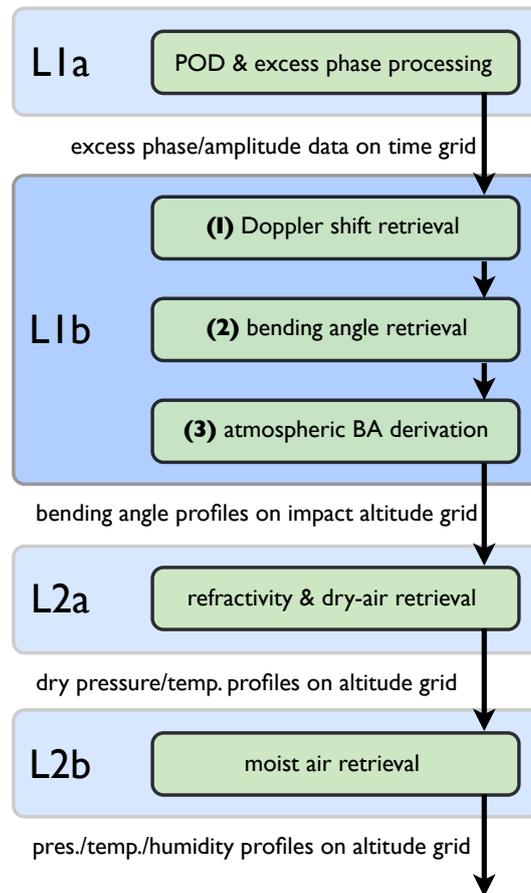


Figure 1.5: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b highlighted, L2a, L2b) and the main operators of the L1b processor, the Doppler shift retrieval (1), the bending angle retrieval (2), and the atmospheric bending angle derivation (3).

of the GO retrieval of the rOPS L1b processor, including the uncertainty propagation algorithm, from excess phase profile to atmospheric bending angle profiles is provided in the Appendix A.

In the L2a processor, the atmospheric bending angle is filtered to a target resolution and interpolated to a sparser grid, before being initializing at high altitudes using background information (Item L2a (1) in Figure 1.6). The initialized bending angle profile is then integrated using the Abel transform, which yields refractivity profiles (Item L2a (2) in Figure 1.6). Finally, from the refractivity profile, the dry-air profiles dry pressure $p_d(z)$

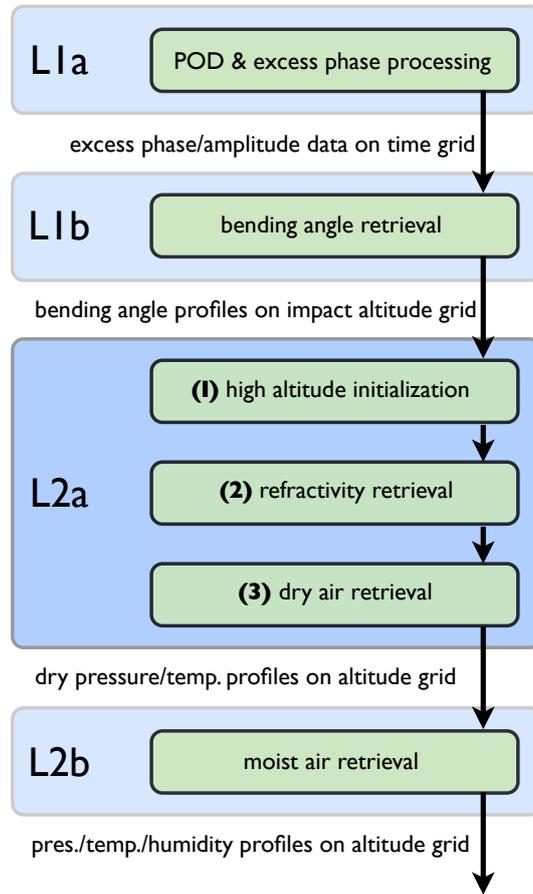


Figure 1.6: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b, L2a highlighted, L2b) and the main operators of the L2a processor, the high altitude initialization (1), the refractivity retrieval (2), and the dry-air retrieval (3).

and dry temperature $T_d(z)$, both relative to a Mean Sea Level (MSL) altitude grid z are derived with the help of the Smith Weintraub equation (Item L2a (3) in Figure 1.6).

The final step is the rOPS moist-air retrieval (Li et al. 2018) to derive ECVs with respect to MSL altitude grid z from dry air variables and background information. First, direct retrievals derive temperature (Item (1a) in Figure 1.7) and specific humidity (Item (1b) in Figure 1.7) profiles from the retrieved dry parameters and background information on temperature and humidity. Secondly these retrieved temperature and specific humidity profiles are combined with background information in a covariance weighted optimal estimation to get the ECVs temperature $T_e(z)$, and specific humidity $q_e(z)$ (Item (2)

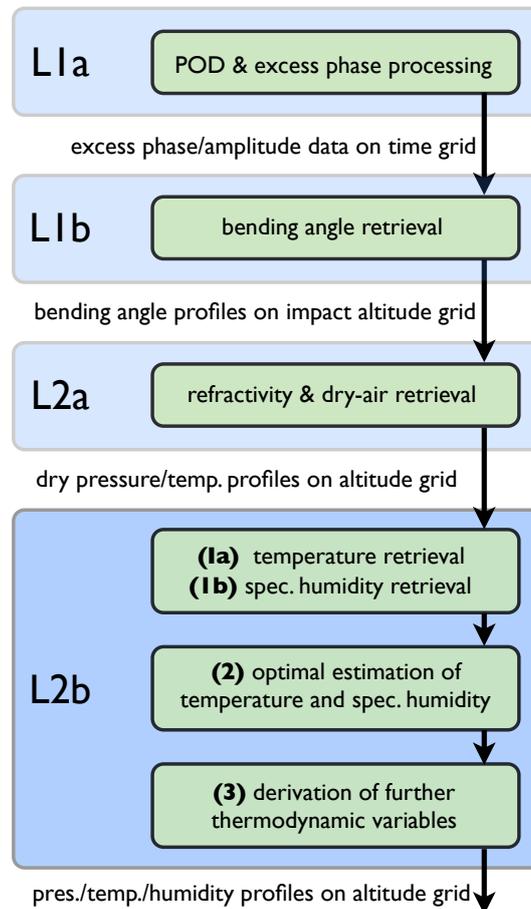


Figure 1.7: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b, L2a, L2b highlighted) and the main operators of the L2b processor (1, 2, 3), which are in the focus of this study.

in Figure 1.7). In the last step of the rOPS moist air retrieval, the optimal estimation temperature and specific humidity profiles are used to calculate further thermodynamic variables, like e.g., pressure $p_e(z)$ and density $\rho_e(z)$ (Item (3) in Figure 1.7).

1.5.4 rOPS Innovations

An overview of the rOPS system layers, design principles and key innovations, is provided in Tables 1.2 to 1.4. The two major system layers, background daily system modeling and occultation event system modeling, operate once per day, and once per event, respectively (Table 1.2). They run across the different processors and contain several subsystems.

Table 1.2: The two main rOPS system layers and their subsystems (updated from: Kirchengast et al. 2016a).

Background and observation system modeling (daily):

- Atmosphere system (ECMWF An, Fc; rOPS Ret-L1b, Ret-L2a)
- Atm.uncertainties system ($F_{\varphi,\lambda,\tau}(\mathbf{x}_a, \mathbf{x}_f; \mathbf{x}_0)$ space-time filters: $\mathbf{u}_b, \mathbf{u}_r, \mathbf{R}_b, \mathbf{R}_r$)
- GNSS Tx system (CODE & IGS $\mathbf{x}_G, \mathbf{v}_G, \tau_G; \mathbf{U}_{s,XG}, \mathbf{U}_{c,XG}$)
- LEO Rx system (\mathbf{X}_G & $\mathbf{X}_{L0nav,L}$ /Bernese & Napeos $\mathbf{x}_L, \mathbf{v}_L, \tau_L; \mathbf{U}_{s,XL}, \mathbf{U}_{c,XL}$)

Occultation event system modeling (per RO event):

- Event geometry modeling (WGS84/EGM2008/ECEF $\mathbf{X}_G, \mathbf{U}_{XG}; \mathbf{X}_L, \mathbf{U}_{XL}; t^{\text{mtp}}, \mathbf{x}^{\text{mtp}}, \mathbf{s}^{\text{mtp}}, R_C^{\text{mtp}}, \delta \mathbf{r}_C^{\text{mtp}}; h_G^{\text{mtp}}, R_E^{\text{mtp}}$ (also for sltp80)
- Event environment modeling (CODE & IGS GNSS vTEC $\mathbf{x}_{ion}, \mathbf{u}_{ion};$ Sfc-Atm-SpW flags and indices)
- L2 data modeling (ECMWF Fc-L2b, Fc-L2a; for ref.An-L2b,-L2a)
- L1 data modeling ($F_k(N \rightarrow \alpha \rightarrow D \rightarrow L)$ operators: Fc-L1b, Fc-L1a)
- Vertical grid modeling ($z \rightarrow z_p, z_a, a_z \leftrightarrow a_t, z_{a,t}, z_t, t_{s,t} \leftarrow t; \mathbf{X}^{\text{tp}}$)
- Processing operators (\mathbf{A}_k models for $L \leftrightarrow D \leftrightarrow \alpha \leftrightarrow N \leftrightarrow p_d, T_d \leftrightarrow \rho, p, T, q$)
- Residual bias modeling (e.g., for spher.symmetry, hydrostaticity, clear air)

The four guiding principles for the data processing design, i.e., the baseband approach, the forward-inverse consistency and reversibility, the integrated uncertainty estimation, and the increased processing speed, are outlined in Table 1.3. The baseband approach helps to decrease residual numerical errors by reducing the dynamic range of the processed profiles, and thus helps to ensure the forward-inverse consistency. Both help to simplify uncertainty propagation, because they reduce the magnitude of algorithmic uncertainty, see Section 2.3.1 (Kirchengast et al. 2016a).

The forward-inverse consistency is also improved by the vertical grid model (Item 2 in Table 1.3), as e.g. the impact altitude grid z_a is constructed from the equidistant output grid z and the background refractivity N_b such that the retrieved altitude grid z_r (calculated from z_a and the retrieved refractivity N_r) matches z closely (*grid matching*) and only a minor interpolation from z_r to z (a *grid adjustment*) is necessary, keeping algorithmic uncertainties from interpolation (see Section 2.3.1 for a categorization of uncertainties) low (Kirchengast et al. 2016a).

To test and demonstrate the operator reversibility of the rOPS, Basic-Update-Control (BUC) tests were run on all operators. In a BUC test, the forward and the retrieval operator are each called two times in a loop, to test the numerical stability of the operator and the forward-inverse consistency. For the results in Figure 1.8. the forward Abel integral was called with a basic refractivity profile as input. The output, i.e., the bending angle profile, was then used as input for the refractivity retrieval. The resulting 'update'

Table 1.3: The four guiding principles of the rOPS processing design (updated from: Kirchengast et al. 2016a).

Four guiding principles for the data processing design:

- Baseband approach: L1/L2 data processing (state retrieval, associated uncertainty estimation, and QC) mainly done on delta-signals obtained from "down-conversion" with L1, L2 data models ($\delta \mathbf{x}_{\mathbf{r}\mathbf{m}} = \mathbf{x}_{\mathbf{r}} - \mathbf{x}_{\mathbf{m}}$); pros: low dynamic range, minimized biases, simplified operators, etc.
 - Forward-inverse consistency and reversibility: vertical grid model ($z \rightarrow z_p, z_a, a_z \leftrightarrow a_t, z_{a,t}, z_t, t_{s,t} \leftarrow t$) and forward-inverse process operators ($L(t) \leftrightarrow D(t) \leftrightarrow \alpha(a) \leftrightarrow N(z) \leftrightarrow p_d, T_d \leftrightarrow \rho, p, T, q$) consistent and reversible within tight constraints ($\Delta x/x < 10^{-6}$ to 10^{-4}); pros: negligible numerical residual errors, etc.
 - Integrated uncertainty estimation: given the base band approach and simple linear (matrix-algebraic) process operators, forward and inverse uncertainty propagation ($\mathbf{u}_x^b, \mathbf{u}_x^s, \mathbf{u}_x^c; \mathbf{R}_x; \mathbf{w}_x$) comes in seamless and is fast; pros: rigorous QC'ed uncertainty trace, explicit vertical resolution, etc.
 - Increased processing speed: given the smart(er) rOPS design and "optimally coded" core modules, aim at increased L1b/L2a/L2b processing speed (vs OPSv5.6), despite much higher rOPS information content.
-

refractivity profile is then again used as input for the forward operator, and the loop repeated (to get to the 'control' profile). As can be seen in Figure 1.8 the relative error stays below 0.05 % between 2 km and 60 km, and further improvements are on-going to increase the relative accuracy to below 10^{-4} (Kirchengast et al. 2015).

Figure 1.9 compares results for numerical errors estimated using one BUC loop from different versions of the refractivity retrieval. It can be seen that the residual numerical errors of the new baseband piecewise-analytical solution to the Abel integral (Syndergaard and Kirchengast 2016), as implemented in the rOPS, are about a magnitude smaller than those of a refractivity retrieval using standard trapezoidal rule discretization of the integral (as e.g., in EGOPsv5.6). The BUC tests of other operators (filtered excess phase L_F to Doppler shift profile D_r , D_r to GO bending angle profile α_G , also refractivity N_r to dry-air variables etc.) showed high forward-inverse consistency, in line with the accuracy target ranges.

Further examples of innovations and improvements in different components of the rOPS retrieval chain are listed in Table 1.4.

1.5 The Reference Occultation Processing System (rOPS) - Ensuring Traceability

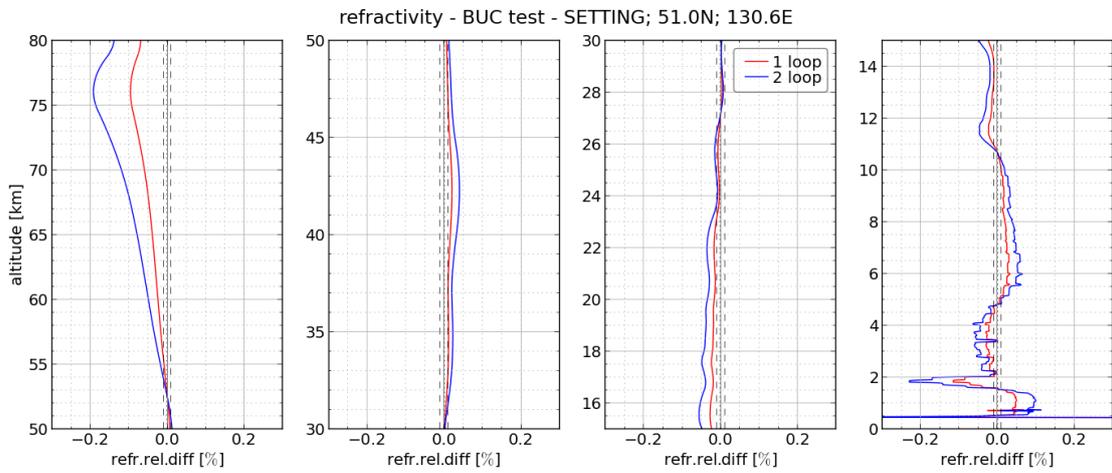


Figure 1.8: Results from the Basic-Update-Control (BUC) test of the refractivity retrieval. Errors after the 1st loop (red) and after the 2nd (blue), relative to refractivity and with respect to altitude. The gray dashed line indicates the target accuracy of 10^{-4} (source: Kirchengast et al. 2015).

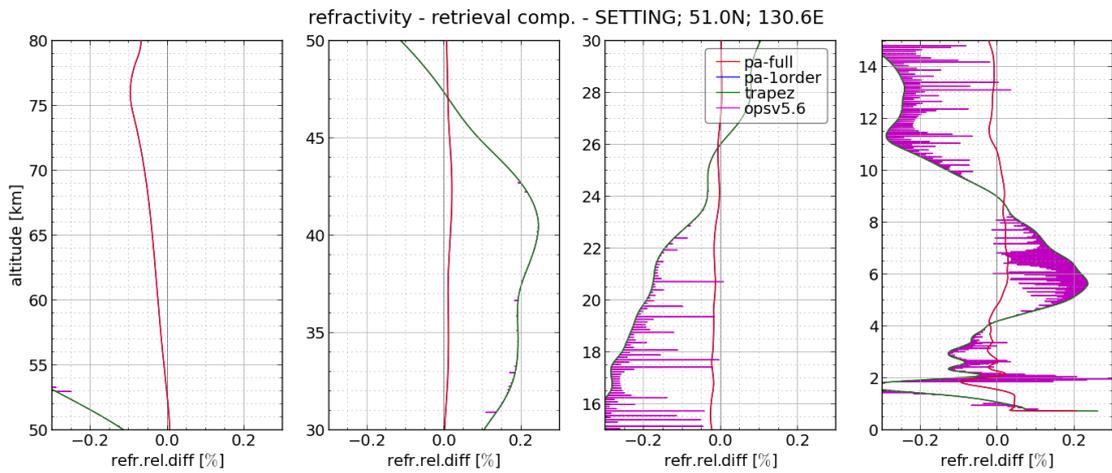


Figure 1.9: Results for relative numerical errors after one forward-inverse loop for 4 different methods for solving the Abel integral in the refractivity retrieval: the rOPS baseband piecewise-analytical solution (red), the baseband piecewise-analytical solution to 1st order (blue), the trapezoidal rule applied in the rOPS (green), and the trapezoidal rule as applied in the EGOPsv5.6 refractivity retrieval (purple)(source: Kirchengast et al. 2015).

Table 1.4: Examples for innovations in rOPS component algorithms (updated from: Kirchengast et al. 2016a).

Innovative and generalized component algorithms—examples:

- do essentially all main QC at excess phase delta-signal level
 - lowpass filters (BWS) with quantified-BW/cutoff-frequency throughout, for vertical resolution trace
 - generalized GO and WO bending angle retrieval/forward modeling, for optimal accuracy and incl. boundary layer bias (BIB) correction
 - dynamic GO-WO bend.ang. merging in upper troposphere, for optimal transition
 - generalized (Sokolovskiy et al. 2009) ionospheric correction, and residual ionospheric bias (RIB) correction
 - dynamic statistical optimization (Li et al. 2013; Li et al. 2015), for optimal subsequent retrieval
 - semi-analytic solutions of Abel and hydrostatic integrals, for optimal accuracy
 - generalized moist-air retrieval for essentially arbitrary (e.g., Thayer-type) refractivity equations, linearized around dry-air state, with dynamic background information
-

1.5.5 Ensuring Traceability

To realize the benchmark quality potential of RO, the rOPS needs to establishing a highly accurate retrieval and integrate uncertainty propagation throughout the retrieval chain. Innovations like the baseband approach are aimed at considerably increasing (state) retrieval accuracy by eliminating residual algorithmic uncertainties from numerical errors, but they also simplify uncertainty propagation, because consequentially algorithmic uncertainties only need to be considered and added to the propagated uncertainty in a few cases.

In order to ensure traceability, the random u_X^r and systematic u_X^s uncertainty profiles need to be traced throughout the four processors to each variable X along the retrieval chain. In addition, correlation length profiles l_X , resolution profiles w_X , and observation-to-background weighting ratio profiles r_X^{obw} are provided by the rOPS. With uncertainties quantified along the retrieval chain all the way to the ECVs, the rOPS also satisfies GCOS climate monitoring Principle 20 (identification of random errors and time-dependent biases of derived products, i.e. ECVs).

This thesis contributed to development, implementation and testing of the uncertainty propagation for the L1b processing (GO only) (Schwarz et al. 2018a), the L2a processing (Schwarz et al. 2017) and the L2b processing (Schwarz et al. 2018b). The research articles

on which this thesis is based, provide the used uncertainty propagation algorithms in their respective appendices, also added as Appendices A to B in this thesis, in line with GCOS requirements concerning the documentation and disclosure of algorithms used for the provision of FCDR and ECV data records (GCOS 2010a). Details on the used uncertainty estimation, propagation and validation methods are introduced in Chapter 2.

1.6 The rOPS Development Context

The work in the context of this thesis, and as documented in Schwarz et al. (2017), Schwarz et al. (2018a) and Schwarz et al. (2018b), were carried out in the framework of the overall development of the rOPS traceable processing, within which multiple projects are simultaneously on-going.

First, the design and implementation of the L1a processing, consisting of the POD for the receiver satellite and the excess phase processing, is on-going. Uncertainty propagation will be integrated, and the corresponding manuscripts are in preparation (Pock et al. 2017, manuscript in preparation for *Journal of Geodesy*, and Innerkofler et al. 2017, manuscript in preparation for *Atmospheric Measurement Techniques*).

Uncertainty propagation for the WO bending angle retrieval was not included in this thesis, but has been implemented and demonstrated by Gorbunov and Kirchengast (2015) for simulated events and by Gorbunov and Kirchengast (2018) for the estimation of random and systematic uncertainties for real events including boundary layer bias correction.

Further rOPS retrieval advancements in this context concern the reduction of residual higher-order ionospheric effects in atmospheric bending angle profiles (based on studies by Syndergaard (2000), Liu et al. (2015), Healy and Culverwell (2015) and Danzer et al. (2013) and Danzer et al. (2015)) and the rOPS implementation of the dynamic high altitude initialization algorithm, introduced by Li et al. (2013) and Li et al. (2015).

Finally, preparation of a new moist air retrieval algorithm (L2b) is on-going (Kirchengast et al. 2017a; Li et al. 2018)).

Estimation, Propagation and Validation of Measurement Uncertainty

Generally, a reported result of a measurement of a physical quantity should provide a quantitative indication of the quality of the result. This ensures that users can assess the reliability of the result, and compare it to other measurement results or with reference values given in a specification or standard (JCGM 2008a).

As described in Section 1.4, traceable uncertainty estimates are also a prerequisite for the provision of climate benchmark data records. For tracing uncertainty through retrieval operators to the ECVs, uncertainty sources need to be identified, the types of uncertainties classified, their magnitudes estimated. The analytical versions of the propagation models need to be found, justified and validated, the corresponding numerical algorithms developed, before implementing them into the rOPS. For all of this, the Guide to the expression of Uncertainty in Measurement (GUM) served as the main source of reference (JCGM 2008a).

2.1 The Guide to the Expression of Uncertainty in Measurement (GUM)

What is referred to as the Guide to the expression of Uncertainty in Measurement (GUM) here, is the original guide (JCGM 2008a) and its supplement 2, the "extension to any number of output quantities" (JCGM 2011), used as basis for the evaluation and propagation of uncertainties through the rOPS multivariate measurement models. Its first supplement (JCGM 2008b) was used as the basis for the Monte Carlo (MC) framework needed to validate the implemented rOPS GUM-type uncertainty propagation.

2.1.1 Motivation for the Establishment of the GUM

Evaluating and expressing uncertainty with the same standardized methods throughout the world allows that the significance of a broad spectrum of measurement results in various fields (science, engineering, commerce, industry, regulation etc.) is understood and interpreted appropriately, similar to how the widespread use of the SI has brought coherence to scientific and technological measurements (JCGM 2008a).

The GUM intends to establish these standardized rules for evaluating and expressing uncertainty in measurement and aims to be applicable to large spectrum of measurements, e.g., measurements required for maintaining quality control and quality assurance in production, or for complying with and enforcing laws and regulations, for calibrating standards and instruments, and for performing tests throughout national measurement system in order to achieve traceability to national and international physical reference standards.

To achieve this aim, the GUM intends to provide a *universal* evaluation method with respect to types of measurements and input data, and a quantity expressing uncertainty that is *internally consistent* (i.e., the way components contributing to the uncertainty are grouped and ordered must not matter) and *transferable* (i.e., the evaluated uncertainty of one measurement can be used as component of the uncertainty of another measurement in which the first is used as input) (JCGM 2008a).

The GUM provides only general rules, not technology-specific instructions and gives no instructions how the GUM-type evaluated uncertainty of a particular measurement result may be used for different purposes, e.g., for establishing tolerance limits in a manufacturing process, or decision making based on the uncertainties. For this, field separate, specific standard would need to be established, building on the GUM (JCGM 2008a).

2.1.2 Terminology for the Measurement Process

The GUM terminology used in this thesis is introduced in Appendix B and D of JCGM (2008a). The terms most frequently used in the next sections are (with minor modifications to JCGM (2008a)):

(measurable particular) Quantity: *The attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively.*

In the rOPS uncertainty propagation, e.g., the temperature of the Earth's atmosphere at a specific time, latitude, longitude and altitude.

Measurement: *The set of operations having the object of determining a value of a quantity.*

In an remote sensing techniques like RO this includes the retrieval, i.e., all steps necessary to actually determine the value of e.g., the temperature.

Measurand: : *The particular quantity which is subject to measurement.*

E.g., the temperature as physical quantity is subject to measurement during the occultation event and processing, and is thus a measurand.

(true) Value (of a measurand): *The magnitude of a measurand generally expressed as a unit of measurement multiplied by a number.*

The term *true value* is replaced by *value* in the GUM, because even the 'true' value of a measurand is always only accurate up to a point of incomplete definition (a perfect definition is impossible). In the rOPS uncertainty propagation, e.g., the value of the atmospheric temperature at the specific time, latitude, longitude and altitude, which we want to determine with the measurement, in Kelvin.

Result of a measurement: *The value attributed to a measurand, obtained by measurement.*

In general, the result of a measurement is only an approximation or estimate of the (true) value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate.

Accuracy of measurement: *The closeness of the agreement between the result of a measurement and a (true) value of the measurand.*

Error (of measurement): *The result of the measurement minus the (true) value of the measurand. In principle unknown.*

Uncertainty (of measurement): *A parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.*

E.g., for atmospheric temperature from RO measurements this is determined by propagation of all sources of uncertainties entering the observation and processing, through the processing chain.

Measurement model: Defined by JCGM (2011) as *a mathematical relation among all quantities known to be involved in a measurement.*

Expressed in a general form by $g(Y, X_1, \dots, X_k, \dots, X_K) = 0$ (or in the explicit case by $Y = f(X_1, \dots, X_k, \dots, X_K)$), where Y , the output quantity in the measurement model, is the measurand. The value of Y is to be inferred from information about input quantities $X_1, X_2, \dots, X_k, \dots, X_K$.

E.g., in the rOPS moist air retrieval the optimal estimation temperature $T_e(z_i)$, at altitude level i is calculated from a background temperature profile $T_b(z)$ and a retrieved temperature profile $T_q(z)$ through optimal estimation, i.e., the measurement model is $T_{e,i} = f(T_{q,1}, \dots, T_{q,i}, \dots, T_{q,N}, T_{b,1}, \dots, T_{b,i}, \dots, T_{b,N})$. In the rOPS and in this thesis, measurement models are referred to as *retrieval operators* or forward modeling operators, or generally *processing operators*.

Figure 2.1 illustrates the meaning of these terms. In the left column the relationship of the 'true' value of the measurand to the result of the measurement is shown. The result of the measurement is obtained from an estimate of the expected value gained through repeated observations (see Equation 2.1), and after applying a correction, the magnitude of which is determined by a Type B evaluation (see Section 2.1.3).

The method for evaluating and expressing uncertainty in measurement of the GUM provides a realistic rather than a 'safe' value of uncertainty, and is based on the idea that there is no inherent difference between an uncertainty component arising from a random effect and one arising from a correction for a systematic effect (see also Section 2.3).

The GUM takes an operational view that the *measurement result* is the value attributed to the measurand, and its evaluated *uncertainty* a measure of the dispersion of the values that could reasonably be attributed to the measurand. The Guide thus avoids the unknowable quantities "*true*" *value* and *error*, because the values of the contributions to the error are unknown and unknowable, but the uncertainties associated with the random and systematic effects that cause the error can be evaluated. However it is to be kept in mind, that even when the evaluated uncertainties are small, there is still no guarantee that the error in the measurement result is small as well, e.g., a systematic effect may have been unrecognized and not been taken into consideration.

The uncertainty of a result of a measurement is thus not an indication of the likelihood that the result of the measurement is near the value of the measurand (having a high accuracy), it is rather an estimate of the likelihood of closeness to the best value at presently available knowledge.

Figure 2.2 illustrates the differences between error and uncertainty for an unknown probability distribution sampled by a finite number of observations.

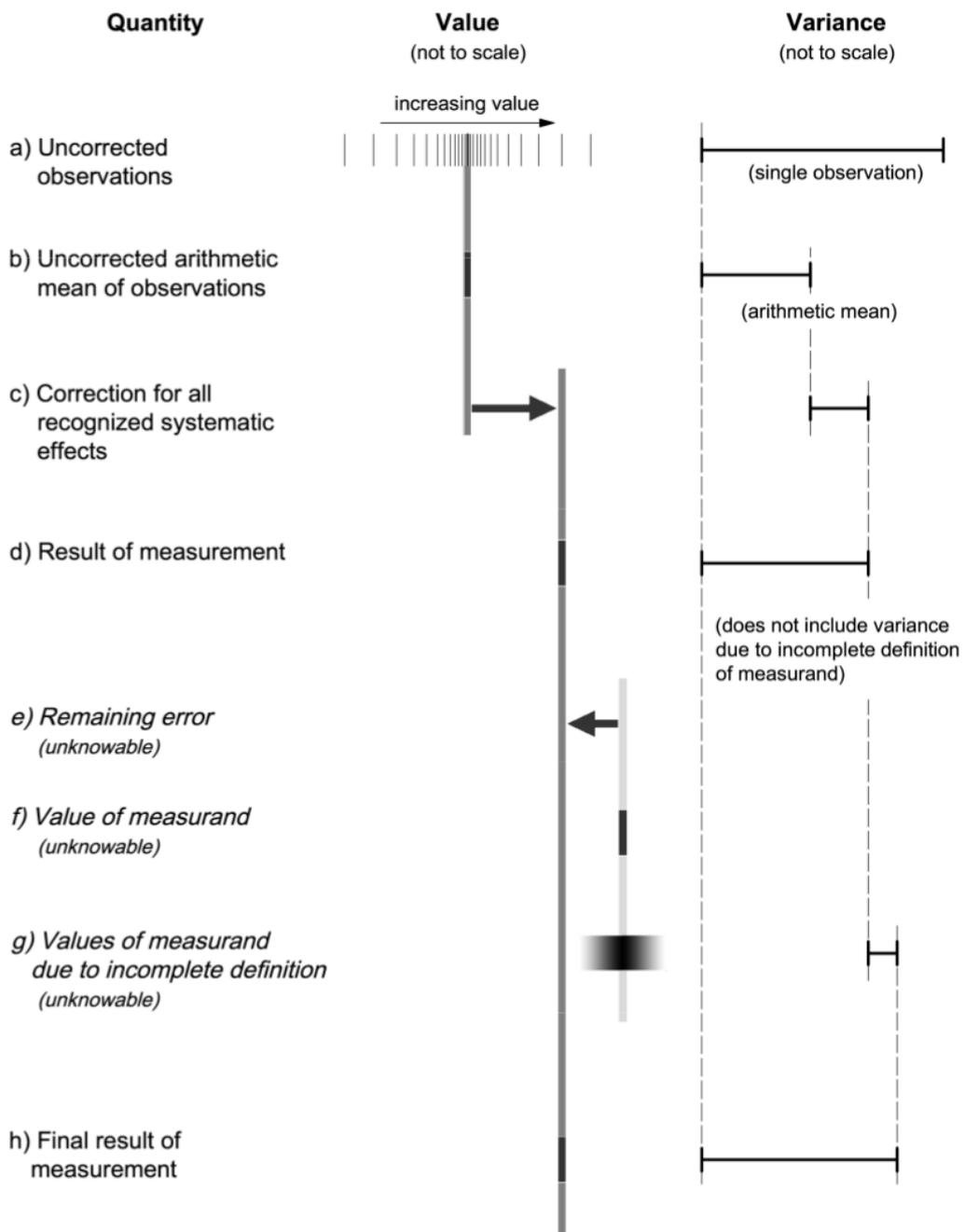


Figure 2.1: Illustration of the meaning of the terms *value of measurand* and *result of measurement, uncertainty* (source: JCGM 2008a).

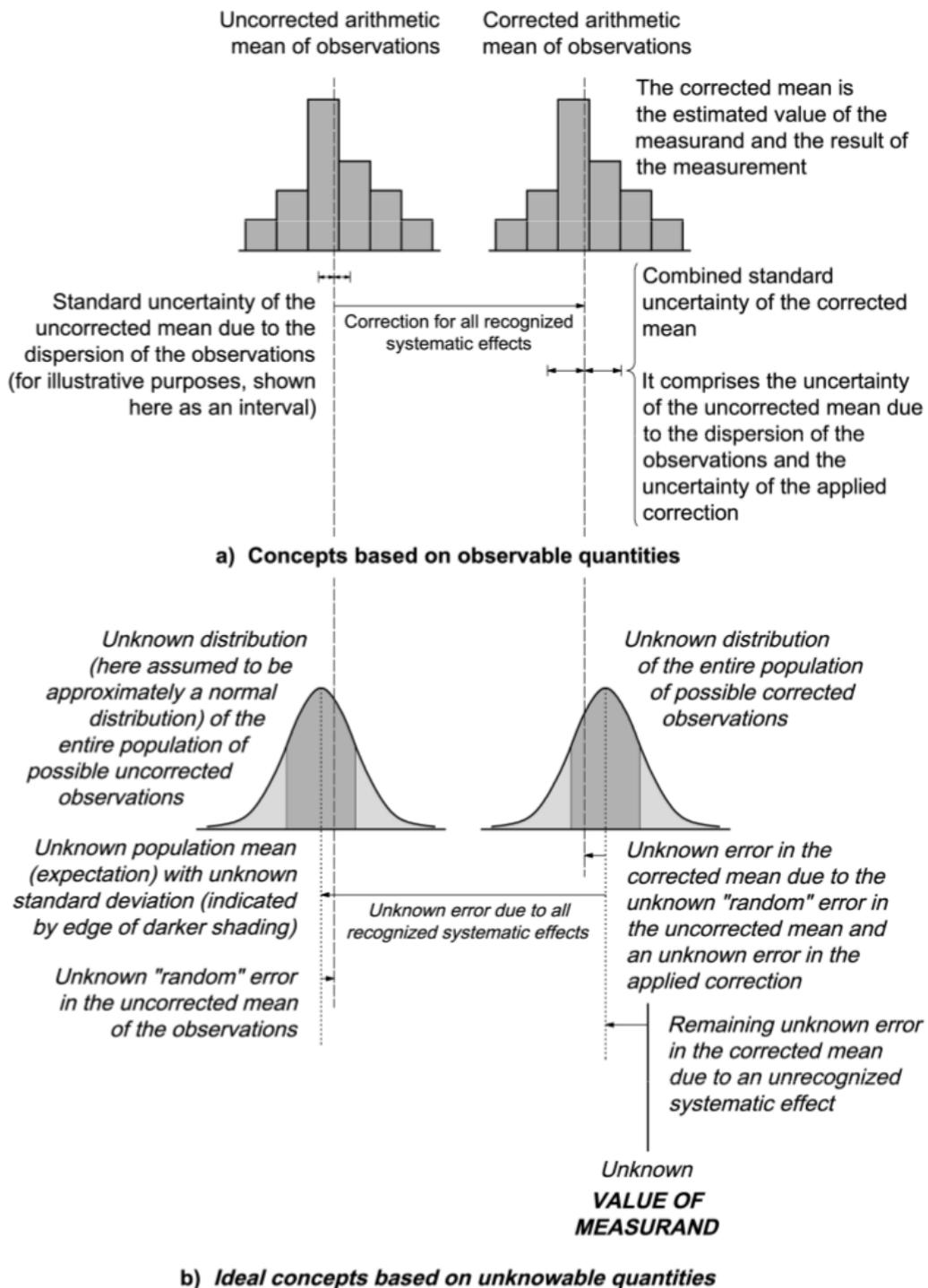


Figure 2.2: Illustration of the difference between error and uncertainty (source: JCGM 2008a).

2.1.3 Evaluating Uncertainty with Type A and Type B Methods

For evaluation of uncertainties, different sources of uncertainties are taken into consideration. Of the many possible sources of uncertainty in a measurement, the GUM gives the following examples:

- incomplete definition of the measurand
- imperfect realization of the definition of the measurand
- nonrepresentative sampling, i.e., the sample measured may not represent the defined measurand
- inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions
- personal bias in reading analogue instruments
- finite instrument resolution or discrimination threshold
- inexact values of measurement standards and reference materials
- inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm
- approximations and assumptions incorporated in the measurement method and procedure
- variations in repeated observations of the measurand under apparently identical conditions

For convenience of discussion, in the GUM uncertainty components are grouped into two categories (Type A and Type B) based on their method of evaluation. There is, however, no inherent difference in the nature of the components resulting from the two types of evaluation. Both types of evaluation are based on underlying probability distributions, and uncertainty components resulting from either type are quantified by variances. The categories are also not meant to substitute the words *random* and *systematic*. E.g., a known systematic effect should be corrected for, and the uncertainty of that correction needs to be evaluated, but can be obtained by either a Type A or a Type B evaluation (the same is true for uncertainties characterizing a random effect).

Generally a measurand X is a random variate, with a certain distribution. If the distribution is assumed to be Gaussian, then X can be characterized by an expected value $E[X]$ and a variance $E[(X - E[X])^2]$. But the expected value and variance of the population are generally unknown and need to be estimated. We term such an estimate for the expected value x and for the variance u_x^2 . In the more general case,

when the measurands are rank-1 random variates \mathbf{X} of length K , the measurand will be characterized by rank-1 estimate \mathbf{x} and a covariance matrix \mathbf{C}_x .

From here on, an element notation is used, i.e, random variables X_k , estimates x_k and (co-)variances $u_{x,kl}$ (where $u_{x,kl} := C_{x,kl} = [\mathbf{C}_x]_{kl}$ and $u_{x,kk} := u_{x,k}^2$) to avoid managing parallel scalar and vector notations. In the rank-1 case, $k, l = 1, 2, \dots, K$.

Often, the best available estimate x_k of the expected value of quantity X_k that varies randomly is determined by taking the arithmetic mean \bar{x}_k from a frequency based distribution (as shown in Figure D1) gained from M independent observations $X_{k,m}$ [check right letter for index] (with $m = 1, 2, \dots, M$) that have been obtained under the same conditions of measurement (JCGM 2008a),

$$\bar{x}_k = \frac{1}{M} \sum_{m=1}^M X_{k,m}. \quad (2.1)$$

As in Equation 2.1, the *uncertainty evaluation method Type A* is using the frequency based distribution information, to calculate a measure for the uncertainty of estimate \bar{x}_k , namely the experimental variance, $u_{\bar{x},k}^2$. For this the experimental variance of the observations, $u_{X_{m,k}}^2$, is determined,

$$u_{X_{m,k}}^2 = \frac{1}{M-1} \sum_{m=1}^M (X_{k,m} - \bar{x}_k)^2, \quad (2.2)$$

and from it the experimental variance of the arithmetic mean \bar{x}_k ,

$$u_{\bar{x},k}^2 = \frac{u_{X_{m,k}}^2}{M} \quad (2.3)$$

The experimental variance of the mean $u_{\bar{x},k}^2$ quantifies how well \bar{x}_k estimates the expectation $E[X_k]$ of variable X_k and may be used as a measure of the uncertainty of x_k . If the single observation is taken as estimate x_k , we take $u_{x,k}^2 = u_{X_{m,k}}^2$ as the variance. For convenience, $u_{x,k}^2$ is referred to as *Type A variance* and its positive square root $u_{x,k}$ as *Type A standard uncertainty* (JCGM 2008a).

The *Type B uncertainty evaluation methods* are not gained from repeated observations, but based on other, *a priori*, information on the distribution, obtained from knowledge like previous measurement data, experience with the behaviour and properties of the relevant materials and instruments, manufacturer's specifications, data provided in calibration certificates, and uncertainties assigned to reference data taken from handbooks (JCGM 2008a).

Even though the proper use of the pool of available information to determine the a priori distribution for a Type B uncertainty evaluation builds on insight based on experience and general knowledge, a Type B evaluation can be as reliable as a Type

A evaluation, when, e.g., a Type A evaluation is based on only a small number of independent observations (JCGM 2008a).

Since there is no difference in the nature of uncertainties estimated through these two types of evaluation, the GUM does not discriminate between them when calculating the uncertainty of a derived result of a measurement (i.e., when propagating uncertainties, as described in Section 2.2).

All Type A and Type B uncertainties of all influence quantities (the input quantities X_k to a measurement model) are added together to determine the *combined (standard) uncertainty* of the measurand (the output quantities Y_i), irrespective whether the uncertainty stems from random effects, or from uncertainty in correcting for systematic effects.

2.2 Measurement Models and Corresponding Uncertainty Propagation Algorithms

When a measurand Y_i is not measured directly, but determined from K other quantities X_k , the uncertainties of x_k , i.e., $u_{x,k}$, need to be propagated to y_i . Here $k = 1, 2, \dots, K$. The propagation algorithm depends on

- the linearity of the functional relationship $f(\cdot)$ between the measurand Y_i and the input quantities X_k ,
- the correlation between the input quantities X_k (in the multivariable case),
- the rank of the output quantity, i.e. either the functional relationship is univariate with a scalar Y , or multivariate, with rank-1 Y_i ($i = 1, 2, \dots, N$), and
- whether the functional relationship is explicit, $y_i = f(x_1, x_2, \dots, x_K)$, or implicit $g(y_1, \dots, y_i, \dots, y_N, x_1, \dots, x_k, \dots, x_K) = 0$.

Once we have evaluated the uncertainties of the input quantities, $u_{x,k}$, we propagate it from the input estimates x_k , through $f(\cdot)$, to the output estimates y_i . $f(\cdot)$ is called *measurement model* in the GUM, but will be referred to as *retrieval operator* here.

This propagation is repeated multiple times for each retrieval operator along the entire RO retrieval chain, to eventually determine the uncertainty of the ECVs.

2.2.1 Univariate, Linear Measurement Model

If an estimate of the measurand Y_i , denoted by y_i (the result of the measurement), is obtained from realizations of other variables, x_k , as is the case in an RO retrieval, we have a generally non-linear measurement model $f(\cdot)$,

$$y_i = f(x_1, x_2, \dots, x_k, \dots, x_K) \tag{2.4}$$

referred to as *state* retrieval operator (as opposed to the uncertainty propagation) in the rOPS.

In the rOPS these state retrieval operators are often linear though, i.e., of the form

$$y_i = \sum_{k=1}^K A_{ik} x_k, \quad (2.5)$$

then, according to the GUM, the uncertainty of the measurement is

$$u_{y,i}^2 = \sum_{k=1}^K A_{ik}^2 u_{x,k}^2, \quad (2.6)$$

when the input quantities x_k are completely uncorrelated (e.g. white noise). If the input quantities are correlated this generalizes to (JCGM 2008a)

$$u_{y,i}^2 = \sum_{k=1}^K A_{ik}^2 u_{x,k}^2 + 2 \sum_{k=1}^{K-1} \sum_{l=k+1}^K A_{ik} A_{il} u_{x,kl}, \quad (2.7)$$

where $l = 1, \dots, K$. Or, formulated with correlation coefficient $R_{x,kl} = u_{x,kl} / (u_{x,k} \cdot u_{x,l})$,

$$u_{y,i}^2 = \sum_{k=1}^K A_{ik}^2 u_{x,k}^2 + 2 \sum_{k=1}^{K-1} \sum_{l=k+1}^K A_{ik} A_{il} u_{x,k} u_{x,l} R_{x,kl}, \quad (2.8)$$

which is easier to interpret than the covariance elements $u_{x,kl}$. The correlation coefficient is an element of the correlation matrix \mathbf{R}_x , i.e., $R_{x,kl} := [\mathbf{R}_x]_{kl}$, and has properties $R_{x,kl} = R_{x,lk}$, and $-1 \leq R_{x,kl} \leq 1$. Equation 2.6 is a special case of Equation 2.8 with $R_{x,kl} = 0$ for all $k \neq l$. Another special case are perfectly correlated input variables, i.e. $R_{x,kl} = 1$ for all k, l ,

$$u_{y,i}^2 = \sum_{k=1}^K A_{ik}^2 u_{x,k}^2 + 2 \sum_{k=1}^{K-1} \sum_{l=k+1}^K A_{ik} A_{il} u_{x,k} u_{x,l}. \quad (2.9)$$

Then, using the multinomial theorem, i.e.,

$$\sum_{l_1+l_2+\dots+l_K=2} \frac{2}{l_1! l_2! \dots l_K!} \prod_{k=1}^K x_k^{l_k} = \left(\sum_{k=1}^K x_k \right)^2 \quad (2.10)$$

this is equal to

$$u_{y,i}^2 = \left(\sum_{k=1}^K A_{ik} u_{x,k} \right)^2, \quad (2.11)$$

and thus, in this case, the uncertainty propagation algorithm is just equal to the functional relationship for the state variables x_k and y_i (the linear state retrieval operator, Equation 2.5)

$$u_{y,i} = \sum_{k=1}^K A_{ik} u_{x,k}. \quad (2.12)$$

These uncertainties $u_{x,k}$ and $u_{y,i}$ are still standard deviations, i.e. statistical quantities, and the actual error realization (deviation from the 'true' value) is within $y \pm u_{y,i}$ to 68 %. If we know the error, $\epsilon_{x,k}$, and aim to find $\epsilon_{y,i}$, we add $\epsilon_{x,k}$ to the state estimate x_k and calculate the erroneous y_i ,

$$y_i + \epsilon_{y,i} = \sum_{k=1}^K A_{ik} (x_k + \epsilon_{x,k}), \quad (2.13)$$

substituting Equation 2.5 for y_i , subtracting it, and splitting up the sum on the right hand side, gives

$$\epsilon_{y,i} = \sum_{k=1}^K A_{ik} x_k + \sum_{k=1}^K A_{ik} \epsilon_{x,k} - \sum_{k=1}^K A_{ik} x_k = \sum_{k=1}^K A_{ik} \epsilon_{x,k}, \quad (2.14)$$

which is the same propagation algorithm as for standard uncertainties in the case of perfectly correlated input quantities.

2.2.2 Multivariate, Linear Measurement Model

Generalizing to the multivariate case, $i = 1, 2, \dots, N$, with correlated input and output variables, we now also want to know not just the variances of output estimates y_i , $u_{y,i}$, but also the covariances $u_{y,i,j} = u_{y,i} \cdot u_{y,j} \cdot R_{y,i,j}$ (here $j = 1, 2, \dots, N$). Hence (see e.g., Arras (1998) for a proof),

$$u_{y,i,j} = \sum_{k=1}^K \sum_{l=1}^K A_{ik} u_{x,kl} A_{jl} = \sum_{k=1}^K \sum_{l=1}^K A_{ik} A_{jl} u_{x,k} u_{x,l} R_{x,kl}, \quad (2.15)$$

where $k, l = 1, 2, \dots, K$. Using $u_{y,i,j} = C_{y,i,j} := [\mathbf{C}_y]_{ij}$ (and analogously for x_i), and $A_{ik} := [\mathbf{A}]_{ik}$, to formulate Equation 2.15 in matrix notation (JCGM 2011),

$$\mathbf{C}_y = \mathbf{A} \mathbf{C}_x \mathbf{A}^T. \quad (2.16)$$

For uncorrelated input quantities, Equation 2.15 would simplify to

$$u_{y,i,j} = \sum_{k=1}^K A_{ik} u_{x,kk} A_{jk}, \quad (2.17)$$

because all off-diagonal elements of matrix \mathbf{C}_x are zero, i.e. $u_{x,kl} = 0$ for all $k \neq l$.

2.2.3 Non-Linear Measurement Models

For non-linear measurement models of type

$$y_i = f(x_1, x_2, \dots, x_k, \dots, x_K), \quad (2.18)$$

Equation 2.15 can be applied, if $f(\cdot)$ is developed into a Taylor series to first order,

$$y_i = y_i(a) + \sum_{k=1}^K \left. \frac{\partial y_i}{\partial x_k} \right|_a (x_k - a) + O(x^2). \quad (2.19)$$

Then Equation 2.15 becomes

$$u_{y,i,j} \approx \sum_{k=1}^K \sum_{l=1}^K J_{ik} u_{x,kl} J_{jl}, \quad (2.20)$$

where \mathbf{J} is the Jacobi matrix with the partial derivatives (or *sensitivity coefficients*) as elements,

$$J_{ik} = \frac{\partial y_i}{\partial x_k}. \quad (2.21)$$

When $f(\cdot)$ is highly non-linear, higher order elements of the Taylor series need to be included (JCGM 2008a), or MC methods to be applied (JCGM 2008b).

2.2.4 Implicit Measurement Models

A final generalization is to consider implicit measurement models of type

$$g(y_1, \dots, y_i, \dots, y_N, x_1, \dots, x_k, \dots, x_K) = 0. \quad (2.22)$$

Equation 2.20 can still be applied for the uncertainty propagation, if the partial derivatives can be obtained through implicit differentiation, i.e., through applying

$$\frac{\partial y_i}{\partial x_k} = - \frac{\partial g(x_k)/\partial x_k}{\partial g(y_i)/\partial y_i}. \quad (2.23)$$

2.2.5 Validation of Propagated Uncertainties

Finally, when covariance propagation as described in the GUM framework is difficult or impossible, in particular when

- a linearization of $f(\cdot)$ provides an inadequate representation of the measurement model, or

- the distribution of the output quantity departs considerably from a Gaussian distribution,

a Monte Carlo (MC) method can replace the uncertainty propagation of the GUM framework.

In case of the rOPS all retrieval operators are sufficiently linear, and the input variables sufficiently normally distributed, and MC methods are only applied to test and validate the GUM-type uncertainty propagation.

For this purpose, the multivariate Gaussian probability distribution of the input quantities \mathbf{X} , characterized by mean $\bar{\mathbf{x}}$ and covariance matrix \mathbf{C}_x , is sampled by a large number M of representative draws from it. These draws are calculated according to (JCGM 2008b),

$$\mathbf{x}_m = \bar{\mathbf{x}} + \mathbf{Q}_x \cdot r_m = \bar{\mathbf{x}} + (\mathbf{U}_x \mathbf{\Lambda}_x^{1/2}) \cdot r_m, \quad (2.24)$$

where r_m is a vector of standard normal distributed random numbers, m the sample index with $m = 1, \dots, M$, and \mathbf{Q}_x , the Cholesky decomposite of \mathbf{C}_x . \mathbf{Q}_x is calculated through spectral decomposition, using \mathbf{U}_x , the matrix of eigenvectors, and $\mathbf{\Lambda}_x^{1/2}$, a diagonal matrix of the square roots of the eigenvalues of \mathbf{C}_x .

These representative draws from the probability distribution of the input quantities \mathbf{X} are then run through all operators of the state retrieval and recombined to covariance matrices after each step. Thus, for each output variable \mathbf{Y} the sample results \mathbf{y}_m are first used to calculate a mean mean profile \mathbf{y}^{MC} , of \mathbf{Y} ,

$$\mathbf{y}^{\text{MC}} = \frac{1}{M}(\mathbf{y}_1 + \dots + \mathbf{y}_M), \quad (2.25)$$

and then the covariance matrix \mathbf{C}_y^{MC} ,

$$\mathbf{C}_y^{\text{MC}} = \frac{1}{M-1}[(\mathbf{y}_1 - \mathbf{y}^{\text{MC}})(\mathbf{y}_1 - \mathbf{y}^{\text{MC}})^T + \dots + (\mathbf{y}_M - \mathbf{y}^{\text{MC}})(\mathbf{y}_M - \mathbf{y}^{\text{MC}})^T]. \quad (2.26)$$

This output covariance matrix \mathbf{C}_y^{MC} is then compared to the GUM-type propagated uncertainty characterized by \mathbf{C}_y .

2.3 Uncertainty Propagation in the rOPS

2.3.1 Sources of Uncertainties in the rOPS

In the rOPS uncertainty trace the main attention is paid to the uncertainty of the measured input quantity of the propagation chain (also called parametric uncertainty or input uncertainty), and to how this uncertainty propagates to the ECVs. For the uncertainty propagation from L1b to L2b (the scope of this thesis), this was the uncorrelated (white) noise and the bias of the excess phase profiles $L_r(t)$ on the one hand, and uncertainties

of the background quantities used in the retrieval, e.g., $\alpha_b(z_a)$, $T_b(z)$, $q_b(z)$, on the other. However, other sources of uncertainty were also taken into consideration, including (Kennedy and O'Hagan 2001):

Algorithmic uncertainty comes from numerical approximations in the implementation of an analytical model.

Most models are too complicated to be solved exactly, thus approximations in the numerical implementation are necessary, but cause errors, e.g., interpolations, numerical integration and infinite sum truncations. In the rOPS particularly the numerical integrations over exponential quantities can cause biases. E.g., through the baseband approach these uncertainties were intended to be reduced to below a relative threshold of 10^{-4} , below which they were not considered in the uncertainty propagation. This improvement of the forward-inverse reversability, can be tested through the BUC tests, as demonstrated in Figures 1.8 and 1.9. Other measures, such as an adjusted grid spacing, or the grid matching explained in Section 1.5.4, also contributed to reducing algorithmic uncertainty considerably.

Structural uncertainty comes from model inadequacy or model bias, which in turn comes from the lack of knowledge of the underlying true physics.

It depends on how accurately a mathematical model describes the true system for a real-life situation, considering the fact that models are almost always only approximations to reality. In the rOPS, these uncertainties (e.g., due to the spherical symmetry assumption in the refractivity retrieval) are added to the propagated uncertainties after the latter have been propagated through the respective retrieval operator.

Parameter uncertainty comes from model parameters and whose exact values are unknown.

These uncertainties are difficult to take into account in the propagation, because they are part of the measurement models $f(\cdot)$ applied onto the uncertainties of the input quantities. In the rOPS different approaches are taken to reduce these parameter uncertainties. In some cases (e.g. the uncertainty in the geoid undulation) the parameter uncertainties are added after the propagation (similar to the structural uncertainty), in other cases these uncertainties are eliminated before (e.g. grid uncertainties are transferred to the state quantity, and thus, when the grid appears as a parameter in the retrieval operator—as e.g. in the Abel integral, or the pressure integral—it can then be considered free of uncertainty by construction).

2.3.2 Estimated Random and Systematic Uncertainties

In the rOPS uncertainty propagation, the uncertainties are categorized into two groups

- the uncertainties that can be evaluated by method Type A (because due to the oversampling we have M independent observations of the same conditions as in Equation 2.2), are estimated and propagated in one group, and are called *estimated random uncertainties*
- the uncertainties which we can not be evaluated by method Type A (because they are due to long-range correlated effects and the M observations in Equation 2.2 are thus not independent and can not capture these effects, which are consequentially evaluated using Type B methods. Since these components are in essence caused by systematic effects - at least for one single event - we call these *estimated systematic uncertainties*.

The (JCGM 2008a) advises to avoid this separation, because uncertainties from systematic effects are not any different in nature to those from random effects and should generally not be treated differently. Known systematic effects should be corrected for, and the uncertainty with respect to the magnitude of the correction should be added to the combined uncertainty.

However, because for these systematic effects the correlation length is higher than the physical atmospheric resolution and thus Type A determination on a single RO event through oversampling is impossible, the 'systematic effects' are essentially identical to the Type B evaluated uncertainties. Because a Type A evaluation might still be possible for those components of the estimated systematic uncertainty that differ between multiple RO events, we further separate the estimated systematic uncertainties, and call these effects *apparent systematic uncertainties*. Effects which are systematic even over a sample of multiple RO events, are referred to as *basic systematic uncertainties*

This separation in the rOPS can be justified through the following arguments:

Consistency Since it can be justifiably assumed that the uncertainty components termed *systematic* and *random* are *uncorrelated*, both can be propagated independently and only combined in the end and still deliver results consistent with a propagation of just one combined uncertainty.

Climatologies Since the main motivation for the rOPS uncertainty propagation is to establish traceable accuracy in benchmark observations for climate applications, it matters whether an effect is likely systematic for all RO events used for a climatology or not, because they need to be aggregated differently. While basic systematic uncertainties cause biases in the climatology, apparent systematic and random uncertainties can partly be offset in ensembles of RO events.

Information Content Particularly because the Type B (estimated systematic) uncertainties are evaluated using experts estimates and best guess quantifications, their magnitudes are more uncertain. The evaluation method is thus likely subject to

improvement in the future, and the separation is helpful to know which part of the final ECV uncertainty depends on these estimates.

For the uncertainty propagation from excess phase to the ECVs (L1b to L2b processing) the input will eventually be provided by the output of the L1a uncertainty propagation, but for this thesis the estimated random and systematic uncertainties had to be evaluated at the excess phase level. The random uncertainty was estimated based on the noise of $L_r(t)$. We did not evaluate the arithmetic mean of multiple observations (Equation 2.1) and take it as estimate $x_k = \bar{x}_k$, but for simplicity either took the actual observation $x_k = X_{k,m}$, or corresponding background information.

Consequentially, the random uncertainty is essentially estimated using Equation 2.2, rather than Equation 2.3. The systematic uncertainty profile of $L_r(t)$ is estimated based on an error model from ESA/EUMETSAT (1998). A detailed description of the estimation of random and systematic uncertainties at excess phase level is provided in Section 3.3.2.

The random uncertainty $u_{y,i}$ is usually propagated in form of the covariance matrix \mathbf{C}_y . After each retrieval operator, the covariance matrix can again be decomposed into the random uncertainty profile $u_{y,i}$ through

$$u_{y,i} = \sqrt{C_{y,ii}}, \quad (2.27)$$

and into a correlation matrix \mathbf{R}_y , using

$$R_{y,ij} = \frac{C_{y,ij}}{u_{y,i} \cdot u_{y,j}}. \quad (2.28)$$

From the correlation matrix the correlation length $l_{y,i}$ can be determined (see Appendix A for more details).

2.3.3 Operator Types and Corresponding Propagation Rules

Three aspects make uncertainty propagation in the rOPS possible (or at least easier):

- thanks to the baseband formulation of most retrieval operators, algorithmic uncertainties are very small (below the 10^{-4} threshold), and thus do not need to be explicitly considered in the uncertainty propagation,
- the chain-like processing avoids feedback loops and iterations and thus simplifies the propagation algorithm to a step by step propagation,
- most retrieval operators in the rOPS are linear. Even the slightly non-linear operators can be easily linearized and higher-order terms in the Taylor expansion do not have to be taken into account.

Depending on the type of uncertainty and retrieval operator type, we apply different uncertainty propagation rules from Section 2.2 in the rOPS.

First, we distinguish between the propagation rules for estimated random and estimated systematic uncertainties. While the general case for the random uncertainty is the linear correlated multivariate case (Equation 2.15 and 2.16) for rOPS processors L1b and L2a, and the non-linear correlated multivariate case (Equation 2.20) in the L2b processor, the *estimated systematic uncertainty* is considered to apply perfectly correlated errors (biases) over the length of the profile. Thus Equation 2.12 can be applied, and the systematic uncertainties are propagated using the state retrieval operators. Even if the systematic effects are not perfectly correlated along the entire profile, this treatment can be practically justified, because most operators give the strongest weight to the same level or adjacent levels. Only the pressure integral weights all above-elements with the same weight, but there the upper uncertainties are considerably smaller than lower elements, and are thus not so relevant.

In case of RO retrievals and particularly in the rOPS, most retrieval operators are multivariate and linear, and (the estimates of) the rank-1 measurand Y_i , i.e., the output quantities y_i , are a (vertical) profile either as a function of time t , impact parameter a , impact altitude z_a or MSL altitude z . The profile is of length N and has elements y_i , with $i = 1, 2, \dots, N$. The input quantities x_k are always either one profile of equal length $K = N$ or multiple input profiles, each of length N , i.e., $K = P \cdot N$, where P is the number of input profiles. Generally, most operators of the L1b and L2a processors have N correlated input variables, and N correlated output variables. For the operators of the L2b processor however, P can be up to 4.

If $P > 1$, a propagation based on Equation 2.12 is inadequate also for estimated systematic uncertainties, i.e., long-range correlated uncertainties, because the correlation typically is only high between the different altitude levels of the same input profile (intra-variable correlation), while the different input profiles are uncorrelated to each other (inter-variable correlation).

Table 2.1 summarizes briefly how random uncertainties were propagated from excess phase profiles $L_r(t)$, to the ECVs density $\rho_e(z)$, pressure $p_e(z)$, temperature $T_e(z)$ and specific humidity $q_e(z)$. The (almost) complete retrieval algorithm and uncertainty propagation algorithm is provided in the Appendices A (L1b), B (L2a), and C (L2b). The variable names used are defined in Tables 3.1, 3.2, for L1b, in Tables 4.1, 4.2, 4.3, for L2a, and in Tables 5.1 for L2b. Figures 3.2 (L1b), 4.2 (L2a), and 5.2 (L2b) provide a schematic illustration of the workflow of the state retrieval and the uncertainty propagation through these processors. Details are described in the respective chapters and related Appendices.

In the rOPS the propagation of the random uncertainties is validated through MC methods as described in Section 2.2.5. The results of these validations are shown in Figures 3.9 (L1b), and 4.8, and 4.9 (L2a).

Table 2.1: Sequence of the retrieval operators of the rOPS retrieval chain, their input profiles and output profiles, the number of input profiles P to the uncertainty propagation step, and the source for the uncertainty propagation rule applied. L1b profiles up to $\alpha_r(z_{at})$ denote profiles for both transmitter frequencies f_{T1} and f_{T2}

Operator input	Output	Operator description	P	Propagation
L1b processor				
$L_r(t)$	$L_F(t)$	Basic BWS filter	1	Equation 2.16
$L_F(t)$	$D_r(t)$	Doppler derivative	1	Equation 2.16
$D_r(t)$	$\alpha_G(a)$	GO retrieval	1	Equation 2.8
$\alpha_G(a)$	$\alpha_G(z_{at})$	Common grid interp.	1	Interpolation
$L_r(t), A_r(t)$	$\alpha_W(z_{at})$	WO retrieval	2	Section A.1.2
$\alpha_G(z_{at}), \alpha_W(z_{at})$	$\alpha_M(z_{at})$	GO-WO merger	2	Equation 2.16
$\alpha_M(z_{at})$	$\alpha_F(z_{at})$	Adaptive BWS filter	1	Equation 2.16
$\alpha_{F1}(z_{at}), \alpha_{F2}(z_{at})$	$\alpha_r(z_{at})$	Ionospheric correction	2	Equation 2.16
L2a processor				
$\alpha_r(z_{at})$	$\alpha_E(z_{at})$	Equalizing BWS filter	1	Equation 2.16
$\alpha_E(z_{at})$	$\alpha_r(z_a)$	Downsampling	1	Interpolation
$\alpha_r(z_a), \alpha_b(z_a)$	$\alpha_s(z_a)$	High altitude initiali.	2	Rodgers (2000)
$\alpha_s(z_a)$	$N_r(z_a)$	Refractivity retrieval	1	Equation 2.16
$N_r(z_a)$	$N_r(z)$	Grid adjustment	1	Interpolation
$N_r(z)$	$\rho_r(z)$	Smith Weintraub eq.	1	Equation 2.16
$\rho_r(z)$	$p_{dr}(z)$	Pressure integration	1	Equation 2.16
$\rho_r(z), p_{dr}(z)$	$T_{dr}(z)$	Equation of State	2	Equation 2.20
L2b processor				
$p_{dr}(z)$	$p_q(z)$	Pressure derivation	1	Equation 2.20
$p_{dr}(z), p_q(z), T_{dr}(z), q_b(z)$	$T_q(z)$	Temperature retrieval	4	Equation 2.23
$p_{dr}(z)$	$p_T(z)$	Pressure derivation	1	Equation 2.20
$p_{dr}(z), p_T(z), T_{dr}(z), T_b(z)$	$q_T(z)$	Specific humidity retr.	4	Equation 2.20
$T_q(z), T_b(z)$	$T_e(z)$	Optimal estimation	2	Rodgers (2000)
$q_T(z), q_b(z)$	$q_e(z)$	Optimal estimation	2	Rodgers (2000)
$q_e(z)$	V_{we}	Volume mixing ratio d.	1	Equation 2.20
$p_{dr}(z)$	$p_e(z)$	Pressure derivation	1	Equation 2.20
$p_e(z), V_{we}(z)$	$e_e(z)$	Partial pressure deriv.	2	Equation 2.20
$q_e(z), T_e(z), p_e(z)$	$\rho_e(z)$	Density derivation	3	Equation 2.20

Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

This chapter presents work that was published in *Atmospheric Measurement Techniques (AMT)* by Schwarz, Kirchengast and Schwärz (2018). The work in the context of this publication included design and implementation of uncertainty propagation in the Reference Occultation Processing System (rOPS) retrieval, starting from excess phase level, through the Doppler shift retrieval, the bending angle retrieval, and the atmospheric bending angle derivation.

My contribution to this study included the analytical derivation of the uncertainty propagation formulas, the conversion into numerical propagation algorithms, the implementation of the algorithms into the rOPS, the testing and validation of the algorithms and the drafting of the manuscript for the publication.

3.1 Abstract

Global Navigation Satellite System (GNSS) radio occultation (RO) observations are highly accurate, long-term stable data sets and are globally available as a continuous record from 2001. Essential climate variables for the thermodynamic state of the free atmosphere – such as pressure, temperature, and tropospheric water vapor profiles (involving background information) – can be derived from these records, which therefore have the potential to serve as climate benchmark data. However, to exploit this potential, atmospheric profile retrievals need to be very accurate and the remaining uncertainties quantified and traced throughout the retrieval chain from raw observations to essential climate variables. The new Reference Occultation Processing System (rOPS) at the Wegener Center aims to deliver such an accurate RO retrieval chain with integrated

uncertainty propagation. Here we introduce and demonstrate the algorithms implemented in the rOPS for uncertainty propagation from excess phase to atmospheric bending angle profiles, for estimated systematic and random uncertainties, including vertical error correlations and resolution estimates. We estimated systematic uncertainty profiles with the same operators as used for the basic state profiles retrieval. The random uncertainty is traced through covariance propagation and validated using Monte Carlo ensemble methods. The algorithm performance is demonstrated using test day ensembles of simulated data as well as real RO event data from the satellite missions CHALLENGING Minisatellite Payload (CHAMP); Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC); and Meteorological Operational Satellite A (MetOp). The results of the Monte Carlo validation show that our covariance propagation delivers correct uncertainty quantification from excess phase to bending angle profiles. The results from the real RO event ensembles demonstrate that the new uncertainty estimation chain performs robustly. Together with the other parts of the rOPS processing chain this part is thus ready to provide integrated uncertainty propagation through the whole RO retrieval chain for the benefit of climate monitoring and other applications.

3.2 Introduction

Observation systems of the free atmosphere, focusing on the range from the top of the atmospheric boundary layer upwards, were historically designed for weather research and forecasting purposes. They have considerable shortcomings from a climate monitoring perspective (Karl et al. 1995), and so the related global climate monitoring infrastructure remains fragile and incomplete even today (Bojinski et al. 2014). The Global Climate Observing System (GCOS) aims to improve the observational foundation for the climate sciences (GCOS 2015). For this purpose the establishment of *climate benchmark data records* is essential. To qualify as climate benchmark, records need to be (1) of global coverage, (2) of high accuracy, (3) long-term stable, (4) tested for systematic errors on orbit, (5) and tied to irrefutable standards, and they need to (6) measure Essential Climate Variables (ECVs) (GCOS 2015; NRC 2007).

Based on the quality and abundance of Global Navigation Satellite System (GNSS) signal sources, in particular from the Global Positioning System (GPS) so far, the GNSS radio occultation (RO) observation record is globally available (continuously since 2001), long-term stable (due to the so-called self-calibration and high signal stability during the event), and highly accurate (accuracy traceable to the SI second). Due to the self-calibrating property, the accuracy is also ensured on orbit; i.e., there is no need for calibration or bias correction in post-processing on ground (Leroy et al. 2006). The basic RO excess phase data can therefore serve as a fundamental climate data record (FCDR) as defined by GCOS (2010a). From this FCDR with its unique properties, ECVs – in particular the thermodynamic ECVs pressure, temperature, and humidity in the free

atmosphere – can be derived using an RO *retrieval* chain.

In order to reliably serve as climate benchmark data record, however, the retrieved ECV profiles and their claimed accuracy – expressed by the uncertainties provided – need to be traceable back to the (small) uncertainties of the FCDR and in turn to the raw data. This requires that (1) the RO retrieval is highly accurate and avoids any undue amplification of uncertainties associated with the quantities in the FCDR and that (2) the uncertainties are propagated through the entire retrieval chain, from the raw data to the ECV profiles, duly accounting for relevant side influences such as from background information. Developed at the Wegener Center of the University of Graz (WEGC), together with international partners, the Reference Occultation Processing System (rOPS) (Kirchengast et al. 2015) aims to establish such a fully traceable RO processing for the first time (Kirchengast et al. 2016a; Kirchengast et al. 2016b).

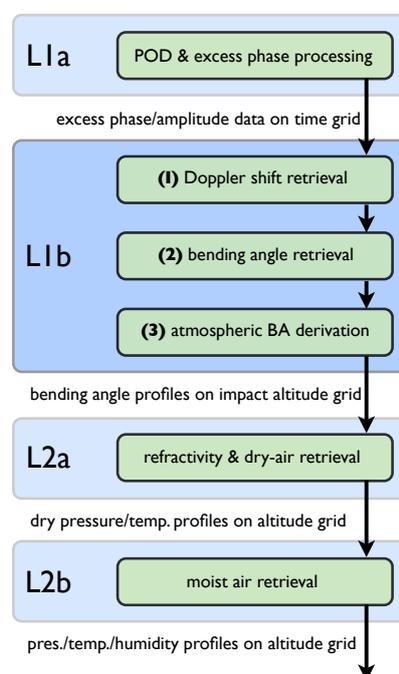


Figure 3.1: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b highlighted, L2a, L2b) and the main operators of the L1b processor (1, 2, 3), which are in the focus of this study.

In Figure 3.1 the basic steps of the RO retrieval chain in the rOPS, i.e., the precise orbit determination (POD) and excess phase processing (labeled “L1a” in Figure 3.1), the subsequent atmospheric bending angle retrieval (“L1b”), the refractivity and dry-air retrieval (“L2a”), and the moist-air retrieval (“L2b”) are sketched.

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

Kursinski et al. (1997) – and more recently Hajj et al. (2002), Anthes (2011), and Steiner et al. (2011) – provided detailed introductions and reviews of the RO technique and its applications in meteorology and climate. Ho et al. (2012) and Steiner et al. (2013) included comparative current RO retrieval chain descriptions of leading international RO processing centers, none of which yet include uncertainty propagation. Empirical error (uncertainty) estimates computed statistically from retrieved RO atmospheric profiles and climatologies have been derived by Kuo et al. (2004), Steiner and Kirchengast (2005), and Scherllin-Pirscher et al. (2011a), Scherllin-Pirscher et al. (2011b) and Scherllin-Pirscher et al. (2017), the last of which with a focus on climate uses also providing simple analytical error models. These studies and many others have described the RO retrieval chain in detail and have shown the high accuracy and quality of RO data, particularly in the upper-troposphere and lower-stratosphere region.

The aim of the integrated uncertainty propagation in the rOPS is to eventually propagate uncertainties along this *entire* retrieval chain from the raw measurement data to the ECVs (Kirchengast et al. 2016a; Kirchengast et al. 2016b), whereby the implementation of the rOPS uncertainty propagation occurs in the sequential blocks illustrated in Figure 3.1. The L2a processing and uncertainty propagation from atmospheric bending angle to dry-air profiles has already been introduced by Schwarz et al. (2017, SKS2017 hereafter).

This study is a direct complement to the work in SKS2017. Using the same propagation and validation methods as applied in SKS2017, it focuses on the uncertainty propagation from excess phase to atmospheric bending angle profiles, i.e., the L1b processing. As in SKS2017, *random uncertainties* are propagated using covariance propagation (CP) and validated using Monte Carlo (MC) ensemble methods. As in the L2a processor, we also propagate (conservative bound) estimates for *systematic uncertainties* along the retrieval chain of the L1b processor. Additionally, *correlation length* profiles and *resolution* profiles are provided.

Uncertainty propagation as covariance propagation from excess phase to bending angle profiles has been outlined and demonstrated in a basic form, by Syndergaard (1999) and Rieder and Kirchengast (2001a), but not been implemented yet in processing center retrieval chains or applied to real RO data. As visible in Figure 3.1, the L1b processor consists of three major retrieval parts, which are expanded into detailed substructure in Figure 3.2. We propagate estimated random uncertainties from excess phase profiles to Doppler shift profiles (Sect. (1) in Figure 3.2); further to geometric-optics (GO) bending angle profiles, merged with wave-optics (WO) bending angle profiles (2); and finally to atmospheric bending angle profiles (3), using a full-CP approach. In combination with the definitions of the main operators and variables in Table 3.1, and of the vertical grid and coordinate variables in Table 3.2, Figure 3.2 provides a concise overview on the detailed workflow of the L1b processor.

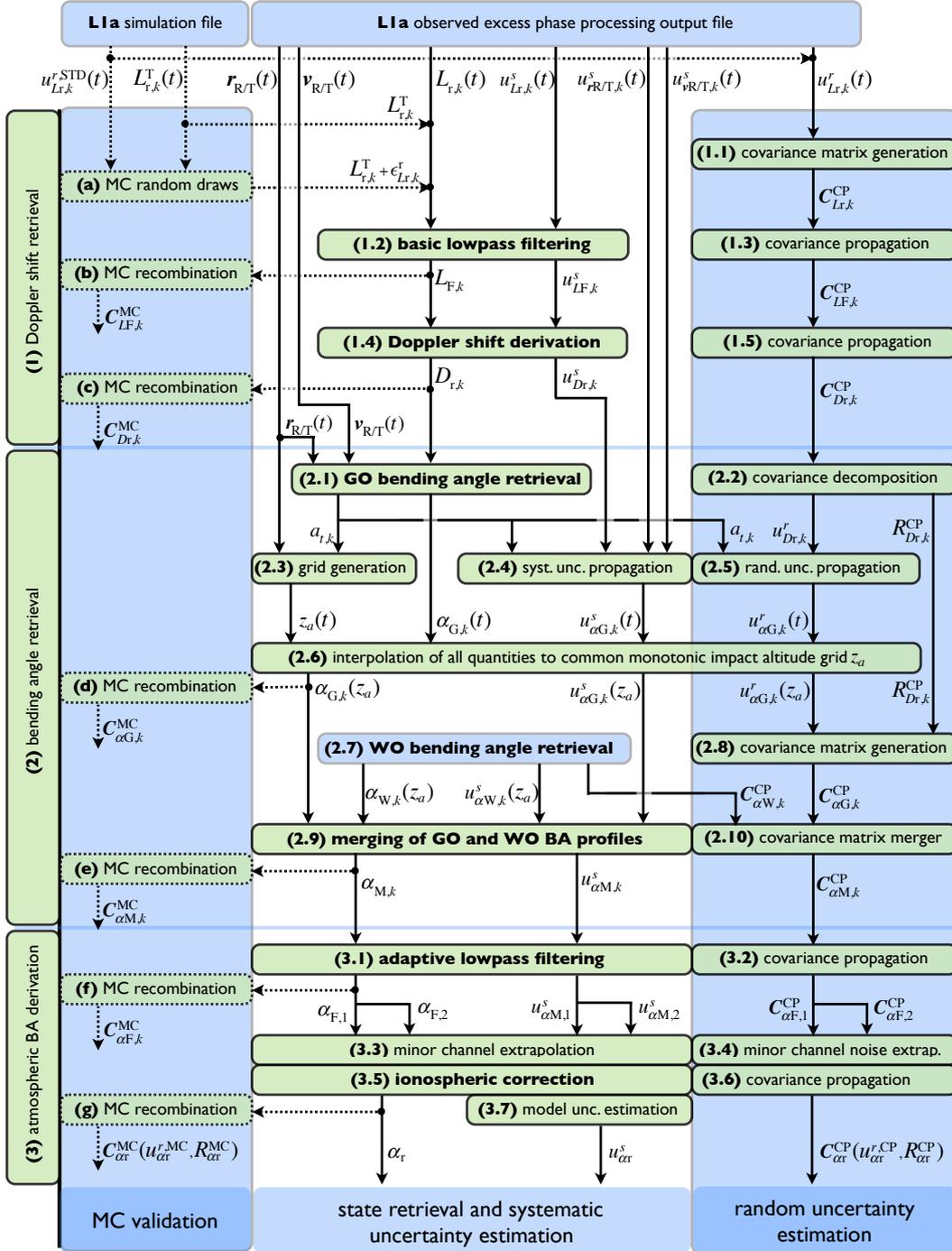


Figure 3.2: Detailed workflow for state retrieval and uncertainty propagation of the main L1b operators from excess phase to atmospheric bending angle profiles (1)–(3) and of the sub-routines used in the MC testing framework (a)–(g). The mathematical notation, including all symbols, is introduced in Tables 3.1 and 3.2.

Uncertainty propagation for the WO bending angle retrieval has been implemented and demonstrated for simulated events by Gorbunov and Kirchengast (2015); estimation of random and systematic uncertainties for real events including boundary layer bias correction is introduced by Gorbunov and Kirchengast (2018).

Other ongoing rOPS retrieval advancements relevant to this study are the inclusion of the high-altitude initialization algorithm, introduced by Li et al. (2013) and Li et al. (2015), in the L2a processor and the reduction of remaining higher-order ionospheric effects in the retrieved bending angle profiles of the L1b processor (based on work by (Syndergaard 2000); (Liu et al. 2015); (Healy and Culverwell 2015); and (Danzer et al. 2013; Danzer et al. 2015)). Furthermore, the precise orbit determination (POD) of the RO receiver satellite and the excess phase processing, also including the associated uncertainty propagation, are part of ongoing work (Innerkofler et al. 2017b).

Finally, related work and manuscript preparation on a new moist-air retrieval algorithm (L2b) and corresponding L2b uncertainty propagation are ongoing (Kirchengast et al. 2017a; Li et al. 2018).

The paper is structured as follows. In Sect. 3.3 we introduce the uncertainty estimation, propagation, and validation methods and the data sources and preparation. In Sect. 3.4, with the help of an example RO event, the uncertainty propagation sequence is introduced. In Sect. 3.5 we present the results from the MC validation of the CP uncertainty estimates. In Sect. 3.6 the performance of the algorithm is then evaluated using test day ensembles with real data from the RO missions CHALLENGING Minisatellite Payload (CHAMP) (Wickert et al. 2001); FORMOSAT-3 Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) (Anthes et al. 2008); and Meteorological Operational Satellite A (MetOp) (Luntama et al. 2008); as well as with simulated data approximating characteristics of the Meteorological Operational Satellite A ((Luntama et al. 2008); simMetOp data hereafter). We close with conclusions and outlook in Sect. 3.7. A detailed description of the implemented uncertainty propagation algorithms can be found in Appendix A.1.

Table 3.1: Principal variables for the rOPS L1b uncertainty propagation.

Variable	Unit	Description
X_r	U	state profile of retrieved excess phase/filtered excess phase/retrieved Doppler shift/retrieved geometric optics bending angle/merged GO WO bending angle/filtered bending angle/retrieved bending angle, with $X_r \in \{L_{r,k}(t), L_{F,k}(t), D_{r,k}(t), \alpha_{G,k}(t), \alpha_{G,k}(z_a), \alpha_{M,k}(z_a), \alpha_{F,k}(z_a), \alpha_r(z_a)\}$, $k \in \{1, 2\}$ (frequencies f_{T1}, f_{T2}), and unit $U \in \{m, m s^{-1}, rad, rad, rad, rad, rad\}$, comprising elements $X_{r,i}$.
u_X^s	U	estimated systematic uncertainty profile of X (with X and U as defined above), comprising elements $u_{X,i}^s$ (including estimated basic and estimated apparent systematic uncertainties, $u_{X,i}^b$ and $u_{X,i}^a$).
u_X^r	U	estimated random uncertainty profile of X (with X and U as defined above), comprising elements $u_{X,i}^r$.
\mathbf{R}_X	1	error correlation matrix of X (with X as defined above), comprising elements $R_{X,ij}$.
\mathbf{C}_X	U ²	error covariance matrix of X (with X and U as defined above), comprising elements $C_{X,ij} = u_{X,i}^r \cdot u_{X,j}^r \cdot R_{X,ij}$.
l_X	m	correlation length profile of X (with X as defined above), comprising elements $l_{X,i}$.
τ_X	s	resolution profile of X (with X as defined above) in time domain, comprising elements $\tau_{X,i}$.
w_X	m	resolution profile of X (with X as defined above) in altitude domain (along impact altitude), comprising elements $w_{X,i}$.
X_m	U	model excess phase/Doppler shift/bending angle profiles based on forward modeling of collocated refractivity profiles from ECMWF short-range forecast fields, with $X_m \in \{L_m(t), D_m(t), \alpha_m(z_a)\}$, and $U \in \{m, m s^{-1}, rad\}$, comprising elements $X_{m,i}$.
\vec{x}_S	U	profiles of Cartesian position/velocity vectors of the receiving/transmitting satellite relative to the center of refraction, with $\vec{x}_S \in \{\vec{r}_T(t), \vec{r}_R(t), \vec{v}_T(t), \vec{v}_R(t)\}$, and unit $U \in \{m, m, m s^{-1}, m s^{-1}\}$, comprising elements $\vec{x}_{S,i}$.
$u_{\vec{x}_S}^s$	U	estimated (systematic) uncertainty profiles of \vec{x}_S (with \vec{x}_S and U as defined above), comprising elements $u_{\vec{x}_S,i}^s$.
\mathbf{A}^{BWS}	1	BWS filter matrix operator, comprising the Blackman windowed sinc (BWS) low-pass filter weights (normalized filter functions) in the form of a band matrix.
\mathbf{A}^{L2D}	s ⁻¹	Doppler differentiation matrix operator, transforming the filtered excess phase profile to the Doppler shift profile.

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

Table 3.2: Vertical grids, coordinate variables, and specific settings for the rOPS L1b processing system.

Variable	Unit	Description
f_T	Hz	transmitter signal carrier frequency, with elements f_{T_k} (for GPS transmitters $k \in \{1, 2\}$ denoting the L-band frequencies $f_{T1} = 1.57542$ GHz and $f_{T2} = 1.22760$ GHz).
f_s	Hz	measurement sampling frequency (also called sampling rate); 50 Hz is generally used for the input excess phase profiles.
f_c	Hz	Blackman windowed sinc (BWS) low-pass filter cutoff frequency; set to 2.5 Hz (but noise-dependent for the $f_{T(1)2}$ filtering for ionospheric correction, with $f_{c(1)2} \in \{2.5, 2, 10/7, 1, 5/7, 0.5 \text{ Hz}\}$).
t	s	time grid of the measurements at sampling rate f_s , with elements t_i , $i \in \{1, 2, \dots, N\}$, where N is the number of grid points of the RO profile.
a_t	m	impact parameter grid corresponding to time grid t .
z_a	m	common monotonic impact altitude grid, calculated from sorted impact parameters $a_{t,i}$ of the leading channel (f_{T1}) bending angle, via $z_{a,i} = a_{t,i} - h_G - R_C$. Used as standard vertical grid after interpolation of all dependent quantities to z_a .
z_t	m	MSL altitude grid corresponding to time grid t , obtained as part of the forward modeling towards α_m , D_m , and L_m (cf. Table 3.1).
z_{aTop}	m	impact altitude of the top of the RO profile; it can lie between 70 and 80 km.
z_{aBot}	m	impact altitude of the bottom of the RO profile; it can lie between 25 km and the Earth's surface. Its value can be different for the different GNSS frequencies (i.e., $z_{aBot,k}$, for $k \in \{1, 2\}$).
z_a^{GW}	m	impact altitude at the center of the sinusoidal transition range of half-width Δz_a^{GW} between the GO and WO bending angle profiles; z_a^{GW} can lie within 9 and 14 km, depending on GO bending angle data quality.
Δz_a^{GW}	m	impact altitude transition half-width of the half-sine-weighted transition between the GO and WO bending angle profile. Set to 2 km.
z_{aGradr}	m	impact altitude at the lower end of the excess phase uncertainty estimation range used in this study, below which the estimated random uncertainties are extended by a linear gradient. Set to 30 km.
z_{aGrads}	m	impact altitude at the lower end of the range with constant excess phase systematic uncertainty used in this study, below which the estimated systematic uncertainties continue with a linear gradient. Set to 8 km.

3.3 Methods and data

3.3.1 Methods

We follow the *Guide to the Expression of Uncertainty in Measurement* (JCGM 2008a; JCGM 2008b; JCGM 2011) [GUM hereafter] and aim to follow terminology as provided by the International Vocabulary of Metrology (JCGM 2012), a terminology also adopted by the GUM. SKS2017 provides a more thorough introduction, including the motivation for using the respective uncertainty estimation, propagation, and validation methods; we refer the particularly interested reader to this companion (open-access) work and provide the essential methods needed more in a summarized form below.

We categorize uncertainties into *estimated random uncertainties* and *estimated systematic uncertainties*. Effects of unpredictable or stochastic temporal and spatial variations in repeated observations, like effects from fluctuations in the atmosphere or the thermal noise of the receiver system, could in principle be estimated by ensemble statistics from multiple RO events. However, since such effects are essentially stationary in a statistical sense, we can estimate their statistics also from individual RO event data, given their high noise-resolving sampling rate. These effects are included in the estimated random uncertainties.

Systematic effects (biases), which can not be quantified using statistical data analysis based on just one individual RO profile, are estimated and corrected for when known, as recommended by the GUM. The remaining residual biases are assumed to stay within a (conservative) bound estimate, which we refer to as *estimated systematic uncertainty* and by which we aim to provide at least 90 % likelihood coverage (confidence) that residual biases stay within the plus/minus envelope range of this uncertainty.

Depending on their nature, components of the systematic uncertainty that we need to estimate can be fundamentally systematic across different RO events, a subtype we term *estimated basic systematic uncertainties*, or appear systematic just for individual RO events, a second subtype that we term *estimated apparent systematic uncertainties*. It is important to distinguish these two subtypes, since the apparent systematic uncertainties will essentially behave as random uncertainties in ensemble averaging over many RO events, such as when generating climatologies, while the basic systematic ones will not average out and therefore fundamentally limit the (absolute) accuracy of ensemble averages such as climatologies.

Since the noise-type effects giving rise to short-range-correlated random uncertainties can be considered uncorrelated to the bias-type effects inducing long-range-correlated apparent systematic uncertainties, and since both are uncorrelated to basic systematic uncertainties, it is insightful and possible with due care to estimate and propagate each of these uncertainties independently.

As for the L2a processor (SKS2017), the operators of the L1b processor (i.e., the boldfaced items 1.2, 1.4, 2.1, 2.7, 2.9, 3.1, and 3.5 in Figure 3.2) qualify as *explicit*,

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

multivariate, and *linear* measurement models, as defined in the GUM, with *correlated input* quantities. They can therefore be formulated as

$$Y = \mathbf{A}^{\text{XY}} \cdot X, \quad (3.1)$$

where the input quantity X and output quantity Y are rank 1 vectors (profiles) of random variables, which we call *state* profiles. According to the GUM, their *random uncertainties* can be propagated using

$$\begin{aligned} \mathbf{C}_Y &= \mathbf{E}[YY^T] = \mathbf{E}\left[\left(\mathbf{A}^{\text{XY}} X\right) \left(\mathbf{A}^{\text{XY}} X\right)^T\right] \\ &= \mathbf{A}^{\text{XY}} \mathbf{E}[XX^T] \left(\mathbf{A}^{\text{XY}}\right)^T = \mathbf{A}^{\text{XY}} \mathbf{C}_X \left(\mathbf{A}^{\text{XY}}\right)^T \end{aligned} \quad (3.2)$$

when the uncertainties are normally distributed. This assumption is reasonably justified, since the receiving system noise (i.e., thermal noise and residual clock estimation noise) and the ionospheric noise (from scintillations induced by ionospheric irregularities) are essentially normally distributed overall (Gorbunov 2002b; Kursinski et al. 1997; Sokolovskiy et al. 2009; Syndergaard 1999). These noise sources are the main contribution to the random uncertainties in the excess phase profiles feeding into the L1b processor.

\mathbf{C}_X and \mathbf{C}_Y are the covariance matrices of the input and output variables, respectively, and \mathbf{A}^{XY} is the linear (or linearized) operator connecting X and Y . Equation (3.3) formulates how the covariance matrix \mathbf{C}_X is calculated from random uncertainty estimates u_X^r and the correlation matrix \mathbf{R}_X :

$$C_{X,ij} = u_{X,i}^r \cdot u_{X,j}^r \cdot R_{X,ij}. \quad (3.3)$$

As a key variable characterizing \mathbf{R}_X , *correlation length* profiles l_X are estimated from the correlation functions assembled in \mathbf{R}_X . The algorithm used estimates l_X by searching for the distances downward and upward of the correlation functions' main peak at which the correlation function has dropped to a value of $1/e$ (≈ 0.378). The adopted correlation length estimate is the arithmetic mean of these two upward and downward estimates (as the peak may be somewhat asymmetric). Additionally the correlation length is constrained by the data domain; i.e., the correlation length can never be larger than the profiles' vertical range.

Since the covariance propagation of random uncertainties requires extensive matrix multiplications for each measurement model along the entire retrieval chain, we also tested simpler variance propagation (VP), for which correlations are ignored; Appendix A.2 summarizes the relevant algorithms. However, as shown in Sect. 3.5, variance propagation unduly overestimates random uncertainties, so that covariance propagation is required.

When the operator is linear, as is the case for the applicable L1b operators, estimated *systematic uncertainties* can be propagated by application of the state retrieval operator

on the estimated systematic input uncertainty:

$$\begin{aligned} u_Y^s &= \mathbf{A}^{XY} \cdot (X + u_X^s) - Y = \mathbf{A}^{XY} \cdot X + \mathbf{A}^{XY} \cdot u_X^s - Y \\ &= Y + \mathbf{A}^{XY} \cdot u_X^s - Y = \mathbf{A}^{XY} \cdot u_X^s, \end{aligned} \quad (3.4)$$

where u_X^s and u_Y^s are the rank 1 systematic uncertainty profiles of the input and output variables.

In addition to random uncertainties, systematic uncertainties, and the correlation length, we also estimate *resolution* profiles w_X as context information along with the provided random uncertainty profiles (necessary, e.g., because smoothing can decrease random uncertainties, while making resolution coarser). This is enabled by careful selection and formulation of low-pass filter operations, in particular explicit filter cutoff frequency specification as the main driver of the resolution remaining after low-pass filtering.

We note that the (half-)Fresnel scale *physical* resolution often ascribed to RO bending angle profiles retrieved by geometric-optics methods (e.g., Gorbunov et al. 2004; Kursinski et al. 1997) will generally be somewhat coarser than the *filter-limited* resolution estimated here. This is intentional to maximize available information in the bending angle profiles provided by the L1b processor. In the rOPS, on input to the L2a processor and before high-altitude initialization by statistical optimization, the resolution of all profiles is brought to a common altitude-dependent resolution, which reflects the half-Fresnel scale (SKS2017).

3.3.2 Data sources and preparation

The input variables needed for the L1b uncertainty propagation, visible in Figure 3.2 and defined in Table 3.1, are the retrieved excess phase profiles $L_{r,k}(t)$ and the associated systematic uncertainty profiles $u_{L_{r,k}}^s(t)$, random uncertainty profiles $u_{L_{r,k}}^r(t)$, and correlation matrices $\mathbf{R}_{L_{r,k}}$, as well as the orbit positions and velocities of receiver and transmitter satellite – $\vec{r}_R(t)$, $\vec{v}_R(t)$, $\vec{r}_T(t)$, and $\vec{v}_T(t)$ – and their (systematic) uncertainties, $u_{\vec{r}_R}^s(t)$, $u_{\vec{v}_R}^s(t)$, $u_{\vec{r}_T}^s(t)$, and $u_{\vec{v}_T}^s(t)$. For due limitation of depth of workflow detail in Figure 3.2 we do not separately show the propagation of the basic and apparent systematic uncertainties as they are both identically propagated through the operator chain shown for $u_{L_{r,k}}^s(t)$. All variables are provided on the time grid t with elements t_i , at $f_s = 50$ Hz sampling rate, and for the two GPS carrier frequencies: f_{T_k} , with $k \in \{1, 2\}$, $f_{T1} = 1.57542$ GHz, and $f_{T2} = 1.22760$ GHz.

We used excess phase state profiles $L_{r,k}(t)$ and the orbit state profiles $\vec{r}_R(t)$, $\vec{v}_R(t)$, $\vec{r}_T(t)$, and $\vec{v}_T(t)$ from 15 July 2008 as a test day ensemble. For CHAMP, COSMIC, and MetOp, orbit state and excess phase profiles were provided by the COSMIC Data Analysis and Archiving Center (CDAAC) of the University Corporation for Atmospheric Research (UCAR), Boulder, Colorado. The End-to-End GNSS Occultation Performance

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

Simulation and Processing System (EGOPS) (Fritzer et al. 2009) was used for generating the simulated MetOp orbit state and excess phase profiles with realistic receiver noise (simMetOp). Figure 3.3 shows $L_{r,k}(t)$ in panel (a), $u_{L_{r,k}}^s(t)$ in panel (b), $u_{L_{r,k}}^r(t)$ in panel (c), and $\mathbf{R}_{L_{r,k}}$ in terms of representative correlation functions in panels (d) and (e), for a typical COSMIC RO event of the test day ensemble from 15 July 2008 (example case).

Exploiting the linearity of the (linearized) retrieval operators, the so-called *baseband approach* (Kirchengast et al. 2016a) is applied throughout the rOPS. Hereby a zero-order model profile is subtracted from the input state profile, and only the remaining delta profile is processed through the operator. After application of the operator, the zero-order model profile of the output state profile is added back to the resulting delta profile. This approach effectively avoids biases from numerical operations on (near-)exponentially varying RO profiles, since the model profiles that we derive from short-range (24 h) forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) skillfully subtract the (near-)exponential variation. The remaining increment profiles that we need to treat numerically then appear to be very linear and with low dynamical range, which leads to very low residual numerical errors of operators such as filters and derivatives.

The model profiles used as zero-order states in the retrieval – i.e., L_m , D_m , and α_m (cf. Table 3.1) – were created from ECMWF short-range (24 h) forecast refractivity fields, accurately forward-modeled to bending angle (α_m), Doppler shift (D_m), and excess phase (L_m) profiles, collocated to the latitude, longitude, and time of the respective RO event processed in the rOPS. The ECMWF fields used have a horizontal resolution of about 300 km (triangular truncation T42) – which corresponds to the approximate horizontal resolution of RO profiles (e.g., Kursinski et al. 1997) – and are available at 91 vertical levels (L91).

ECMWF fields were chosen for their proven leading quality (Bauer et al. 2015; Untch et al. 2006) and thus high suitability for serving as zero-order state profiles; any other reasonable model profiles could be used as well since the retrieval results negligibly depend on the exactly chosen zero-order model profiles. For comparison we plotted $L_m(t)$ for the COSMIC example case into Figure 3.3a, which demonstrates that the ECMWF short-range forecast lies very close to $L_{r1}(t)$ and $L_{r2}(t)$ and thus is well suited as the model profile.

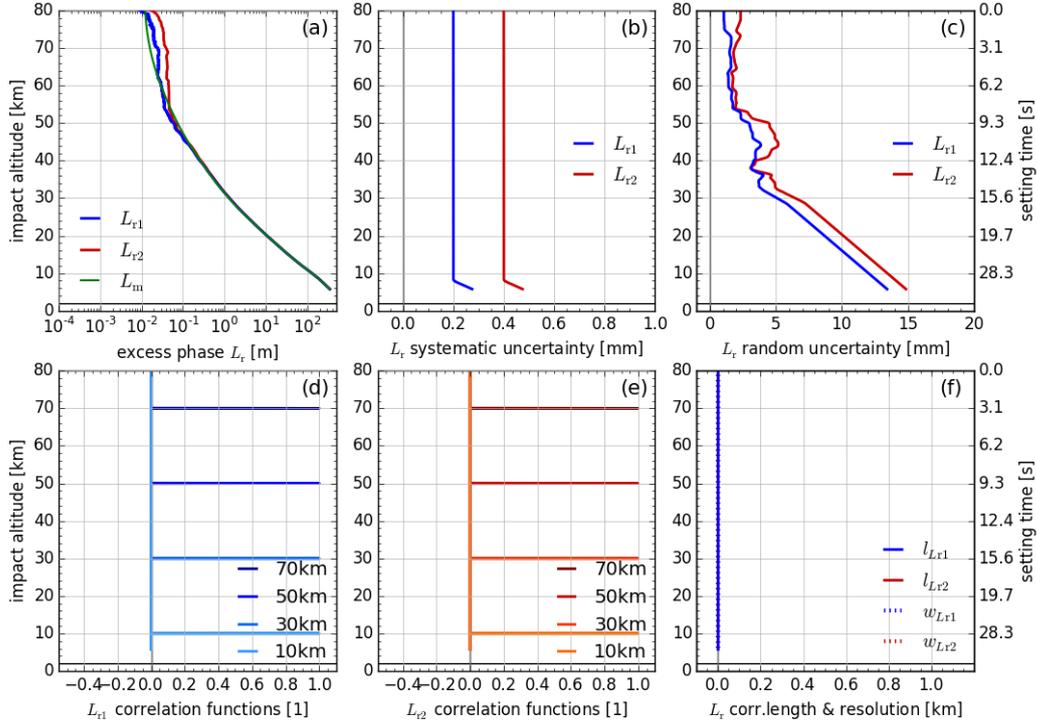


Figure 3.3: Input profiles of retrieved excess phase L_r (with model profile L_m for comparison) in panel (a), estimated systematic uncertainty profiles $u_{L_r}^s$ in panel (b), estimated random uncertainty profiles $u_{L_r}^r$ in panel (c), representative correlation functions $R_{L_r,i}$ (at 10, 30, 50, and 70 km) in panels (d) and (e), and correlation length l_{L_r} (solid) and resolution profiles w_{L_r} (dotted) in panel (f), which are set zero for these initial essentially uncorrelated input data. All profiles are shown for both GPS carrier frequencies: f_{T1} (blue) and f_{T2} (red).

While in the future the excess phase random and systematic uncertainty profiles will be more rigorously estimated by the rOPS L1a processor (Innerkofler et al. 2017b) and provided as input to the L1b processor, they had to be estimated for this study from existing excess phase profiles with realistic noise and simplified modeling. To this end, each *estimated random uncertainty* profile $u_{L_r,k}^r(t)$ was estimated based on the noise of the respective retrieved excess phase profile $L_{r,k}(t)$. The noise was determined following the estimation scheme for bending angle observation errors described by Li et al. (2015, Sect. 2.2 therein), so we just briefly summarize how we used it here.

First, for both, the retrieved profile $L_{r,k}$ and for the model profile L_m , the mean over all grid points between 60 and 70 km was determined. Then L_m was offset-corrected towards $L_{r,k}$ by subtracting the difference of these two means from L_m , giving the offset-corrected model profile $L_{\tilde{m}}$. Next, the delta profile $\delta L_{r\tilde{m},k} = L_{r,k} - L_{\tilde{m}}$ was calculated. After

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

smoothing $\delta L_{\text{r}\bar{m},k}$ with a 10 km moving-average boxcar (BC) filter, the smoothed profile was subtracted from $\delta L_{\text{r}\bar{m},k}$ again, to get $\delta\delta L_{\text{r}\bar{m},k}$, the random noise profile component of $L_{\text{r},k}$ isolated in this way. Finally, the estimated random uncertainty was determined as

$$u_{L_{\text{r}},ik}^r = \sqrt{\sum_{j=i-M/2}^{i+M/2} \delta\delta L_{\text{r}\bar{m},jk}^2}, \quad (3.5)$$

where M is the number of grid points equivalent to a window width of 10 km. To avoid boundary effects of the filter, $u_{L_{\text{r}},k}^r$ was only determined up to $z_{\text{aTop}} - 5$ km and down to z_{aGradr} at 30 km. It was constantly extended at the upper end and extended by a linear gradient below z_{aGradr} , using (in units [m])

$$u_{L_{\text{r}},ik}^r = u_{L_{\text{r}},k}^r(z_{\text{aGradr}}) + \frac{z_{\text{aGradr}} - z_{\text{a},i}}{3 \times 10^6}, \quad (3.6)$$

for all elements of $u_{L_{\text{r}},k}^r(t)$ below z_{aGradr} , roughly following estimates of ESA/EUMETSAT (1998) and the overall behavior of estimates from real excess phase profiles (the latter became too vulnerable to biases and fluctuations to continue using them below 30 km).

Since the noise components responsible for the random uncertainty at excess phase level are essentially uncorrelated at a sampling rate of 50 Hz (Hajj et al. 2002; Syndergaard 1999), the correlation matrix $\mathbf{R}_{L_{\text{r}}}$ is set to unity in the diagonal and to zero outside (i.e., a Kronecker- δ assignment) for both channels (Figure 3d–f). In case the future excess phase data from the rOPS L1a processor exhibit non-negligible correlations for some data from some of the RO missions, we will account for these correlations in $\mathbf{R}_{L_{\text{r}}}$, since our L1b algorithm (Sect. 3.4) is prepared for full covariance propagation. The elements of the covariance matrix $\mathbf{C}_{L_{\text{r}}}$ are hence (item 1.1 in Figure 3.2)

$$C_{L_{\text{r}},ijk} = u_{L_{\text{r}},ik}^r \cdot u_{L_{\text{r}},jk}^r \cdot R_{L_{\text{r}},ijk} = u_{L_{\text{r}},ik}^r \cdot u_{L_{\text{r}},jk}^r \cdot \delta_{ij}. \quad (3.7)$$

For the MC validation of the CP, error profile realizations $\epsilon_{L_{\text{r}}}^r$ were superimposed onto simulated “true” excess phase profiles $L_{\text{r},k}^T(t)$. As a source for $L_{\text{r},k}^T$, we used an EGOPS-simulated “error-free” CHAMP event from 8 August 2008 (i.e., no receiver system errors superimposed). Using an “error-free” profile as the basis, with the particular simulated profile just serving as a representative RO profile to illustrate the MC validation, allows us to strictly ensure the consistency of the random uncertainty of the input profile with the ensemble of superimposed error profile realizations.

To create the error profiles, a representative $u_{L_{\text{r}}}^{r,\text{Std}}$ uncertainty profile was selected from a COSMIC ensemble of uncertainty profiles, created according to Equations (3.5) and (3.6). The error profile realizations are random draws from a distribution characterized by these uncertainties, again assuming that $R_{L_{\text{r}},ij} = \delta_{ij}$, i.e., that there are no correlations between the individual grid levels (item (a) in Figure 3.2; Figure 3.3f). The same standard

profile $u_{Lr}^{r,Std}$ was used as input for the CP to which the MC results are then compared. This MC validation method applied to test the rOPS L1b uncertainty propagation steps is essentially the same as in SKS2017, and it is described therein in more detail.

The *estimated systematic uncertainty* $u_{Lr,k}^s$ was determined based on a simple model roughly following error estimates from ESA/EUMETSAT (1998), with constant uncertainty from 80 km down to z_{aGrads} at 8 km and a linear uncertainty gradient in the troposphere; as noted above, this simplified modeling will be replaced in the future by realistic uncertainty estimates received as L1b retrieval input from the L1a processor (Innerkofler et al. 2017b).

The constant $u_{Lr,k}^s$ above z_{aGrads} is 0.1 mm for $k = 1$ and 0.2 mm for $k = 2$ for MetOp and simMetOp. This uncertainty is interpreted as an estimated basic systematic uncertainty, i.e., as a lower-bound estimate of available accuracy.

For CHAMP and COSMIC we set $u_{Lr,1}^s = 0.2$ mm and $u_{Lr,2}^s = 0.4$ mm, to roughly reflect the fact that these RO receivers are lower-cost instruments with lower gain, and thus somewhat lower tracking performance, than the RO receiver on MetOp (e.g., Angerer et al. 2017; Luntama et al. 2008). From z_{aGrads} downwards, $u_{Lr,k}^s$ (in units [m]) increases by

$$u_{Lr,ik}^s = u_{Lr,k}^s(z_{aGrads}) + \frac{z_{aGrads} - z_{a,i}}{3 \times 10^7}. \quad (3.8)$$

In order to avoid a sharp kink in the $u_{Lr,k}^r$ profiles at z_{aGrads} , and in the $u_{Lr,k}^s$ profiles at z_{aGrads} , a 2 km width moving-average boxcar filter was applied to smooth these simple uncertainty models around these transition altitudes (for the example profile u_{Lr}^s is visible in Figure 3.3b).

The orbit position and velocity uncertainties of the transmitter and the receiver satellites show little variation within the short duration of an individual RO event of about 45 s to 2 min (Innerkofler et al. 2017b) and can be assumed to be constant biases. They are thus counted to the systematic uncertainties, more precisely the apparent systematic uncertainties, since the actual values of the orbit-borne biases will generally change in a pseudo-random manner from event to event.

We set the transmitter position and velocity uncertainties to $u_{\vec{r}T}^s = 3$ cm and $u_{\vec{v}T}^s = 0.01$ mm s⁻¹, consistent with accuracies for GPS orbits available from GNSS orbit providers like the International GNSS Service (IGS). The receiver position and velocity uncertainties, $u_{\vec{r}R}^s = 5$ cm and $u_{\vec{v}R}^s = 0.05$ mm s⁻¹ for CHAMP and MetOp, are adopted 4 times smaller than those for COSMIC with $u_{\vec{r}R}^s = 20$ cm and $u_{\vec{v}R}^s = 0.2$ mm s⁻¹, as found by ongoing rOPS-related POD studies (Innerkofler et al. 2017b), consistent with previous literature (e.g., Montenbruck et al. 2009; Schreiner et al. 2010).

3.4 Algorithm sequence and example results

In this section the L1b uncertainty propagation algorithm sequence is introduced. We illustrate the effects of the algorithm on the main uncertainty variables by way of the COSMIC example case already used for Figure 3.3.

For each L1b retrieval step – i.e., segments (1), (2), and (3) in Figure 3.2 – the results for the principal variables are shown in Figs. 3.4 to 3.8. These variables are the state profiles X_r (with $X_r \in \{L_{F,k}(t), D_{r,k}(t), \alpha_{G,k}(t), \alpha_{G,k}(z_a), \alpha_{M,k}(z_a), \alpha_{F,k}(z_a), \text{ and } \alpha_r(z_a)\}$), the estimated systematic uncertainty profiles $u_{X_r}^s$, the estimated random uncertainty profiles $u_{X_r}^r$, representative correlation functions $R_{X_r,i}$ (with i such that $z_{a,i} \in \{10, 30, 50, 70 \text{ km}\}$), and the correlation length profiles l_{X_r} and resolution profiles w_{X_r} . Along with the dual-frequency state profiles, we also show the collocated forward-modeled short-range forecast profiles, i.e., model profiles X_m with $X_m \in \{L_m, D_m, \alpha_m\}$ for comparison.

A concise definition of the variables involved is provided in Table 3.1, as introduced above. The summary description in this section is complemented by a complete step-by-step description of the algorithm along the entire L1b retrieval chain in Appendix A.1, which is organized for convenience into the same sequence of subsections.

To simplify the notation in the description, we suppress index k whenever steps are applied in an identical way to the data of both GNSS L-band channels with frequencies f_{T1} and f_{T2} . Only if the two channels are treated differently, such as in Sect. 3.4.3, is the index considered again. For conciseness we also do not illustrate both the estimated basic and estimated apparent systematic uncertainty but rather the total estimated systematic uncertainty as the overall result.

3.4.1 Doppler shift retrieval

Basic low-pass filtering

A Blackman windowed sinc (BWS) low-pass filter with a filter cutoff frequency $f_c = 2.5 \text{ Hz}$ (boxcar-equivalent filter width of 0.2 s) (item 1.2 in Figure 3.2) is applied onto the excess phase profile $L_r(t)$, before the Doppler differentiation (item 1.4 in Figure 3.2), to avoid an amplification of high-frequency noise in the phase profile by the derivative operation. This filter suppresses the noise; consequentially the filtered excess phase profile $L_F(t)$, shown in Figure 3.4a, is expected to have random uncertainties $u_{L_F}^r$ of smaller magnitude, but correlated over the length of the filter window. The uncertainties obtained through the implemented algorithm confirm these expectations, i.e., random uncertainty profiles in Figure 3.4c are less than a third in magnitude of those in Figure 3.3c, and Figure 3.4d–e show how the correlation functions widened and the correlation length and vertical resolution reached ~ 0.5 and $\sim 0.6 \text{ km}$, respectively, above about 30 km impact altitude (Figure 3.4f).

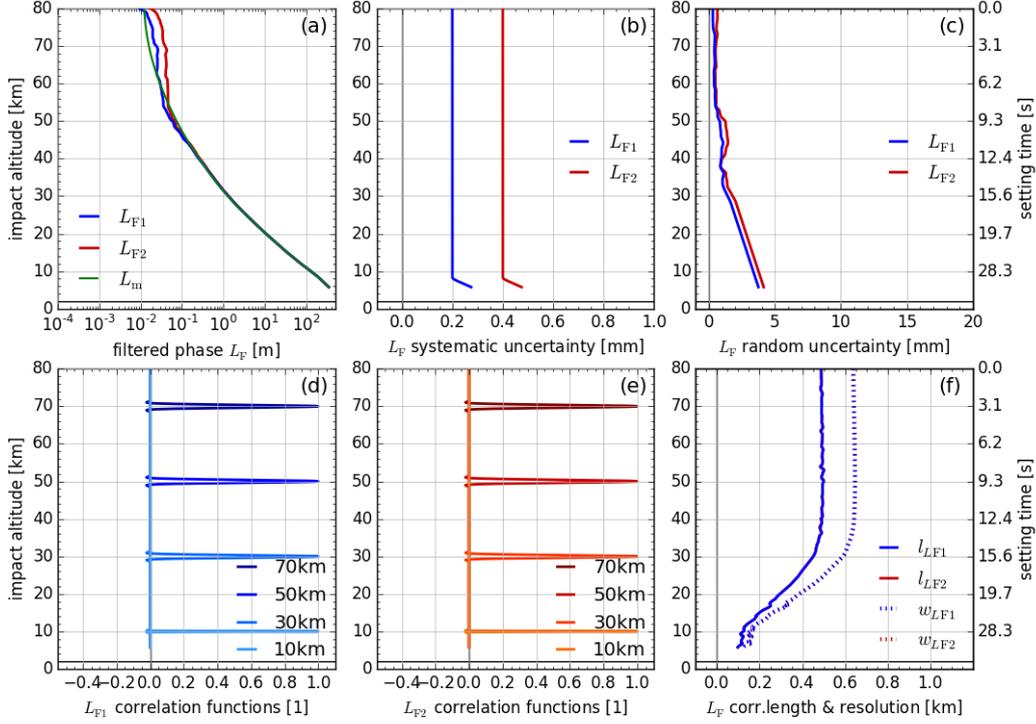


Figure 3.4: Results for filtered excess phase profiles L_F (with model profile L_m for comparison) in panel (a), estimated systematic uncertainty profiles $u_{L_F}^s$ in panel (b), estimated random uncertainty profiles $u_{L_F}^r$ in panel (c), representative correlation functions $R_{L_F,i}$ (at 10, 30, 50, and 70 km) in panels (d) and (e), and correlation length l_{L_F} (solid) and resolution profiles w_{L_F} (dotted) in panel (f). All profiles are shown for both GPS carrier frequencies: f_{T1} (blue) and f_{T2} (red).

The random uncertainty propagation algorithm, i.e., the covariance propagation from \mathbf{C}_{L_r} to \mathbf{C}_{L_F} , is described by Equation (A.6) and item 1.3 in Figure 3.2 and is justified by Equation (3.2). To obtain $u_{L_F}^r$ and \mathbf{R}_{L_F} , we use Equations (A.7) and (A.8).

To propagate the estimated systematic uncertainty $u_{L_r}^s$, which characterizes long-range-correlated offsets or biases, we use the same BWS filter as for the state profile, i.e., making use of Equation (3.4). Because the input uncertainty profile $u_{L_r}^s$ is chosen to be constant down to z_{aGrads} , the filter has little effect, and $u_{L_F}^s$, shown in Figure 3.4b, is essentially equal to $u_{L_r}^s$, shown in Figure 3.3b.

The resolution profile w_{L_r} is determined by the filter width according to Equations (A.11) and (A.13). After the BWS filtering, the resolution is roughly equal to the correlation length l_{L_F} , amounting to ~ 0.6 km above about 30 km impact altitude and becoming finer downwards due to the increasing refraction (Figure 3.4f).

Doppler shift derivation

The next step is a five-point differentiation operation (item 1.4 in Figure 3.2) used to calculate the Doppler shift profile $D_r(t)$ from the filtered excess phase profile $L_F(t)$. The resulting dual-frequency Doppler shift profiles are plotted along with the model profile $D_m(t)$ in Figure 3.5a for the example case.

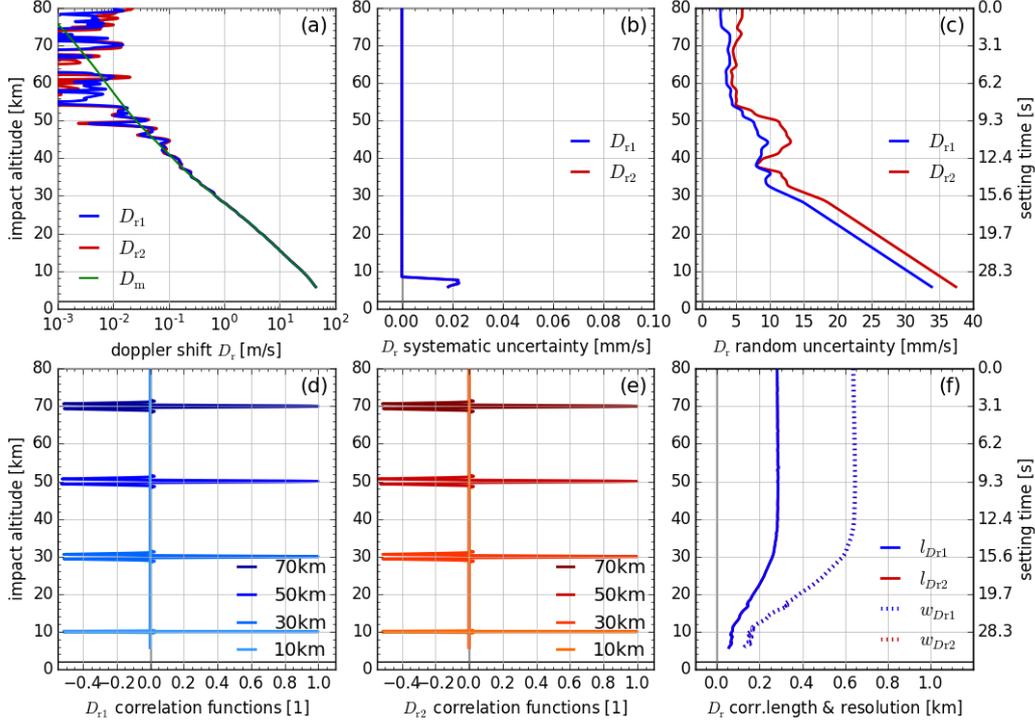


Figure 3.5: Results for retrieved Doppler shift profiles D_r (with model profile D_m for comparison) in panel (a), estimated systematic uncertainty profiles $u_{D_r}^s$ in panel (b), estimated random uncertainty profiles $u_{D_r}^r$ in panel (c), representative correlation functions $R_{D_r,i}$ (at 10, 30, 50, and 70 km) in panels (d) and (e), and correlation length l_{D_r} (solid) and resolution profiles w_{D_r} (dotted, estimated for main peak) in panel (f). All profiles are shown for both GPS carrier frequencies: f_{T1} (blue) and f_{T2} (red).

As for the filtered excess phase, we apply CP (Equation A.18, item 1.5 in Figure 3.2) to first calculate the covariance matrix \mathbf{C}_{D_r} and then extract $u_{D_r}^r$ (shown in Figure 3.5c) and \mathbf{R}_{D_r} (Figure 3.5d and 3.5e) from it. The choice of the x-axis range shows the random uncertainties increased, but the differentiation actually does increase relative random uncertainties (relative to the state profile). It also causes anti-correlation with neighboring elements, as visualized by the negative side peaks of the correlation functions

in Figure 3.5d and 3.5e. The correlation length l_{D_r} (of the main correlation function peak) decreases accordingly (now smaller than 0.3 km throughout), because the correlation functions fall off more steeply on both sides of the main peak (Figure 3.5f).

For calculating the estimated systematic uncertainty, we use the state operator; i.e., we just differentiate u_{LF}^s and get $u_{D_r}^s$ (shown in Figure 3.5b). With the current illustrative choice of input uncertainties the systematic uncertainty of the Doppler shift profile is zero above the transition to the troposphere, where the estimated systematic uncertainty of the excess phase is assumed constant; in the troposphere a Doppler shift offset of $\sim 0.02 \text{ mm s}^{-1}$ occurs.

The resolution profile w_{D_r} shows that the vertical resolution stays unaffected by this operator (cf. Figs. 3.5f and 3.4f), because the BWS filter width of the preceding low-pass filtering (intentionally) stretched beyond the five neighboring points involved in the differentiation.

3.4.2 Bending angle retrieval

GO bending angle retrieval

The next operator is the GO bending angle retrieval in which retrieved GO bending angle profiles $\alpha_G(t)$ are calculated from Doppler shift profiles $D_r(t)$ and the orbit position and velocity vectors $\vec{r}_T(t)$, $\vec{r}_R(t)$, $\vec{v}_T(t)$, and $\vec{v}_R(t)$ (item 2.1 in Figure 3.2) and then interpolated to the (common monotonic) impact altitude grid z_a (item 2.6 in Figure 3.2).

Figure 3.6a shows retrieved α_G profiles together with the model profile α_m . The mildly nonlinear implicit-type bending angle retrieval operator needs to be solved iteratively, and it requires linearization for both random and systematic uncertainty propagation, as described in detail in Appendix A (Sect. A.1.2). Because this retrieval step is performed level by level, keeping levels independent, the GO bending angle retrieval leaves correlation functions and resolution unchanged (cf. Figs. 3.6d–f and 3.5d–f).

The estimated random uncertainties $u_{\alpha_G}^r$, as shown in Figure 3.6c, now increase more strongly in the lower stratosphere and troposphere (to about 40 to 50 μrad near 10 km), because they are dependent on the vertical gradient of the impact parameter a_t , which is increasingly larger towards lower altitudes from the increasing refraction.

The main contributions to the estimated systematic uncertainty $u_{\alpha_G}^s$ are induced by systematic uncertainties in orbit velocity and position of the transmitter and the receiver satellite (details in Sect. A.1.2), which in total amount to about 0.05 μrad (Figure 3.6b). Compared to this magnitude, the systematic uncertainty contributed by the Doppler shift uncertainty is very small.

WO bending angle retrieval

Due to strong refractivity gradients and multipath effects, the GO bending angle retrieval can be inadequate in the troposphere, and therefore WO algorithms are applied to

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

reconstruct the geometric optical ray structure of the wave field (e.g., Gorbunov 2002a; Gorbunov and Lauritsen 2004).

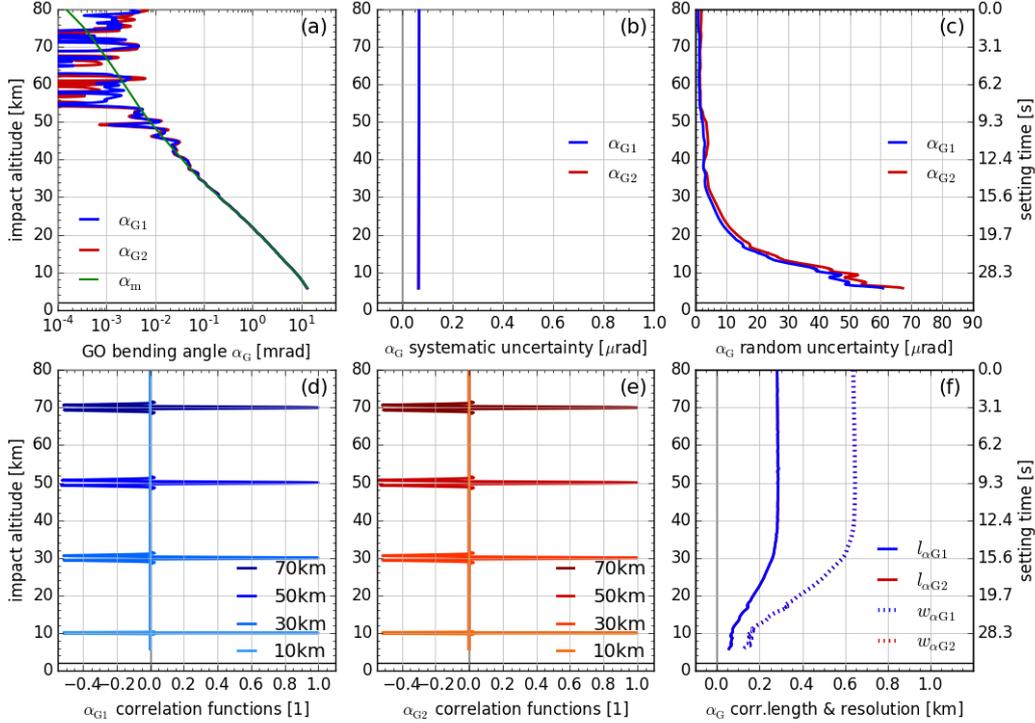


Figure 3.6: Results for geometric optics bending angle profiles α_G (with model profile α_m for comparison) in panel (a), estimated systematic uncertainty profiles $u_{\alpha_G}^s$ in panel (b), estimated random uncertainty profiles $u_{\alpha_G}^r$ in panel (c), representative correlation functions $R_{\alpha_G,i}$ (at 10, 30, 50, and 70 km) in panels (d) and (e), and correlation length l_{α_G} (solid) and resolution profile w_{α_G} (dotted, estimated for main peak) in panel (f). All profiles are shown for both GPS carrier frequencies: f_{T1} (blue) and f_{T2} (red); in panels (b) and (f) both profiles are essentially identical (so that blue shadows the red color).

In the rOPS, along with the WO bending angle profile $\alpha_W(z_a)$, the systematic uncertainty profile $u_{\alpha_W}^s$, the random uncertainty profile $u_{\alpha_W}^r$, the correlation matrix \mathbf{R}_{α_W} , and the resolution profile w_{α_W} are retrieved (item 2.7 in Figure 3.2).

The WO bending angle retrieval algorithm used is a canonical transform (CT2) algorithm (Gorbunov et al. 2004), and the associated uncertainty propagation algorithm is not described here, but separately by Gorbunov and Kirchengast (2015) and Gorbunov and Kirchengast (2018). The WO retrieval and uncertainty propagation results are supplied up to 20 km impact altitude by the WO algorithms.

Merging of GO and WO bending angle profiles

In the rOPS bending angle retrieval the results from the WO retrieval, α_W , are merged with GO retrieval results, α_G , at around a transition altitude z_a^{GW} in a transition range $z_a^{GW} \pm \Delta z_a^{GW}$, to get merged profiles α_M (item 2.9 in Figure 3.2). The determination of the transition altitude and the merging algorithm are described in Appendix A.1.2. We use a specialized covariance propagation to propagate the GO and WO uncertainties, expressed by the covariance matrices $\mathbf{C}_{\alpha G}$ and $\mathbf{C}_{\alpha W}$, to properly obtain the covariance matrix of the merged bending angle $\mathbf{C}_{\alpha M}$ (Equations A.37 and A.38, item 2.10 in Figure 3.2).

Because the rOPS implementation of the WO uncertainty propagation (Gorbunov and Kirchengast 2018) was still in test phase and not yet available for integration into the simulations here, all examples in this study are GO-only; i.e., only the GO retrieval is performed. Results for α_M are thus unchanged from those shown in Figure 3.6 and not separately illustrated.

In order to nevertheless test and validate the uncertainty propagation of the merging algorithm, WO retrieval results were artificially substituted by the GO results for the MC validation (Sect. 3.5); i.e., GO was used as a proxy for WO since reasonably capturing expected WO variability as indicated by tests of Gorbunov and Kirchengast (2018).

3.4.3 Atmospheric bending angle derivation

Adaptive low-pass filtering and minor-channel extrapolation

To prepare the merged bending angle profiles $\alpha_{M,k}$ for the ionospheric correction, they are first filtered by another BWS filter operation (item 3.1 in Figure 3.2) in order to ensure adequately smoothed bending angle profiles $\alpha_{F,k}$, with $k \in \{1, 2\}$.

The chosen filter cutoff frequency for $k = 1$ is $f_{c1} = 2.5$ Hz, same as the basic filtering (Sect. 3.4.1), just to ensure clearness of any higher-frequency effects from operators after the initial excess phase filtering (e.g., from Doppler shift derivation that induces short-range anti-correlation effects). For $k = 2$, the cutoff frequency f_{c2} is set noise-dependent, between 2.5 and 0.5 Hz (boxcar-equivalent width of 0.2 to 1.0 s). In events in which the α_{F2} profile does not reach down as far as α_{F1} , it is extrapolated down to the bottom of α_{F1} , z_{aBot} . The results for the filtered bending angle state profiles $\alpha_{F,k}$ are displayed in Figure 3.7a, together with the associated model bending angle profile α_m . The filter has considerably reduced the noise of the profile, particularly for α_{F2} , where the algorithm selected a cutoff frequency $f_{c2} = 10/7$ Hz in this example case.

The relevant covariance-propagated random uncertainties $u_{\alpha F,k}^r$ are shown in Figure 3.7c (blue and red), illustrating the reduced noise, especially for α_{F2} . In return, the peaks of the correlation functions broaden (cf. Figs. 3.7d–e and 3.6d–e), with correlation lengths $l_{\alpha F,k}$ at near 0.4 km for α_{F1} and above 0.5 km for α_{F2} (Figure 3.7f).

The estimated systematic uncertainty remains largely unchanged (Figure 3.7b) due to its smooth character.

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

The resolution of the filtered bending angle profiles (according to Equations A.11 and A.13) is determined by the cutoff frequencies $f_{c,k}$ of the BWS filters. In the example case it is therefore essentially unchanged for α_{F1} , while it is significantly decreased for α_{F2} (cf. Figs. 3.7f and 3.6f) since $f_{c2} = 10/7$ Hz. That is, the resolution $w_{\alpha F2}$ in the upper stratosphere for example, where the vertical scanning velocity of this RO event is about 3.2 km s^{-1} , is near 1.1 km (Figure 3.7f).

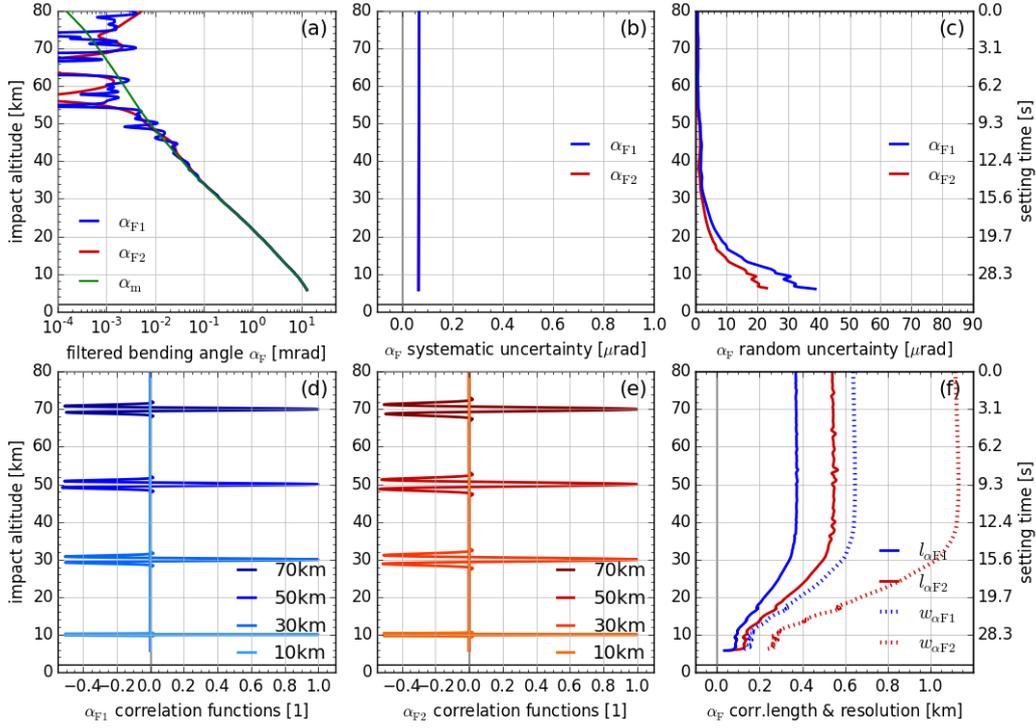


Figure 3.7: Results for filtered bending angle profiles α_F (with model profile α_m for comparison) in panel (a), systematic uncertainty profiles $u_{\alpha F}^s$ in panel (b), random uncertainty profiles $u_{\alpha F}^r$ in panel (c), representative correlation functions $R_{\alpha F,i}$ (at 10, 30, 50, and 70 km) in panels (d) and (e), and correlation length $l_{\alpha F}$ (solid) and resolution profiles $w_{\alpha F}$ (dotted, estimated for main peak) in panel (f). All profiles are shown for both GPS carrier frequencies: f_{T1} (blue) and f_{T2} (red).

Ionospheric correction

The final step of the L1b processor is the ionospheric correction (item 3.5 in Figure 3.2). The atmospheric bending angle α_r is obtained by applying a linear dual-frequency combination of α_{F1} and α_{F2} , such that ionospheric effects are largely removed (details are described in Sect. A.1.3). The final retrieved atmospheric bending angle α_r of the

example case is shown in Figure 3.8a. The propagation results for the estimated random uncertainty are shown in Figure 3.8c. The linear combination of the ionospheric correction amplifies noise and $u_{\alpha r}^r$ is therefore considerably larger than $u_{\alpha F1}^r$, and $u_{\alpha F2}^r$ (cf. Figs. 3.8c and 3.7c).

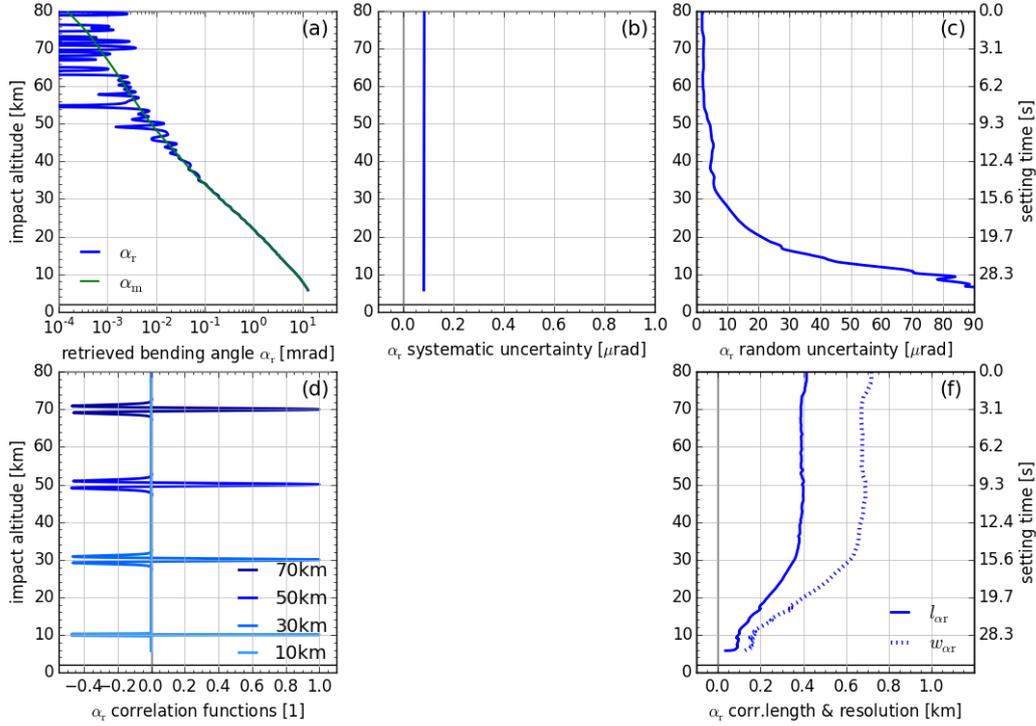


Figure 3.8: Results for atmospheric bending angle profile α_r (with model profile α_m for comparison) in panel (a), systematic uncertainty profile $u_{\alpha r}^s$ in panel (b), random uncertainty profile $u_{\alpha r}^r$ in panel (c), representative correlation functions $R_{\alpha r, i}$ (at 10, 30, 50, and 70 km) in panel (d), and correlation length $l_{\alpha r}$ (solid) and resolution profile $w_{\alpha r}$ (dotted, estimated for main peak) in panel (e).

Figure 3.8d shows how the correlation functions – as obtained through covariance propagation – combine the characteristics of the correlation functions from the two matrices $\mathbf{R}_{\alpha F1}$ and $\mathbf{R}_{\alpha F2}$, while essentially inheriting the α_{F1} behavior, since the α_{F2} influence on the ionospheric correction is comparatively minor (see Sect. A.1.3).

The residual higher-order ionospheric effects are accounted for by a “conservative best-guess” value ($0.05 \mu\text{rad}$, reflecting results of (Liu et al. 2015), and (Danzer et al. 2013; Danzer et al. 2015)) and added (in root-mean-square form) to the systematic uncertainty profile $u_{\alpha r}^s$, leading to a total estimated systematic uncertainty in this

example case of $\sim 0.07 \mu\text{rad}$ (Figure 3.8b). Within this uncertainty, the one dominating component from orbit uncertainties ($\sim 0.05 \mu\text{rad}$; cf. Figure 3.6b) can be considered an apparent systematic uncertainty that will essentially average out in ensemble averaging (e.g., climatologies), while the other dominating component from residual higher-order ionospheric biases (also estimated $\sim 0.05 \mu\text{rad}$ as noted above) can be considered a basic systematic uncertainty. For the latter it is therefore useful and prepared for in the rOPS – in line with GUM recommendations and as discussed in the introductory Sect. 1 – to correct for the quantifiable part of it in the future so that the total basic systematic uncertainty may be mitigated down to the $\sim 0.01 \mu\text{rad}$ level.

The resolution profile w_{or} of the retrieved bending angle (Figure 3.8f) is dominated by the contribution of α_{F1} that strongly dominates (intentionally by construction) the ionospheric correction results in terms of the small-scale bending angle variability. Similar to the correlation length profile l_{or} , it is therefore very close to $w_{\alpha_{\text{F1}}}$ and only slightly larger.

3.5 Algorithm validation

The GUM advises to use a MC method for uncertainty propagation if the retrieval operators do not fulfill the criteria for a GUM-type CP. In our case the MC method is put to another beneficial use, to validate the results of the CP, as recommended by JCGM (2011).

For the validation of the covariance propagation by the MC method, we sampled the input excess phase profile random error distribution, statistically described by

$$C_{L_r,ij}^{\text{MC}} = u_{L_r,i}^{r,\text{Std}} \cdot u_{L_r,j}^{r,\text{Std}} \cdot \delta_{ij}, \quad (3.9)$$

by a large number M of draws $L_r^{\text{T}} + \epsilon_{L_r,j}^r$ (with $j \in \{1, \dots, M\}$ and $M = 1000$). For each of these M profile realizations, the state retrieval is run through the L1b retrieval chain, to give M realizations of the output variable X_j (with $X_j \in \{L_{\text{F},kj}(t), D_{\text{r},kj}(t), \alpha_{\text{G},kj}(z_a), \alpha_{\text{F},kj}(z_a), \alpha_{\text{r},j}(z_a)\}$ and $k \in \{1, 2\}$). From these individual realizations the mean profiles X^{MC} and the covariance matrices \mathbf{C}_X^{MC} ,

$$\begin{aligned} \mathbf{C}_X^{\text{MC}} = & \frac{1}{M-1} \left[\left(X_1 - X^{\text{MC}} \right) \left(X_1 - X^{\text{MC}} \right)^{\text{T}} + \dots \right. \\ & \left. + \left(X_M - X^{\text{MC}} \right) \left(X_M - X^{\text{MC}} \right)^{\text{T}} \right], \end{aligned} \quad (3.10)$$

are calculated (items b–g in Figure 3.2). Using the same input profile and uncertainty information as used to specify the MC runs (described in Sect. 3.3.2), the retrieval is then also run with covariance-based uncertainty propagation, and the resulting CP-propagated covariance matrices \mathbf{C}_X^{CP} are compared to the MC-derived matrices \mathbf{C}_X^{MC} . In order to be

able to attribute potential changes between CP and MC covariance matrices better, we decompose \mathbf{C}_X into u_X^r and \mathbf{R}_X (Equations A.7 and A.8), and compare them separately.

Figure 3.9 shows the different steps along the retrieval chain from $L_{F,k}(t)$ to $D_{r,k}(t)$, $\alpha_{G,k}(z_a)$, $\alpha_{F,k}(z_a)$, and $\alpha_r(z_a)$ in the rows, for $k = 1$ (GPS f_{T1} frequency) in the left column and for $k = 2$ (GPS f_{T2} frequency) in the middle column. The right column shows multiple representative correlation functions, from near 10 to near 70 km. Due to the limited number of MC draws, the MC results (black lines) show some jitter both in the estimated random uncertainty and in the correlation functions. Since the purpose of the MC results is only to demonstrate the correctness of the CP result, we can disregard this behavior.

Figure 3.9a (light blue) and b (orange) show the random uncertainties $u_{Lr,1}^r$ and $u_{Lr,2}^r$, respectively, which characterize the input distribution and from which the random error profiles $\epsilon_{Lr,j}^r$ are drawn. They also show the CP results for the random uncertainty u_{LF1}^r (dark blue in Figure 3.9a) and u_{LF2}^r (red in Figure 3.9b), compared to the MC propagated random uncertainties (black).

The CP and MC lines match very well and show that the implemented CP algorithm delivers correct results for the basic filtering step. For f_{T2} , the MC uncertainties do not reach down as far as the CP uncertainties, because the shortest of all draws of the large ensemble of size M determines how far down the recombined MC covariance matrix (Equation 3.10) reaches. Figure 3.9c compares CP correlation functions $R_{LF,i1}$ (blue) and $R_{LF,i2}$ (red) to the corresponding MC correlation functions (black dashed).

The CP and MC correlation functions also agree well. Both capture the narrow peak, broadened by the BWS filter. Again the MC correlation functions fluctuate around zero left and right of the peak, from the finite ensemble size, but it is obvious that the CP delivers the correct off-peak results (i.e., zero; the off-peak elements outside the BWS filter window must nominally be zero). The MC validation (black) of u_{Dr1}^r (Figure 3.9d), u_{Dr2}^r (Figure 3.9e), and $\mathbf{R}_{Dr,i}$ (Figure 3.9f, blue and red) demonstrates that the CP through the Doppler shift derivation performs correctly as well.

The next row, Figure 3.9g to i, shows the results for the GO bending angle $\alpha_G(z_a)$, i.e., after the interpolation of all quantities to the (common monotonic) impact altitude grid z_a . For comparison, in Figure 3.9a to f, all quantities have been computed on the common time grid (“setting time” relative to time zero at 80 km altitude) with 50 Hz sampling rate; the corresponding impact altitude of the “true” profile L_r^T is shown for additional convenience on the right-hand-side (RHS) axis. In Figure 3.9g to o, these bending angle quantities have been computed on the impact altitude grid; in these cases therefore the corresponding setting time of the “true” profile is shown for additional convenience on the RHS axis.

The results for filtered α_F follow in Figure 3.9j to l. Also here the MC results match the CP result well. Due to the lower BWS cutoff frequency for α_{F2} , now $u_{\alpha F2}^r$ is smaller than $u_{\alpha F1}^r$, even though $u_{\alpha G2}^r$ was larger than $u_{\alpha G1}^r$. Correspondingly the peak of the correlation functions $R_{\alpha F,i2}$ widened more than those of $R_{\alpha F,i1}$ (cf. Figure 3.9l and i).

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

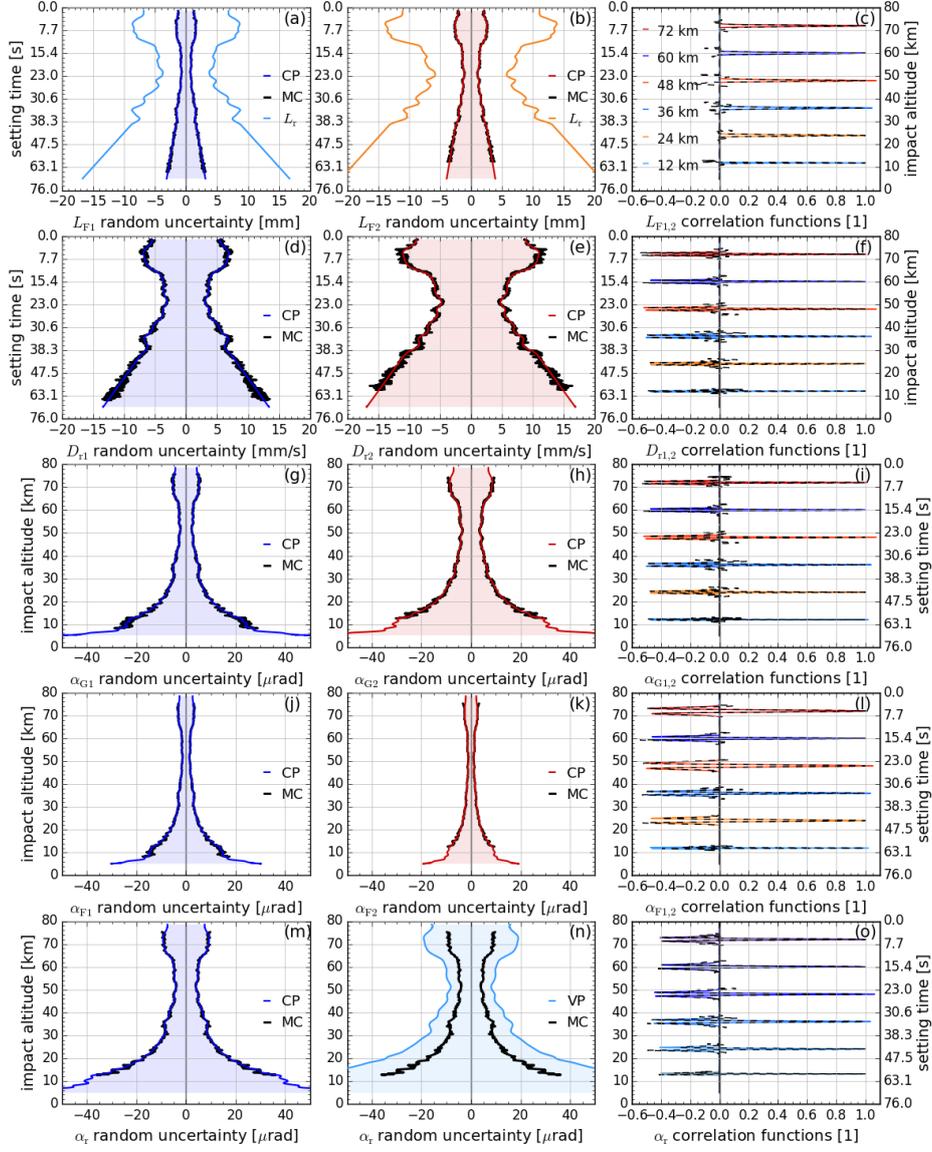


Figure 3.9: Results from the validation of CP covariance matrices $\mathbf{C}_{L_r}^{\text{CP}}$ (“CP”) by MC covariance matrices $\mathbf{C}_{L_r}^{\text{MC}}$ (“MC”) (first four rows): the consecutive retrieval steps are shown for L_F (a–c; in panels a–b also for L_r) and D_r (d–f) relative to setting time t , and for α_G (g–i) and α_F (j–l) relative to impact altitude z_a . The left column shows estimated random uncertainties for f_{T1} (CP in blue, MC in black, in panel a $u_{L_r1}^r$ in light blue); the middle column for f_{T2} (CP in red, MC in black, in panel b $u_{L_r2}^r$ in orange); the right column representative correlation functions at 60, 36, and 12 km for f_{T1} (CP in blue, MC in black) and 72, 48, and 24 km for f_{T2} (CP in red, MC in black). The last row (m–o) shows CP (blue) and MC (black) results for estimated α_r random uncertainties (m) and representative correlation functions at 72, 60, 48, 36, 24, and 13 km (o), as well as variance propagation (“VP”) results (light blue) for α_r in panel (n).

Finally, Figure 3.9m to o show the CP results for retrieved atmospheric bending angle α_r , where Figure 3.9n is included as a special cross-comparison in the case of only variance propagation being used instead of CP. Figure 3.9m and o confirm that CP results are also correct for this final L1b variable, in terms both of random uncertainty and correlation functions.

In order to demonstrate that a full CP is necessary to propagate random uncertainties correctly, we also calculated random uncertainties $u_{\alpha_r}^r$ based on mere VP from α_G to α_r for comparison. A description of this VP algorithm (i.e., only diagonal elements of the covariance matrices are considered) is provided in Appendix A.2. Figure 3.9n clearly shows that VP would overestimate random uncertainties in α_r considerably, pointing to the importance of the complete CP implementation in the L1b retrieval chain, even though the correlation lengths involved in the processing steps are rather small.

3.6 Performance demonstration

To statistically evaluate the performance of the new L1b uncertainty propagation algorithm, we also processed a complete test day of real (CHAMP, COSMIC, MetOp) and simulated (simMetOp) data from GNSS RO satellite missions. Figure 3.10 shows the results for estimated systematic and random uncertainty profiles, as well as correlation length and resolution profiles for filtered excess phase profiles $L_{F,1}$. Figure 3.11 subsequently illustrates the ensemble mean of the same variables for $L_{F,1}$, $D_{F,1}$, $\alpha_{G,1}$, and α_r for the test day ensemble. In Figure 3.10 we also co-illustrate the number of events processed for each of the RO missions (middle column).

About 5% of the total number of processed profiles for each mission have been discarded, because they were detected as outliers based on the magnitude of their random uncertainty profiles (these outliers are not included in the number of profiles shown). All profiles are shown as function of impact altitude, because each of the profiles in the ensembles needed to be interpolated to the same (standard) impact altitude grid, to orderly calculate their mean profiles.

Figure 3.10a shows $u_{L_{F,1}}^r$ and $u_{L_{F,1}}^s$ for all ~ 100 CHAMP events. It is visible (also in Figure 3.10d and g) that the random uncertainty is estimated based on excess phase noise between 30 and 75 km and synthetically extended above and below, as described in Sect. 3.3.2. For the large majority of events, $u_{L_{F,1}}^r$ lies between about 0.5 and 3 mm in the range between 30 and 75 km. Note that these results show the random uncertainties after the application of the basic BWS filter (Sect. 3.4.1), but the input uncertainties $u_{L_{F,1}}^r$ are of similar shape (though larger in magnitude).

Figure 3.10b shows that the correlation length profiles of the CHAMP ensemble (gray) and its ensemble mean (yellow) are of relatively constant magnitude from 35 to 80 km but then get smaller downward, because the RO event's scan velocity decreases (see Equation A.13). Since the BWS filter determines the vertical resolution and the correlation length at the same time, the resolution profiles $w_{L_{F,1}}$ (Figure 3.10c) are quite similar to the correlation length profiles $l_{L_{F,1}}$ (Figure 3.10b).

The number-of-events profile shows that most CHAMP events end between 5 and 12 km (Figure 3.10b, black). This is because the GO profiles illustrated here are cut off right at the lower end of the GO–WO transition range at $z_a^{\text{GW}} - \Delta z_a^{\text{GW}}$ (cf. Table 3.2).

Compared to CHAMP, the mean random uncertainty $u_{L_{F,1}}^r$ (Figure 3.10d) for the ~ 1500 events of the COSMIC ensemble is smaller, particularly above 30 km, indicating the improved data quality of this later mission. The mean of the correlation length profiles $l_{L_{F,1}}$ (Figure 3.10e) is higher than for CHAMP (Figure 3.10b), and correspondingly the resolution of the COSMIC profiles is also somewhat coarser (Figure 3.10f and c). The cutoff frequency and sampling rate – and thus the resolution in time – are set to be the same in the rOPS, irrespective of the missions; these differences hence are due to the different vertical scan velocities of the missions induced by the differences in orbit altitudes (CHAMP ~ 400 km, COSMIC ~ 700 km).

For the real MetOp data (available here as a data set from UCAR/CDAAC, as for CHAMP and COSMIC), u_{LF1}^r appears similar to COSMIC (cf. Figure 3.10d, g), while for simMetOp (with best possible simulated MetOp-type receiver noise) it is clearly smaller than for COSMIC. From 35 to 80 km the mean random uncertainty profile for simMetOp stays below 1 mm (Figure 3.10j). Three individual profiles exhibit comparatively high uncertainties of larger than 2 mm within about 40 to 55 km, however, reflecting that the simMetOp error simulations are capable of partly generating higher-noise profiles of the type more frequently seen in the real MetOp data (Figure 3.10g).

On the other hand, the average correlation length/resolution profile of the ~ 500 real MetOp and ~ 700 simMetOp ensemble members is very similar, driven by the orbit being essentially the same for the real data and the simulations (Figure 3.10h, i, k, l). Compared to COSMIC (Figure 3.10e, f), the correlation length and resolution are again somewhat larger/coarser, due to an even somewhat higher scan velocity of the MetOp satellite (~ 820 km orbit altitude). The systematic uncertainty u_{LF1}^s – just co-illustrated for completeness in Figure 3.10a, d, g, and j – is almost left unchanged by the BWS filter and is essentially equal to the preset input uncertainty for all three missions (Sect. 3.3.2).

Figure 3.11 shows how the u_{LF1}^s , u_{LF1}^r , l_{LF1} , and w_{LF1} profiles are on average affected by the uncertainty propagation. The color code for the different satellite missions is the same as in Figure 3.10. The propagation effects visible are similar to those already seen in Figs. 3.3 to 3.8. The Doppler shift derivation increases the relative uncertainties and reduces correlation length (of the main peak), while the resolution stays the same (Figure 3.11d–f). The GO bending angle retrieval leaves correlation length and resolution unchanged, while random uncertainties increase strongly in the lower stratosphere and troposphere due to the increasing refractive effects (Figure 3.11g–i).

Finally, the BWS filtering before the ionospheric correction decreases random uncertainties and increases correlation length, and resolution somewhat. However, the linear combination of the two bending angle profiles α_{F1} and α_{F2} then increases the random uncertainty again (cf. Figure 3.11j and g). The adaptive minor-channel cutoff frequency f_{c2} for the relatively noisy CHAMP profiles is generally lower than for the other two missions, and the filter effect is therefore stronger for CHAMP (indicated by the larger $l_{\alpha r}$ in Figure 3.11k)

The estimated systematic uncertainty of the atmospheric bending angle $u_{\alpha r}^s$, indicated for completeness in Figure 3.11 (left column, inflated by a factor of 10 in panel a and 100 in panels d, g, and j for somewhat better visibility), stays below $0.1 \mu\text{rad}$ for all three missions.

3 Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

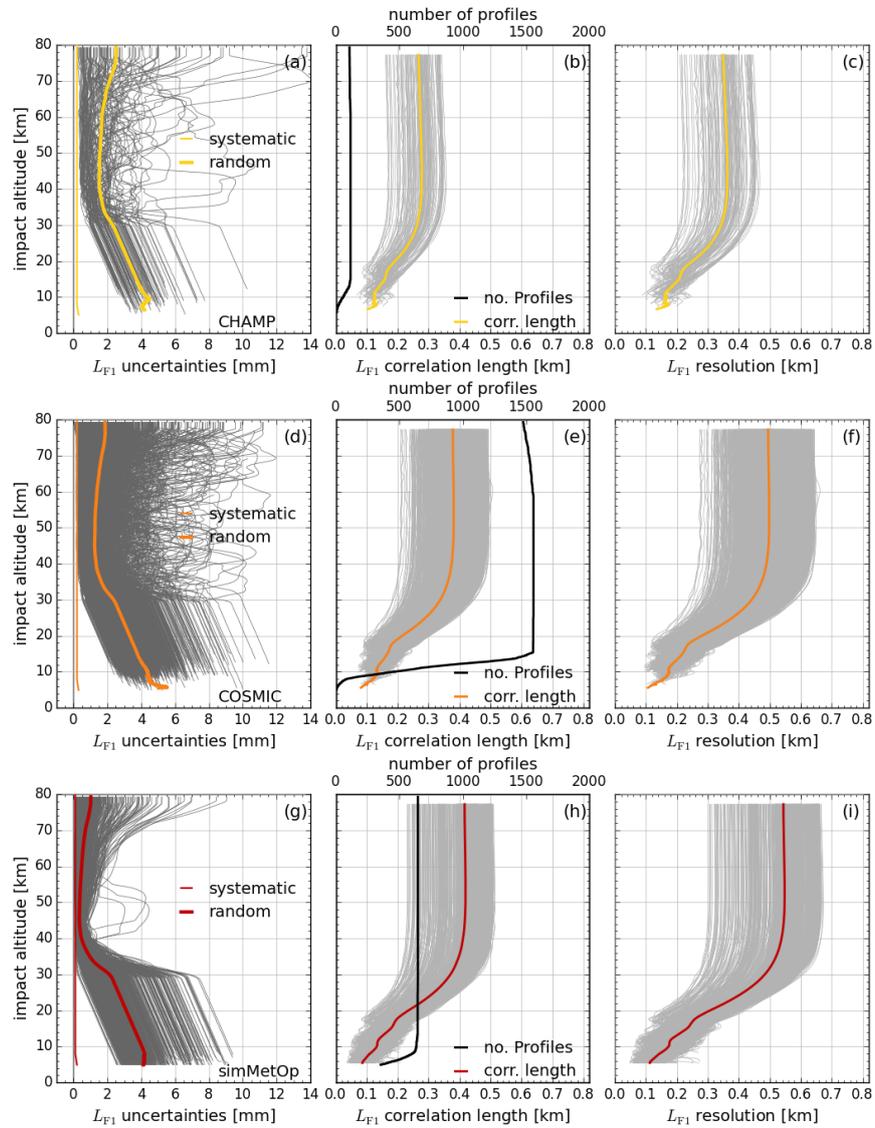


Figure 3.10: Uncertainty propagation results for real-data ensembles from 15 July 2008, for the filtered excess phase profile L_{F1} of the leading channel (f_{T1} , GPS L1 frequency). Left column: estimated random $u_{L_{F1}}^r$ (heavy) and systematic $u_{L_{F1}}^s$ (light) uncertainty profiles of each ensemble member (gray), and the ensemble mean (color) for CHAMP (a), COSMIC (d), MetOp (g), and simMetOp (j). Middle column: correlation length profiles $l_{L_{F1}}$ of each ensemble member (gray), the ensemble mean (color), and the ensemble size profile (black, scale at upper axis) for CHAMP (b), COSMIC (e), MetOp (h), and simMetOp (k). Right column: estimated resolution profile $w_{L_{F1}}$ of each ensemble member (gray) and the ensemble mean (color) for CHAMP (c), COSMIC (f), MetOp (i), and simMetOp (l).

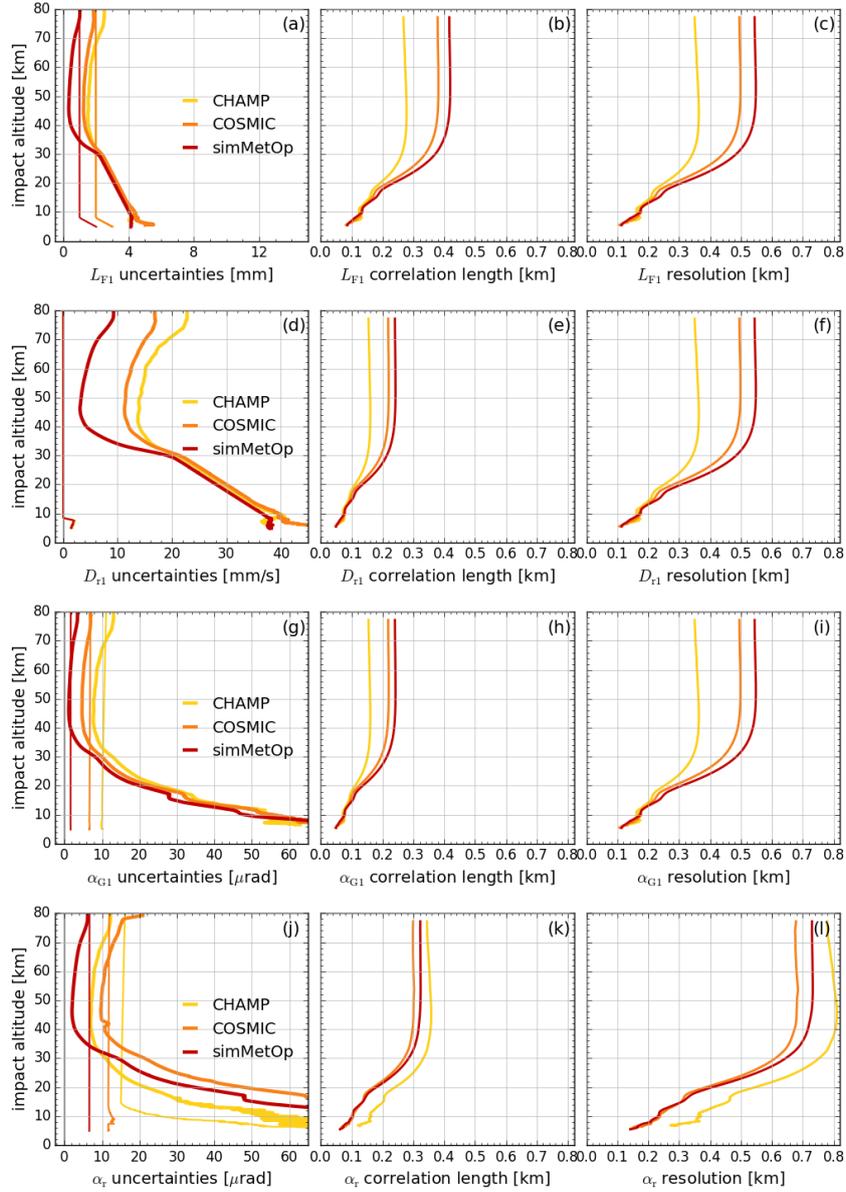


Figure 3.11: Uncertainty propagation results for real-data ensembles from 15 July 2008 for output profiles of the leading channel (f_{T1} , GPS L1 frequency). The first row shows results for L_{F1} (a–c), the second for D_{r1} (d–f), the third for α_{G1} (g–i), and the fourth for α_r (j–l). The different ensemble mean profiles are shown in colors (CHAMP, yellow; COSMIC, orange; MetOp, red; and simMetOp, violet). Left column: mean random uncertainty $u_{X_r}^r$ (heavy) and mean systematic uncertainty $u_{X_r}^s$ (light) profiles (a, d, g, j); the latter is shown as $10 \times u_{L_{F1}}^s$ (in a) and $100 \times u_{X_r}^s$ (d, g, j) for enabling visibility of these small quantities. Middle column: correlation length profiles l_{X_r} (b, e, h, k). Right column: vertical resolution profiles w_{X_r} (c, f, i, l).

3.7 Conclusions

In order to deliver climate benchmark data sets, it is essential to integrate uncertainty propagation in RO retrievals. In this study we presented the uncertainty propagation algorithm chain from excess phase profiles to atmospheric bending angle profiles (L1b processing), as newly implemented in the rOPS at the WEGC. Along with the basic profile retrieval, we provide estimates for systematic and random uncertainties, error correlation matrices, and vertical resolution profiles, which is unique amongst all existing RO processing systems so far (Ho et al. 2012; Steiner et al. 2013).

We validated the implemented algorithm via comparison to Monte Carlo sample propagation results and demonstrated the performance of the algorithm using real-data ensembles. We find close agreement between the implemented covariance propagation of random uncertainties and the Monte Carlo validation runs, verifying the correctness of the implemented algorithm. The test day ensembles for three different missions (CHAMP, COSMIC, MetOp) show reliable, robust, and consistent results that provide valuable insight and understanding of retrieval chain details.

Together with the integration of the uncertainty propagation algorithm from atmospheric bending angle profiles to dry-air profiles (L2a processing) presented by Schwarz et al. (2017), the rOPS can now provide estimates of systematic and random uncertainty profiles, of error correlation matrices and resolution, and of observation-to-background weighting ratio profiles from excess phase to dry-air profiles.

The next step towards the final atmospheric profiles, currently ongoing, is the introduction of integrated uncertainty propagation for the moist-air retrieval (L2b processing). Implementation of uncertainty propagation for the wave-optics bending angle retrieval and for the orbit determination and excess phase processing (L1a processing) is ongoing as well.

Once completed, the full rOPS retrieval chain will run with integrated uncertainty estimation, a major step towards climate benchmark data provision, and beneficial for the wide diversity of uses in atmospheric and climate science and applications.

3.8 Acknowledgements

We thank UCAR/CDAAC Boulder for access to their RO excess phase and orbit data (available at <http://cdaac-www.cosmic.ucar.edu/>) as well as ECMWF Reading for access to their analysis and forecast data (available at <http://www.ecmwf.int/en/forecasts/datasets>). To access the relevant result files of the uncertainty propagation, please contact the corresponding author. The developed algorithm is provided in the appendix. The work was funded by the Austrian Aeronautics and Space Agency of the Austrian Research Promotion Agency (FFG-ALR) under projects OPSCLIMPROP and OPSCLIMTRACE and by the European Space Agency (ESA)

under project MMValRO-E.

Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

This chapter presents work that was published in *Earth and Space Science (ESS)* by Schwarz, Kirchengast and Schwärz (2017). The work in the context of this publication included design and implementation of uncertainty propagation in the rOPS retrieval, starting from retrieved and background bending angle profiles, through the high altitude initialization, the refractivity retrieval, and the dry-air retrieval.

My contribution to this study included the analytical derivation of the uncertainty propagation formulas, the conversion into numerical propagation algorithms, the implementation of the algorithms into the rOPS, the testing and validation of the algorithms and the drafting of the manuscript for the publication.

4.1 Abstract

Global Navigation Satellite System (GNSS) Radio Occultation (RO) observations, globally available as a continuous record since 2001 are highly accurate, and long-term stable data records. Essential climate variables for the thermodynamic state of the free atmosphere, such as temperature and tropospheric water vapor profiles (involving background information), can be derived from these records, which consequentially have the potential to serve as climate benchmark data. In order to exploit this potential, atmospheric profile retrievals need to be very accurate and the remaining uncertainties need to be quantified and traced throughout the retrieval chain. The new Reference Occultation Processing System (rOPS) at the Wegener Center aims to deliver such an accurate retrieval chain with integrated uncertainty propagation. Here we introduce and demonstrate the algorithms implemented for uncertainty propagation from RO bending

angle profiles to dry-air variables (pressure, temperature), for estimated random and systematic uncertainties and for co-estimates of observation-to-background weighting ratio profiles. We estimated systematic uncertainty profiles with the same operators as used for the basic profiles retrieval. The random uncertainty propagation was integrated by a covariance propagation approach and validated using Monte-Carlo ensemble methods. We present the results of the validation, and demonstrate how the algorithm performs for individual simulated RO events and for ensembles of real RO events. We also compare the new results from the integrated uncertainty propagation to previous ones from empirical error analyses for RO-retrieved atmospheric profiles. We find that the new uncertainty estimation chain shows robust performance and is in good agreement with previous comparable results.

4.2 Introduction

Historically, observation systems of the free atmosphere were designed for weather research and forecasting purposes, and have considerable shortcomings when looking at them from a climate monitoring perspective (Karl et al. 1995). According to Bojinski et al. (2014) till today much of the global climate monitoring infrastructure remains fragile and incomplete. The observational foundation for the climate sciences needs to be improved by the establishment of climate benchmark data records, i.e. data records which 1. are of global coverage, 2. are of high accuracy, 3. are long-term stable, 4. are tested for systematic errors on-orbit, 5. are tied to irrefutable standards, and 6. measure Essential Climate Variables (ECVs) (NRC 2007) (GCOS 2015).

Observations from space are needed to obtain these data records with global coverage (CEOS 2012). Current satellite missions allow for the observation of thermodynamic ECVs in the free atmosphere, but are unable to provide them as climate benchmark dataset (Leroy et al. 2006) (Kirchengast and Schweitzer 2011). The Global Navigation Satellite System (GNSS) Radio Occultation (RO) remote sensing technique has the potential to provide it (Kirchengast et al. 2016a; Kirchengast et al. 2016b; Leroy et al. 2006; Steiner et al. 2011).

4.2.1 The RO Method

In an RO event, signals of GNSS satellites—in practice so far from the Global Positioning System (GPS)—propagate through the Earth’s atmosphere in limb sounding geometry and are received by a Low Earth Orbit (LEO) satellite’s GNSS receiver after exiting from the atmosphere. Due to the vertical density gradient of the atmosphere, the GNSS signals are refracted and follow a bent trajectory, resulting in an excess phase path before reaching the receiver. Depending on the motion of the transmitting GNSS and the receiving LEO satellite relative to the Earth, consecutive measurements scan the atmosphere from top downwards (setting event) or from bottom upwards (rising event),

yielding near vertical profiles of excess phase measurements on a dense time grid of typically 50 Hz sampling rate (Hajj et al. 2002; Kursinski et al. 1997; Schreiner et al. 2010). The accuracy of the measurement depends on the accuracy of the clocks in the LEO and GNSS satellites (effectively the accuracy of the SI second), which are commonly highly stable to fractions of a nanosecond. At the top of the neutral gas atmosphere at around 80 km (at the minimum between residual neutral gas and residual ionospheric refractive effects), the signal passes almost unattenuated and unrefracted from the transmitter to the receiver and therefore the instrument is in this sense naturally calibrated at each individual RO event; up to ionospheric influence which is calibrated by dual-frequency ionospheric correction of GNSS signals (Healy and Culverwell 2015; Ladreiter and Kirchengast 1996; Liu et al. 2015; Sokolovskiy et al. 2009; Syndergaard 2000; Vorob'ev and Krasil'nikova 1994).

The RO observation record can therefore serve as a Fundamental Climate Data Record (FCDR) as defined by GCOS (2010a): It is globally available (due to the abundance of signal sources), long-term stable (due to self-calibration and high signal stability during the event), highly accurate (accuracy traceable to the SI second), and due to the self-calibrating property, the accuracy is also ensured on-orbit (Leroy et al. 2006).

4.2.2 Essential Climate Variables from RO

With the help of a retrieval, ECVs (such as the thermodynamic ECVs pressure, temperature and humidity in the free atmosphere) can be derived from the RO FCDR with its unique properties (in the moist troposphere also including background information). The basic steps of an RO processing chain, as applied to each individual RO event, are schematically summarized in Figure 4.1. It starts with the precise orbit determination (POD) and excess phase processing (labeled 'L1a' in Figure 4.1), which involves accurate and precise GNSS and LEO orbit provision and basic geodetic corrections to isolate the atmospheric excess phase at two GNSS frequencies as time series, serving as basic RO event FCDR for the subsequent retrieval (alternatively the Doppler shift profile or the bending angle profile on the measurement grid may be considered FCDRs). The subsequent bending angle retrieval yields vertical bending angle profiles, representing the cumulative bending of the GNSS signal attributed to each vertical impact altitude gridpoint ('L1b'). The low signal-to-noise ratio of the measurements at high altitudes, requires a *high altitude initialization* ('L2a (1)'), otherwise the integrals in the following retrieval steps (Abelian integral, hydrostatic integral) transport too much noise to lower altitude levels. To reduce the noise, raw bending angle profiles are initialized with smooth background bending angle data by means of a statistical optimization leading to optimized bending angle profiles. In the *refractivity retrieval* ('L2a (2)'), using an Abel integral transform, refractivity profiles as function of mean-sea-level (MSL) altitude (termed *altitude* hereafter) are derived from these optimized bending angle profiles. The refractivity profiles are used in a *dry air retrieval* ('L2a (3)') to calculate dry density,

dry pressure and dry temperature profiles as a function of altitude. In dry-air profiles, atmospheric moisture is disregarded by neglecting the humidity-dependent 'wet term' in the refractivity equation (Smith and Weintraub 1953) irrespective of the actual humidity of the atmosphere (for details on the difference between dry-air and actual pressure and temperature profiles see Scherllin-Pirscher et al. (2011b) and Scherllin-Pirscher et al. (2017)). Finally, using background information on humidity and/or temperature in the troposphere, atmospheric profiles of density, pressure, temperature and humidity are calculated in a moist air retrieval ('L2b').

Numerous studies have described this retrieval chain in detail and have shown the high accuracy of RO data, particularly in the upper troposphere and lower stratosphere (UTLS) region, see e.g. (Gobiet et al. 2007; Hajj et al. 2002; Ho et al. 2012; Kuo et al. 2004; Kursinski et al. 1997; Scherllin-Pirscher et al. 2011a; Scherllin-Pirscher et al. 2011b; Steiner and Kirchengast 2005; Steiner et al. 2013).

RO-derived ECVs and their uncertainties can potentially be traced to the (small) uncertainties of the FCDR (and in turn to the raw data), but this requires that 1. the retrieval is highly *accurate* and avoids any undue amplification of uncertainties associated with the quantities in the FCDR and 2. these uncertainties need to be propagated through the retrieval, in order to trace the claimed accuracy (expressed by the uncertainties provided) of the ECV data back to the FCDR and the raw data.

4.2.3 Reference Occultation Processing System (rOPS)

Developed at the Wegener Center of the University of Graz (WEGC), the Reference Occultation Processing System (rOPS) project (Kirchengast et al. 2017b) aims to establish such a fully traceable processing (Kirchengast et al. 2015; Kirchengast et al. 2016b). It aims to 1. create an RO retrieval chain of high accuracy (residual relative numerical errors at all steps $< 10^{-6}$ to 10^{-4}), and to 2. integrate uncertainty propagation from raw GNSS orbits and RO measurements to atmospheric variables along the retrieval chain.

Concerning retrieval advancements as relevant to this study, rOPS inclusion of the high altitude initialization algorithm introduced by Li et al. (2013) and Li et al. (2015) is ongoing. Following work of Syndergaard (2000), Liu et al. (2015), Healy and Culverwell (2015) and Danzer et al. (2013) and Danzer et al. (2015), remaining higher-order ionospheric effects in the retrieved bending angles are planned to be significantly reduced. Uncertainty propagation for the wave-optics bending angle retrieval has been implemented and demonstrated for simulated events by Gorbunov and Kirchengast (2015); work and manuscript preparation based on real RO data is on-going.

Further studies with advanced progress will introduce the uncertainty propagation for the entire 'L1a' and 'L1b' processing parts (Figure 4.1 and Section 4.2.2 above) of the rOPS.

This study is a key contribution focussing on the uncertainty propagation from bending angle to dry-air profiles ('L2a' processing), after which uncertainty propagation for the

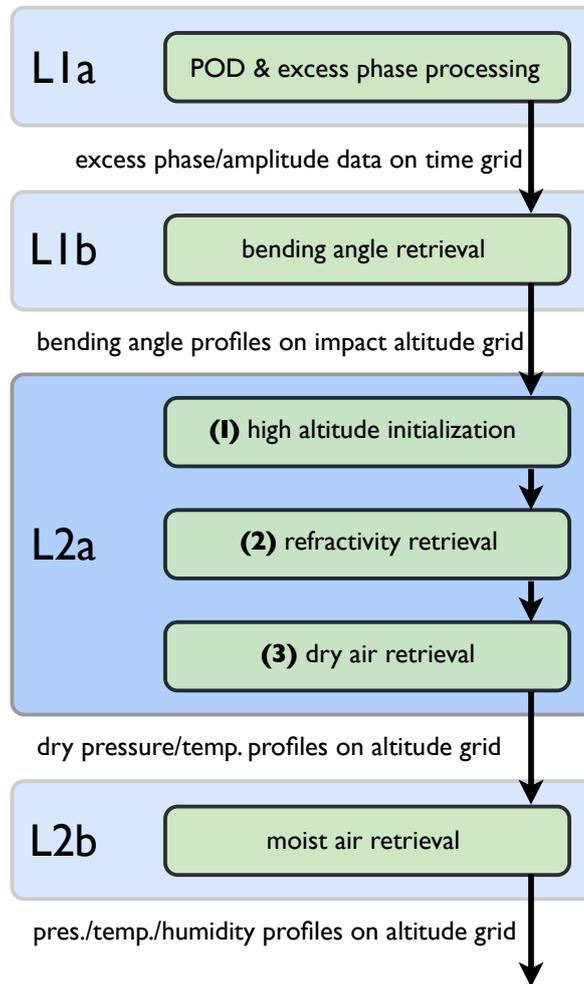


Figure 4.1: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b, L2a, L2b) and the main operators of the L2a processor (1, 2, 3), which are in the focus of this study.

moist air retrieval ('L2b') will be introduced. Related work and manuscript preparation on an advanced moist air retrieval algorithm, including dynamical background information in the troposphere—similar to its use by Li et al. (2013) and Li et al. (2015) at high altitudes—is on-going.

4.2.4 Study Approach and Outline

Uncertainty propagation as covariance propagation (CP) from bending angle to dry-air variables has been outlined, e.g., by Syndergaard (1999) and Rieder and Kirchengast (2001a), but as yet not been implemented in retrievals and applied to real RO data. We propagate estimated random uncertainties from retrieved bending angles to 1. optimized bending angles, 2. refractivity, 3. dry pressure and dry temperature profiles, using a full CP approach. We also propagate conservative bound estimates for systematic uncertainties. We use individual simulated RO events and ensembles of real RO events to test and demonstrate the algorithm.

The propagated uncertainties are put into context with the results of an *empirical* error analysis using simulated RO data by Steiner and Kirchengast (2005) [SK2005 hereafter], who calculated errors of RO profiles based on ensembles of profiles from end-to-end simulations with the End-to-End GNSS Occultation Performance Simulation and Processing System (EGOPS) (Fritzer et al. 2009). We will also compare the results to those of Scherllin-Pirscher et al. (2011a) [SP2011 hereafter]. SP2011 created an error model for the relevant RO variables bending angle, refractivity, dry pressure and dry temperature, based on an empirical analysis of real RO data from different RO missions and data centers.

Section 4.3 summarizes the applied uncertainty propagation methods, gives an overview of the used input data sources, and describes the basic setup and input data of the testing framework used for the algorithm description and the performance demonstration. In Section 4.4, we present the sequence of the new uncertainty propagation algorithms and show how the algorithms work by way of an example case. A detailed description of the algorithms can be found in Appendix B.1. In Section 4.5 we demonstrate the appropriate functioning and performance of the algorithm. First we show the results of a validation of the covariance-based propagation by Monte-Carlo (MC) methods. Then we demonstrate the geographic robustness by means of three geographically dispersed simulated example cases. Additionally we run test-day ensembles with real data from the RO missions CHALLENGING Minisatellite Payload (CHAMP) (Wickert et al. 2001), the FORMOSAT-3 Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) (Anthes et al. 2008), and the Meteorological Operational Satellite A (MetOp) (Luntama et al. 2008). For MetOp we additionally used an ensemble of EGOPS simulated input data [*simMetOp* hereafter]. We close with conclusions and outlook in Section 4.6.

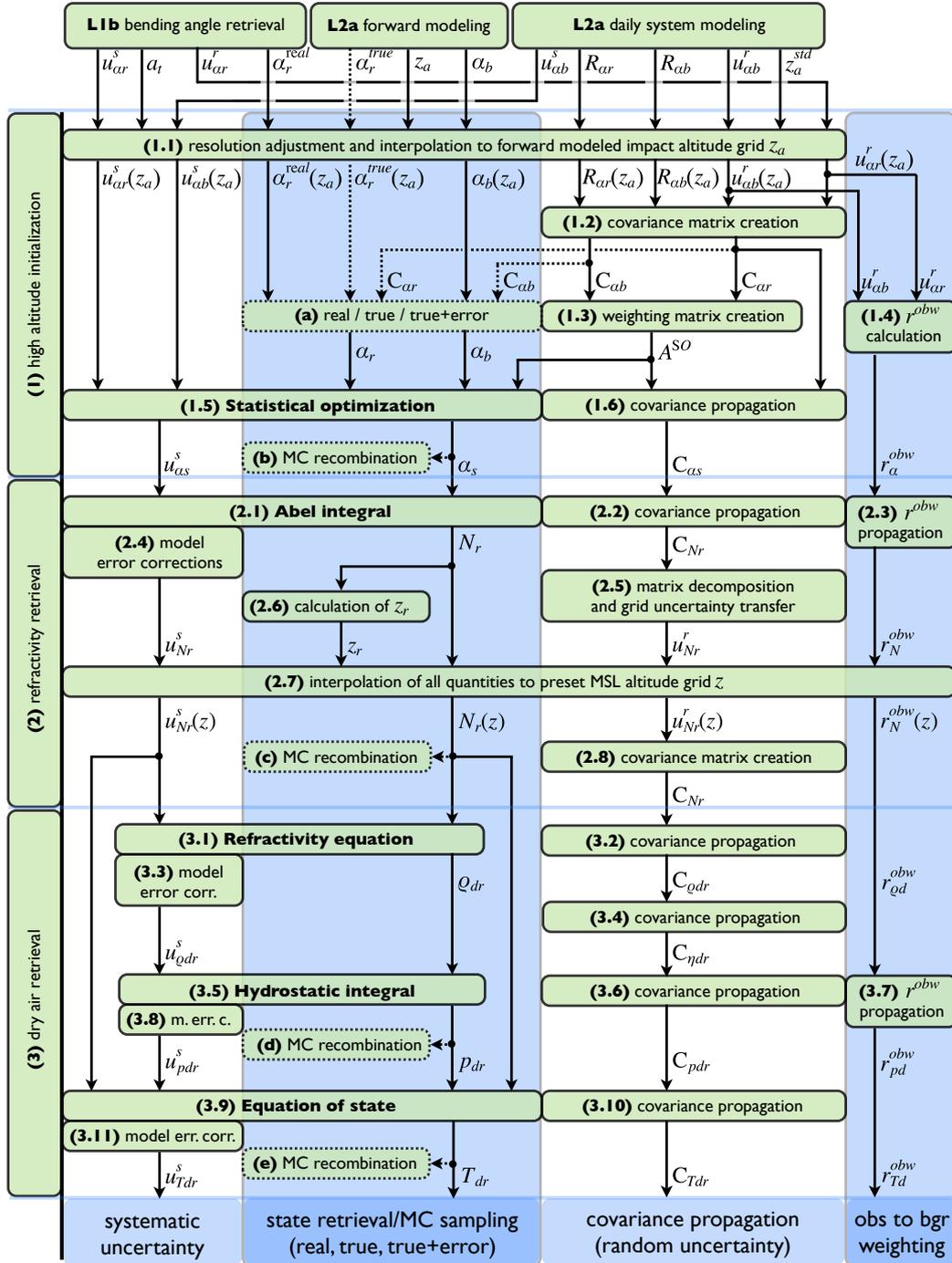


Figure 4.2: Detailed workflow for state retrieval and uncertainty propagation of the main L2a operators from bending angle to dry-air profiles (1)-(3) and of the subroutines used in the MC testing framework (a)-(e). The mathematical notation, including all symbols, is introduced in Tables 4.1, 4.2 and 4.3.

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

Table 4.1: Principal input variables for uncertainty propagation

Variable	Unit	Description
$X(z_a)$	rad	with $X \in \{\alpha_r, \alpha_b\}$, retrieved/ background bending angle profile at z_a grid, comprising elements X_i .
$u_X^s(z_a)$	rad	with $X \in \{\alpha_r, \alpha_b\}$, estimated systematic uncertainty profile of $X(z_a)$, comprising elements $u_{X,i}^s$.
$u_X^r(z_a)$	rad	with $X \in \{\alpha_r, \alpha_b\}$, estimated random uncertainty profile of $X(z_a)$, comprising elements $u_{X,i}^r$.
\mathbf{R}_X	1	with $X \in \{\alpha_r, \alpha_b\}$, error correlation matrix of $X(z_a)$, comprising elements $R_{X,ij}$.
\mathbf{C}_X	rad ²	with $X \in \{\alpha_r, \alpha_b\}$, error covariance matrix of $X(z_a)$, comprising elements $C_{X,ij} = u_{X,i}^r \cdot u_{X,j}^r \cdot R_{X,ij}$.

4.3 Methods and Data

4.3.1 Methods

We distinguish uncertainty propagation from error propagation, following the *Guide to the expression of Uncertainty in Measurement* (GUM) (JCGM 2008a) (JCGM 2008b) (JCGM 2011) [GUM hereafter]. We recommend the GUM as a key reference for readers interested in a thorough and exhaustive introduction to uncertainty propagation. We also aim to follow terminology as provided by the International Vocabulary of Metrology (VIM) (JCGM 2012), a terminology also adopted by the GUM.

The aim of the integrated uncertainty propagation in the rOPS is to eventually propagate uncertainties along the entire retrieval chain from the raw measurement data to the ECVs (Kirchengast et al. 2016a; Kirchengast et al. 2016b). However, for the purpose of this study, we first prepare input data at the beginning of the L2a processor. The workflow of the L2a processor is sketched in Figure 4.2, expanding the respective schematic block of Figure 4.1. The mathematical symbols used for the main operators and output variables are defined in Table 4.2, to provide a concise overview.

The effects of unpredictable or stochastic temporal and spatial variations of the influencing quantities, called *random effects* by the GUM, give rise to variations in repeated observations of a measurand, and the uncertainty due to these effects can be estimated by a so-called *type A evaluation* and expressed by a statistical standard deviation estimate (JCGM 2008a). We use the term *estimated random uncertainty* to stress the stochastic character of this uncertainty estimate. In RO data, the main contributor to these effects is the receiving system noise, in particular from thermal noise and residual clock estimation noise, which is essentially normally distributed (Kursinski

Table 4.2: Principal operators and output variables

Variable	Unit	Description
\mathbf{A}^{SO}	1	weighting matrix, operator of the statistical optimization.
$\alpha_s(z_a)$	rad	statistically optimized bending angle profile at z_a grid
$u_{\alpha_s}^s(z_a)$	rad	estimated systematic uncertainty profile of $\alpha_s(z_a)$.
$u_{\alpha_s}^r(z_a)$	rad	estimated random uncertainty profile of $\alpha_s(z_a)$.
\mathbf{R}_{α_s}	1	error correlation matrix of $\alpha_s(z_a)$.
\mathbf{C}_{α_s}	rad ²	error covariance matrix of $\alpha_s(z_a)$.
$r_{\alpha}^{\text{obw}}(z_a)$	%	observation-to-background weighting ratio of $\alpha_s(z_a)$.
\mathbf{A}^{AbI}	1	Abel transform matrix, operator of Abel integral.
\mathbf{A}^{Hyd}	1	Hydrostatic transform matrix, operator of hydrostatic integral.
\mathbf{A}^{N2T}	1	Refractivity-to-temperature linearized matrix operator of dry temperature retrieval.
$Y_r(z)$	unit	with $Y \in \{N, \rho_d, p_d, T_d\}$, retrieved refractivity/dry density/dry pressure/dry temperature profile at z grid, comprising elements $Y_{r,i}$.
$u_{Y_r}^s(z)$	unit	with $Y \in \{N, \rho_d, p_d, T_d\}$, estimated systematic uncertainty profile of $Y(z)$, comprising elements $u_{Y_r,i}^s$.
$u_{Y_r}^r(z)$	unit	with $Y \in \{N, \rho_d, p_d, T_d\}$, estimated random uncertainty profile of $Y(z)$, comprising elements $u_{Y_r,i}^r$.
\mathbf{R}_{Y_r}	1	with $Y \in \{N, \rho_d, p_d, T_d\}$, error correlation matrix of $Y(z)$, comprising elements $R_{Y_r,ij}$.
\mathbf{C}_{Y_r}	unit ²	with $Y \in \{N, \rho_d, p_d, T_d\}$, error covariance matrix of $Y(z)$, comprising elements $C_{Y_r,ij} = u_{Y_r,i}^r \cdot u_{Y_r,j}^r \cdot R_{Y_r,ij}$.
$r_Y^{\text{obw}}(z)$	%	with $Y \in \{N, \rho_d, p_d, T_d\}$, observation-to-background weighting ratio of $Y_r(z)$, comprising elements $r_{Y,i}^{\text{obw}}$.

et al. 1997; Syndergaard 1999).

The GUM advises to make every effort to identify *systematic effects*, typically requiring *type B evaluation* (methods other than statistical data analysis), and correct for all such effects recognized (JCGM 2008a). We follow this, and for the residuals that will always remain we provide and propagate so-called conservative bound estimates. We assume that all residual potential systematic effects lie within these bound estimates (based on best educated-guess estimates and previous point-wise studies, but without quantitative knowledge about the exact values for each individual RO event). We call these conservative bound profiles *estimated systematic uncertainty*.

As summarized in Figure 4.1, the rOPS RO retrieval has a chain-like workflow. The

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

Table 4.3: Vertical grids, coordinate variables, and specific vertical levels

Variable	Unit	Description
z	m	Preset, equidistant, MSL altitude grid. Used in top-downward direction like all other vertical grids in the rOPS, with $z_1 = 80000$ m, $z_n = 100$ m, $z_i - z_{i+1} = 100$ m. All quantities are interpolated to this grid upon output.
z_a	m	impact altitude grid, with $z_a = z + 10^{-6} \cdot N_b(z) \cdot (z + h_G + R_C) = f(z, N_b)$, or equivalently $z_a = f(z_r, N_r)$.
a_t	m	impact parameter at the 50 Hz measurement time grid. After the L1b bending angle retrieval, α_r and its uncertainties are provided on a_t .
z_a^{std}	m	standard impact altitude grid, based on the refractivity of the International Standard Atmosphere, i.e. $z_a^{\text{std}} = f(z, N^{\text{std}})$. The daily quantities are provided on z_a^{std} .
a_z	m	impact parameter at z_a grid, $a_z = z_a + h_G + R_C$.
z_r	m	retrieved MSL altitude grid, calculated from z_a and the retrieved refractivity $N_r(z_a)$, i.e., $z_r = f(z_a, N_r)$.
z_{ar}	m	impact altitude profile at z grid, $z_{ar} = f(z, N_r)$. All bending angle related quantities are interpolated to this grid for output and can thus also be used as function of MSL altitude z .
$z_{a\text{TopMax}}$	m	maximum top impact altitude for all atmospheric profiles, set to 80000 m.
$z_{a\text{TopMeas}}$	m	impact altitude of the top of the atmospheric profile, can lie between 70000 and 80000 m.
$z_{a\text{ST}}$	m	impact altitude of the bottom of the statistically optimized height range, set to 30000 m, center of the sinusoidal transition range of half-width $\Delta z_{a\text{ST}}$.
$\Delta z_{a\text{ST}}$	m	impact altitude transition half-width of the half-sine-weighted transition between the statistically optimized and the pure retrieved bending angle. Set to 2000 m.
$z_{\text{obw}50}$	m	MSL altitude at which r^{obw} reaches 50 %, i.e., the altitude above which the background information and below which the observation information dominates.

main measurement/retrieval models—or *operators*—of the L2a processor (i.e. boldfaced Items 1.5, 2.1, 3.1, 3.5 and 3.9 in Figure 4.2) typically have the following properties:

They are 1. explicit (i.e., they allow for an explicit expression $Y = f(X)$ of the output variable Y in terms of the input variable X); they are 2. multivariate (i.e., Y is a rank-1 vector of random variables); they are 3. linear (i.e., $f(X)$ can be formulated as matrix multiplication $Y = \mathbf{A}^{\text{XY}} \cdot X$) with the exception of the dry temperature retrieval (Item 3.9 in Figure 4.2); and they have 4. multiple correlated input variables (i.e., X is a rank-1 vector of random variables and their covariances need to be considered).

Therefore we follow the GUM uncertainty framework for explicit, multivariate, linear measurement models with correlated input quantities (JCGM 2011). The estimated random uncertainties are thus propagated according to

$$\mathbf{C}_Y = \text{E}[YY^T] = \text{E}[(\mathbf{A}^{\text{XY}}X)(\mathbf{A}^{\text{XY}}X)^T] = \mathbf{A}^{\text{XY}}\text{E}[XX^T](\mathbf{A}^{\text{XY}})^T = \mathbf{A}^{\text{XY}}\mathbf{C}_X(\mathbf{A}^{\text{XY}})^T. \quad (4.1)$$

Here \mathbf{C}_X and \mathbf{C}_Y are the covariance matrices of the input and output variables, respectively, and \mathbf{A}^{XY} is the operator in question. The input covariance matrix is calculated using random uncertainty estimates u_X^r and correlations \mathbf{R}_X as summarized in Table 4.1. This multivariate uncertainty propagation is referred to as *covariance propagation* (CP) method, as opposed to the *Monte Carlo* (MC) method.

The *state retrieval* is the calculation of an individual rank-1 realization (e.g., an individual vertical RO *profile*) y of random variable Y , from an individual rank-1 realization x of random variable X , by means of the operator \mathbf{A}^{XY} ,

$$y = \mathbf{A}^{\text{XY}} \cdot x. \quad (4.2)$$

The MC method (introduced in Section 4.3.3) is used to validate the CP method by applying the state retrieval to a large number of individual realizations.

Thanks to the linearity of \mathbf{A}^{XY} , the conservative bound estimate profiles of the systematic uncertainties do not need to be transferred from step to step by using

$$y + u_Y^s = f(x + u_X^s), \quad (4.3)$$

but rather can simply be propagated by application of the state retrieval operator on the systematic uncertainty estimates

$$u_Y^s = \mathbf{A}^{\text{XY}} \cdot u_X^s, \quad (4.4)$$

where u_X^s and u_Y^s are the rank-1 systematic uncertainty profiles of the input and output variables. We refer to this propagation as the *systematic uncertainty propagation*.

Furthermore, in order to characterize the influence of the background information used in the statistical optimization on all output profiles, we also propagate the observation-to-background weighting ratio r^{obw} , as defined by Li et al. (2015) (see below in Section B.1.1).

Algorithmically in the implementation, the *state retrieval*, the *propagation of the estimated systematic uncertainty* and the *MC validation* make use of (almost) the same chains of subroutines (see Figure 4.2), since they are all based on the same operators \mathbf{A}^{XY} . The *covariance propagation for the estimated random uncertainty*, and the *r^{obw} propagation* use separate subroutines.

4.3.2 Data Sources

The input quantities needed for the L2a uncertainty propagation, i.e. state X , systematic uncertainty u_X^s , random uncertainty u_X^r , and correlation matrices \mathbf{R}_X for the retrieved and the background bending angle, i.e. $X \in \{\alpha_r, \alpha_b\}$, are listed in Table 4.1. The construction of the retrieved bending angle α_r (real or simulated, with or without error superimposed) depends on the algorithm test type, and is modified accordingly at the beginning of each run (see Item (a) in Figure 4.2). For the uncertainties of the retrieved state we—for the purposes of this study—resort to results and experiences from previous studies and estimates. These quantities are estimated in a reasonable manner for now, but will later be provided by the rOPS L1b uncertainty propagation. Correlations and background quantities are provided by a prototype of the *daily system modeling processor* introduced by Kirchengast et al. (2016a) and Kirchengast et al. (2016b) and applied by Li et al. (2015).

For the algorithm test types with simulated data—as applied to illustrate the algorithm sequence (Section 4.4), the MC validation (Section 4.5.1) and the performance evaluation based on individual events (Section 4.5.2)—we use three simulated CHAMP event cases from August 8th, 2008, listed in Table 4.4. These cases are termed TRO (tropical), MID (mid-latitude), and POL (polar) hereafter.

Table 4.4: Simulated test cases August 08th, 2008.

No.	Label	GPS-ID	Mission	Lat.	Lon.	Time	Characterization
1	TRO	02	CHAMP	6.8S	138.5E	23:42	tropical event
2	MID	08	CHAMP	34.9S	126.6E	10:57	mid-latitude event
2	POL	22	CHAMP	79.9S	79.4W	08:11	polar event

The latitude, longitude and time of TRO, MID, POL were used in the rOPS software (Kirchengast et al. 2017b) to create co-located *true* bending angle profiles forward modeled from European Centre for Medium-Range Weather Forecasts (ECMWF) *analysis* fields. ECMWF analysis fields were chosen for their proven leading quality (Bauer et al. 2015; Untch et al. 2006). The ECMWF analysis fields used have a horizontal resolution of about 300 km (triangular truncation T42)—which corresponds to the approximate horizontal resolution of RO profiles—and are available at 91 vertical levels (L91). Co-located refractivity profiles were calculated from the ECMWF analysis fields. Then these *true* refractivity profiles were forward-modeled to bending angle profiles using an Abel transform. These *true* bending angle profiles were then used as source for the input variable α_r^{true} . Additionally to these *true* profiles, we created bending angle profiles with an observation error superimposed, to be more representative for realistic events. This *realistic* bending angle profile is shown as 'ret MID' in panel (a) of Figure 4.3. Section 4.3.3 explains in detail how the superimposed error was determined and added. For the

purpose of increasing comprehensibility, all input quantities (MID case) used for the algorithm description (Section 4.4) are shown in Figure 4.3.

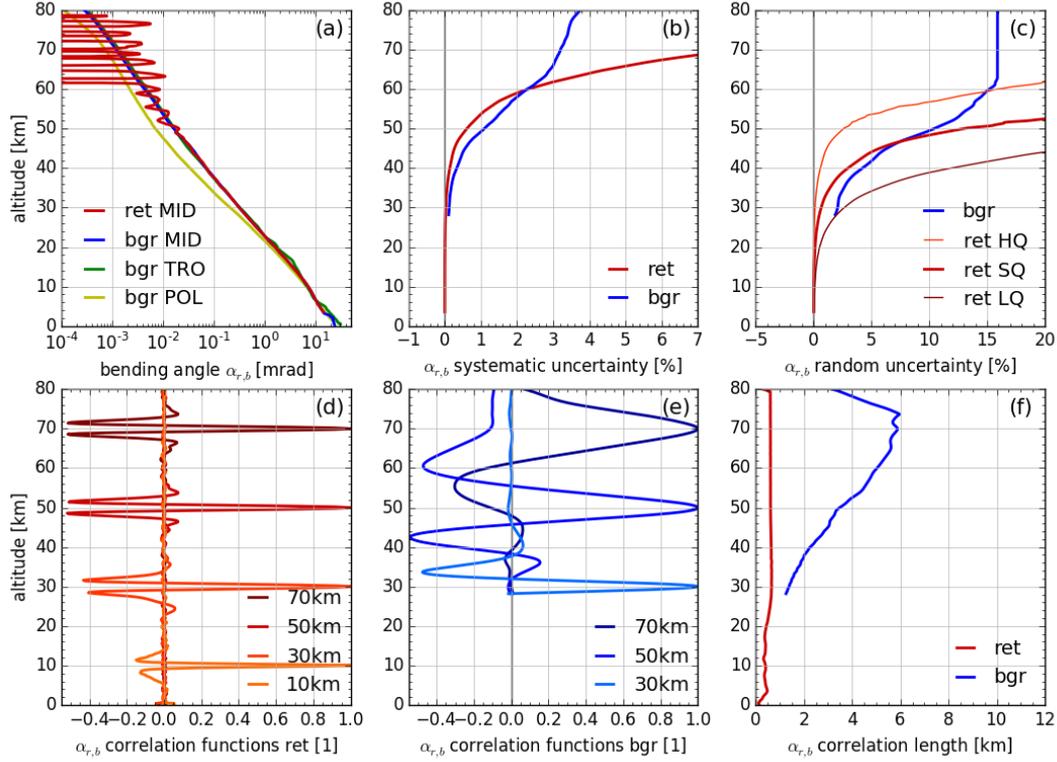


Figure 4.3: Input profiles for retrieved bending angle α_r and background bending angle α_b (case MID and scenario SQ, if not stated otherwise): (a) Bending angle α_r for MID ('ret MID') and α_b for all cases ('bgr MID', 'bgr POL', 'bgr TRO'), (b) relative systematic uncertainties $100 \cdot u_{\alpha_r}^s / \alpha_b$ ('ret') and $100 \cdot u_{\alpha_b}^s / \alpha_b$ ('bgr') for MID, (c) relative random uncertainties $100 \cdot u_{\alpha_r}^r / \alpha_b$ for MID and all scenarios ('ret HQ', 'ret SQ', 'ret LQ') and $100 \cdot u_{\alpha_b}^r / \alpha_b$ ('bgr') for MID, (d,e) representative correlation functions for MID at [10,] 30, 50 and 70 km, $R_{\alpha_r, i}$ and $R_{\alpha_b, i}$, and (f) correlation lengths L_{α_r} ('ret') and L_{α_b} ('bgr') for MID.

The arrows indicating the use of $\alpha_r^{\text{true}}(z_a)$ between Item 1.1 and Item (a) in Figure 4.2 are intentionally shown as dotted, because simulated data or ECMWF analysis data are only used for the test runs presented here (in Sections 4.4, 4.5.1 and 4.5.2). In the operational rOPS retrieval and also in runs with the test-day ensembles (in Section 4.5.3) real observed bending angle profiles α_r^{real} from the rOPS L1b bending angle retrieval are used. For the *real* bending angle profiles α_r^{real} in Section 4.5.3 we used ensembles of real RO events from the 15th of January and 15th of July 2008 (only those from July are shown below in Figures 4.10 and 4.11, those from January served as crosscheck). For

CHAMP, COSMIC, and MetOp, orbit data and excess phase profiles were provided by the COSMIC Data Analysis and Archiving Center (CDAAC) of the University Corporation for Atmospheric Research (UCAR). EGOPS was used for generating the simulated orbit data and excess phase profiles with realistic receiver noise for simMetOp. Based on these excess phase profiles from UCAR/CDAAC and EGOPS, we retrieved the *real* bending angle profiles with the rOPS. Although the bending angle profiles used as input for the simMetOp ensemble are based on *realistic* simulations rather than *real* measurement data, they will be called *real* for simpler reference.

Analogously to α_r^{true} , co-located profiles forward modeled from ECMWF short-range (24 h) *forecast* fields were used as source for input profiles α_b . The chosen 1-day forecast lead time is sufficiently long to ensure independence of the background from assimilated RO measurements in the analysis, and at the same time sufficiently short to ensure a skilled high quality forecast (sensitivity tests showed weak sensitivity to lead times within 12 h to 36 h, so 24 h is well suitable). For the MID case, α_b , denoted 'bgr MID', is visible as blue line in panel (a) of Figure 4.3.

All input variables are given on an impact altitude grid z_a (see Table 4.3), which is forward modeled from an equidistant 100 m altitude grid z and the co-located refractivity profile of the ECMWF short-range forecast.

For the ensemble based performance demonstration (in Section 4.5.3), the random uncertainties of the retrieved bending angle $u_{\alpha_r}^r(z_a)$ were estimated based on the noise of retrieved bending angle profiles $\alpha_r(a_t)$ on the native measurement grid. The noise was determined following the estimation scheme for observation errors described by Li et al. (2015, Section 2.2 therein).

For the other tests (in Sections 4.4, 4.5.1, 4.5.2) the random uncertainties were chosen such as to represent three different receiver *scenarios*, which approximately cover the range of random uncertainties of existing GNSS RO receiver systems: A *high quality* scenario ('HQ', with an uncertainty of $0.5 \mu\text{rad}$), a *standard quality* scenario ('SQ', $2.5 \mu\text{rad}$), and a *low quality* scenario ('LQ', $10.0 \mu\text{rad}$). In order to obtain some typical vertical variation in the random uncertainty profiles, we took a representative noise profile and scaled it such that the average vertical random uncertainty matched the 0.5, 2.5, or $10.0 \mu\text{rad}$ respectively. The estimated random uncertainties of these three scenarios are illustrated relative to α_b in Figure 4.3c (red lines).

The measurement was assumed to be essentially bias free, with systematic uncertainty $u_{\alpha_r}^s(z_a)$ set to $0.1 \mu\text{rad}$ (vertically constant), roughly in line with empirical residual bias estimates related to the ionosphere (Danzer et al. 2013; Danzer et al. 2015; Liu et al. 2013; Liu et al. 2015). The solar flux and thus residual errors were low in July 2008; the latter are therefore well covered by the systematic uncertainties assumed for the test-day ensembles. At days with a high solar and geomagnetic activity, which can induce strong ionospheric perturbations, these exemplary test-day uncertainties might underestimate real effects. The systematic uncertainty relative to α_b is plotted as red line in Figure 4.3b.

Both, estimated random $u_{\alpha_r}^r$ and systematic $u_{\alpha_r}^s$ uncertainties will be propagated

onwards from the L1b processor output as soon as uncertainty propagation for L1a and L1b is fully implemented.

Global bending angle uncertainty fields, derived from ECMWF short-range forecast and analysis fields, and calculated according to Li et al. (2013) for the 15th of July 2008, are used as basis for the uncertainties $u_{\alpha b}^r(z_a)$ and $u_{\alpha b}^s(z_a)$ of the background bending angle profiles. The fields for the 15th of July 2008 are sufficiently close to those of the 8th of August 2008 for the purpose of this study, since temporal variabilities are low (Li et al. 2015) and this performance demonstration requires input data to be *realistic*, but not necessarily *real*. To get the vertical profiles at the desired location, the appropriate horizontal bins of the global fields are selected, and the profiles are interpolated from their standard impact altitude grid z_a^{std} —which is based on the refractivity of the International Standard Atmosphere taken from the Fast Atmospheric Signature Code (FASCODE) (Anderson et al. 1986)—to the target grid z_a of the particular profile.

The estimated relative systematic and random background uncertainties are plotted as blue lines in Figure 4.3b (systematic) and 4.3c (random). Down to about 60 km $u_{\alpha b}^s$ is smaller, below 60 km larger than $u_{\alpha r}^s$. For the random uncertainty, the retrieved exceeds the background uncertainty above about 30 km for LQ, 48 km for SQ, and 60 km for HQ, respectively.

Similar to the background uncertainties the correlation matrices $\mathbf{R}_{\alpha r}$ and $\mathbf{R}_{\alpha b}$ are calculated according to Li et al. (2013) for the 15th of July 2008 and provided with respect to z_a^{std} (defined in Table 4.3). Because correlation lengths and correlation function shapes vary insignificantly along small changes of height (Li et al. 2015), the matrices are not interpolated, but directly taken as applicable on z_a , avoiding sophisticated matrix re-adjustment that is irrelevant in practice.

The correlation matrices are visualized in Figure 4.3d and 4.3e in terms of representative correlation functions, and their correlation lengths $L_{\alpha r}$ and $L_{\alpha b}$ are plotted in Figure 4.3f. All background uncertainty profiles are only shown for the altitude range from 28 to 80 km, which is the domain where they are used in the statistical optimization.

The correlation length of the retrieved bending angle is seen to be relatively constant with height (around 0.5 km), while the correlation length of the background increases with altitude to ~ 6 km near 70 km altitude, in line with Li et al. (2015). Thus, the background error correlation functions become broader with increasing altitude.

The used algorithm estimates the correlation length by searching for the distances downward and upward of the correlation functions' asymmetric peak at which the correlation function has dropped to below $1/e$. The returned correlation length is the arithmetic mean of these two distances. Additionally the correlation length is constrained by the data domain, i.e., the correlation length can never be larger than the distance from the peak to the end of the data at 80 km and 0 km. Consequentially, as visible in Figure 4.3f, the correlation length of the background decreases again above 70 km. The correlation length results in Section 4.4 partly depend on the specific settings used in this estimation scheme.

4.3.3 Data Preparation

In the rOPS L2a operational retrieval, the high altitude initialization (Segment (1) in Figure 4.2) starts with the resolution adjustment to the 'L2a' target resolution and the interpolation of all quantities from grids a_t , z_a and z_a^{std} to target grid z_a of the given occultation event (Item 1.1), followed by the calculation of the covariance matrices \mathbf{C}_{or} and \mathbf{C}_{ab} , using the random uncertainty profiles and the error correlation matrices (Item 1.2). However, as mentioned above, to carry out the tests based on simulated input data we also need to superimpose randomly distributed errors onto the *true* profiles. The probabilistic distributions of these errors are characterized by \mathbf{C}_{or} and \mathbf{C}_{ab} , and therefore we count the resolution adjustment, the interpolation to z_a , the creation of these covariance matrices, and the superimposition of the error profile realization (i.e., Items 1.1, 1.2 and (a) in Figure 4.2) as part of the data preparation, described in this Section.

It can be shown that additional relative errors introduced by the interpolation stay below a threshold of 10^{-4} and, therefore, effects of the interpolation on the uncertainty of the bending angle stay small and are not considered. Instead, like for the bending angle profile itself, the uncertainties u_{or}^s and u_{or}^r are just interpolated to the z_a grid. It is assumed that there are no random or systematic uncertainties associated with the z_a grid that need to be considered. The application of a lowpass filter to adjust all 'L1b' profiles—which have different vertical resolution due to different obliquity of RO event tangent point trajectories—to a common target resolution affects the magnitude of the estimated random uncertainty profile u_{or}^r . This effect is taken into account by applying the corresponding uncertainty propagation. The target resolution of RO profiles ranges from near 200 m in the lower troposphere via about 1 km in the lower stratosphere, to about 1.5 km at higher altitudes, consistent with wave-optics processing in the troposphere (Gorbunov et al. 2004) and geometric-optics processing in the stratosphere (Kursinski et al. 1997).

To avoid a loss of processing speed by the needed series of matrix operations in the uncertainty propagation, all relevant quantities are also downsampled, such that there are 400 vertical levels from 0.2 km and 80 km altitude. This vertical grid with 200 m spacing is still well sufficient to properly resolve the vertical variation of uncertainty estimates.

As described by Li et al. (2013), in the new dynamical statistical optimization, the bending angle is only optimized down to an impact altitude z_{aST} (defined in Table 4.3). In order to ensure a consistent representation of uncertainties (consistent with respect to how the uncertainties of the bending angle propagate through the statistical optimization in the state retrieval), and smooth transitions at z_{aST} , $u_{\text{ab}}^r(z_a)$ is artificially enforced to increase from 40 km downwards to z_{aST} , such that the ratio of the observation to the background uncertainty $u_{\text{or}}^r/u_{\text{ab}}^r$ decreases linearly and reaches 0.1 at z_{aST} (Li et al. 2015, Section 2.3 therein).

Another modification is needed because bending angle profiles of different RO events

reach different maximum impact altitude levels $z_{a\text{TopMeas}}$, which may lie below $z_{a\text{TopMax}}$ as defined in Table 4.3. In order to perform the optimization over a consistent altitude range for all events, the value of the estimated random uncertainty of the measurement at the top of the retrieved profile, i.e. $u_{\alpha_r}^r(z_{a\text{TopMeas}})$, is extended constantly up to $z_{a\text{TopMax}}$ as needed. Since $u_{\alpha_b}^r \ll u_{\alpha_r}^r$ at these altitudes > 70 km, the practical effects of this modification are negligible.

Then the observation and background error covariance matrices needed for the statistical optimization can be calculated (Item 1.2 in Figure 4.2) according to

$$C_{X,ij} = u_{X,i}^r \cdot u_{X,j}^r \cdot R_{X,ij}, \quad (4.5)$$

with $X \in \{\alpha_r, \alpha_b\}$, and using these modified random uncertainties, now both reaching from $z_{a\text{TopMax}}$ to $z_{a\text{ST}}$ for any given occultation event.

The final preparatory step and starting point of the testing algorithm is the superimposition of representative error profiles onto α_r^{true} (Item (a) in Figure 4.2). This way we ensure that the prepared random uncertainty $u_{\alpha_r}^r$ —and hence the error covariance matrix \mathbf{C}_{α_r} —and the noise of the actually used retrieved bending angle profile α_r are consistent. By simulating three different receiver noise scenarios (HQ, SQ, LQ) we also aim to test how higher or lower noise affects the statistical optimization and output quantities along the retrieval chain.

We create error profiles representative for the multivariate normal probability distribution characterized by \mathbf{C}_{α_r} . We add these error realizations to α_r^{true} , and hence create representative ensembles $\alpha_{r,k}$ for the distribution of α_r (JCGM 2008b),

$$\alpha_{r,k} = \alpha_r^{\text{true}} + \mathbf{Q}_{\alpha_r} \cdot r_k = \alpha_r^{\text{true}} + (\mathbf{U}_{\alpha_r} \mathbf{\Lambda}_{\alpha_r}^{1/2}) \cdot r_k. \quad (4.6)$$

Here r_k is a vector of standard normal distributed random numbers, $k = 1, \dots, m$ is the sample index to create m realizations of the vector α_r , and \mathbf{Q}_{α_r} is the Cholesky decomposite of \mathbf{C}_{α_r} . \mathbf{Q}_{α_r} is calculated via a spectral decomposition, using the matrix of eigenvectors \mathbf{U}_{α_r} , and $\mathbf{\Lambda}_{\alpha_r}^{1/2}$, a diagonal matrix filled with the square roots of the eigenvalues of \mathbf{C}_{α_r} .

The higher m , the better the drawn samples will represent the sampled distribution. However, such a representation is only needed for the MC validation of the covariance propagation (the results of which are presented in Section 4.5.1). For the algorithm description in Section 4.4 (and for the tests presented in Section 4.5.2) the sample size is limited to 1 and $\alpha_{r,k=1} =: \alpha_r$, just to have an illustrative profile.

For the MC validation we set $m = 1000$ and also sample the distribution of the background bending angle in the same manner, by using α_b as mean, superimposing an error sample drawn from \mathbf{C}_{α_b} , and applying Equation 4.6.

4.4 Algorithm Sequence and Example Results

In this Section the L2a uncertainty propagation algorithm sequence is introduced starting with α_r and α_b as input and proceeding to α_s , N , p_d and finally T_d . We illustrate the effects of the algorithm on the main uncertainty variables of the MID case, and the SQ scenario. A step-by-step description of the algorithm along the retrieval chain can be found in Appendix B.1.

Along the absolute state profiles of the MID case, we also show the background bending angle α_b for the TRO and POL cases for comparison in Figure 4.3a. While α_b is significantly smaller for POL than for MID, TRO and MID are relatively similar (the same is true for α_r , but not shown).

For each L2a retrieval step, i.e. segments (1), (2), and (3) in Figure 4.2, the results for the principal variables (i.e. the state Y with $Y \in \{\alpha_s, N_r, p_{dr}, T_{dr}\}$, the estimated systematic uncertainty u_Y^s , the estimated random uncertainty u_Y^r , the correlation functions $R_{Y,i}$, the correlation length L_Y , and the observation-to-background weighting ratio r_Y^{obw}) are shown in Figures 4.4 to 4.7. These output quantities are defined in Table 4.2.

4.4.1 High Altitude Initialization

We start with the high altitude initialization, in which the retrieved and the background bending angle, i.e., α_r and α_b , are combined in an statistically optimal way, weighted by the covariances \mathbf{C}_{α_r} and \mathbf{C}_{α_b} , to obtain the statistically optimized bending angle profile α_s (Item 1.5 in Figure 4.2).

We propagate estimated random and systematic uncertainties to additionally obtain the covariance matrix \mathbf{C}_{α_s} (and with it random uncertainty $u_{\alpha_s}^r$ and correlation matrix \mathbf{R}_{α_s}) and systematic uncertainty $u_{\alpha_s}^s$. We also calculate the observation-to-background weighting ratio r_{α}^{obw} .

Figure 4.4 illustrates the relevant quantities. Line 'SQ MID' in Figure 4.4a shows that the superimposed error, visible in the upper part of α_r in Figure 4.3a, has been smoothed out by the statistical optimization. Comparing Figures 4.3c and 4.4c shows how the statistical optimization reduced the random uncertainty over the entire height range. $u_{\alpha_s}^r$ is always smaller than the smaller of the two incoming random uncertainties $u_{\alpha_r}^r$ and $u_{\alpha_b}^r$, also at the intersection (e.g., at about 28 km and 2% for LQ). The black line 'err' stands for the relative superimposed error, i.e., $100 \cdot (\alpha_r - \alpha_r^{\text{true}}) / \alpha_r^{\text{true}}$.

The result for $u_{\alpha_s}^s$ (Figure 4.4b) might seem unexpected, as it increases up to 70 km and decreases afterwards. However, this is the consequence of the inverse covariance weighting in the statistical optimization, causing $u_{\alpha_s}^s$ to be (to first order) a linear combination of $u_{\alpha_r}^s$ and $u_{\alpha_b}^s$, with the random uncertainties as weights. Looking at Figure 4.3b and 4.3c, makes apparent that $u_{\alpha_s}^s$ is equal to $u_{\alpha_r}^s$ and $u_{\alpha_b}^s$ at 60 km, because the latter two have approximately the same value ($\sim 2.3\%$). Since both increase upwards of 60 km, $u_{\alpha_s}^s$ will increase independent of their weighting ratio, compared to its value at 60 km. At

80 km, $u_{\alpha s}^s$ is essentially equal to $u_{\alpha b}^s$, because $u_{\alpha r}^r \gg u_{\alpha b}^r$, and it thus decreases again (to $\sim 3.7\%$).

The Figures 4.4d and 4.4e with correlation functions $R_{\alpha s, i}$ and the correlation length $L_{\alpha s}$, show how the statistical optimization combines $\mathbf{C}_{\alpha r}$ and $\mathbf{C}_{\alpha b}$ to $\mathbf{C}_{\alpha s}$: at high altitudes $\mathbf{C}_{\alpha s}$ replicates $\mathbf{C}_{\alpha b}$ with its long correlation length, then $L_{\alpha s}$ quickly decreases to reach $L_{\alpha r}$ at about 30 km.

In Figure 4.4f the effect of the different $u_{\alpha r}^r$ scenarios on the observation-to-background weighting ratio r_{α}^{obw} is visible. For a α_r with a very low quality, there is a relatively abrupt edge near 30 km, i.e., at the center of the half-sine-weighted transition.

All quantities in Figure 4.4 are given with respect to altitude z , as opposed to z_a , because they have been re-adjusted to the equidistant z grid after the refractivity retrieval.

4.4.2 Refractivity Retrieval

The next step is the refractivity retrieval, where refractivity $N_r(z)$ is derived from $\alpha_s(z_a)$, by solving an Abel integral (Item 2.1 in Figure 4.2). Again we propagate the covariance matrix \mathbf{C}_{N_r} (and thus $u_{N_r}^r$, \mathbf{R}_{N_r}), the systematic uncertainty $u_{N_r}^s$, and the observation-to-background weighting ratio r_N^{obw} .

Figure 4.5 shows the relevant variables for the refractivity retrieval and refractivity uncertainty propagation. Typically the relative uncertainties stay in the same order of magnitude as for the bending angle, but the maxima are slightly smoothed out and shifted downwards (visible in Figure 4.5b and 4.5c). Due to the effects of the integration, the correlation length (Figure 4.5e), and the share of background information at low altitudes (Figure 4.5f) both increase.

4.4.3 Dry-Air Retrieval

In the dry-air state retrieval, the so-called *dry-air variables* are calculated from N_r , using the Smith-Weintraub equation (Kursinski et al. 1997; Smith and Weintraub 1953) (Item 3.1 in Figure 4.2), the hydrostatic integral (Item 3.5) and the equation of state (Item 3.9).

Figure 4.6 shows the results for the dry pressure state retrieval and the uncertainty propagation. The pressure integral further shifts the typical features of u^s and u^r downwards. Due to the hydrostatic integral, the correlation length (in Figure 4.6e) first increases with decreasing height from 80 km downwards, but reaches a maximum of ~ 37 km near 60 km altitude. Below 60 km altitude the data-domain-constrained downward correlation length decreases at roughly the same rate as the upward correlation length increases, and therefore L_{pd} stays roughly constant. r_{pd}^{obw} shows that the background info is now stretching down to below 10 km (for LQ).

Figure 4.7 illustrates the results for the dry temperature. In Figure 4.7a the differences between the three cases, and the effect of the superimposed error are more distinct than

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

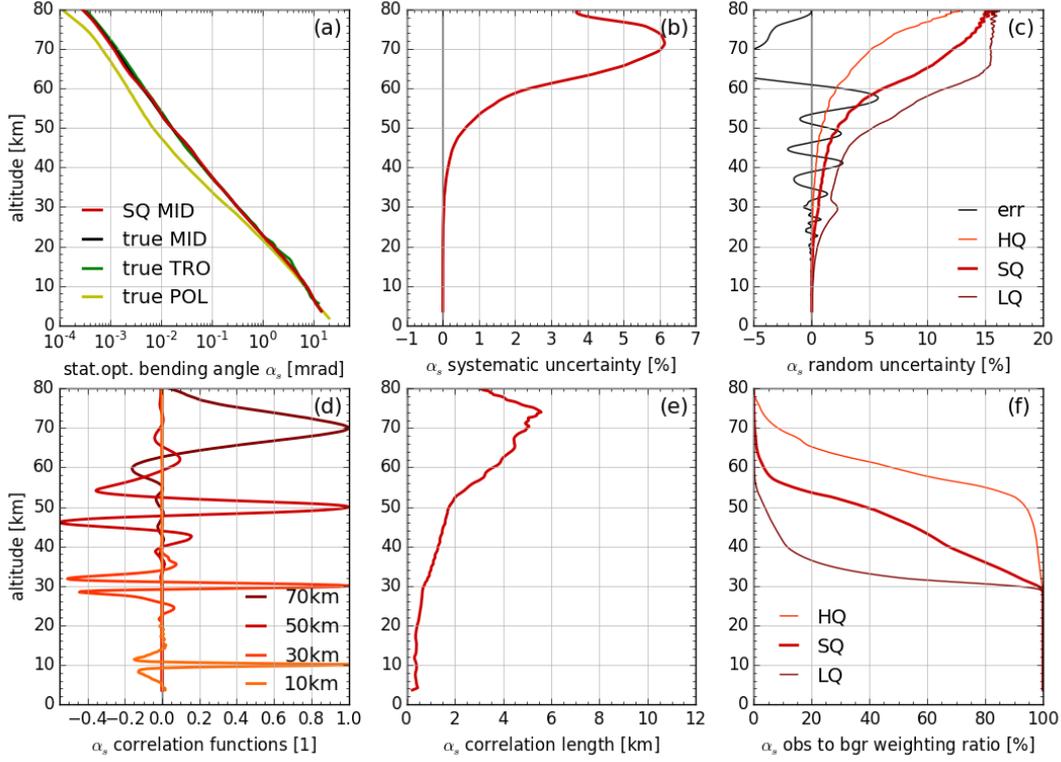


Figure 4.4: Results for statistically optimized bending angle (case MID and scenario SQ, if not stated otherwise): (a) Stat.opt.bend.angle α_s with error ('SQ MID') and without error for all cases ('true MID', 'true POL', 'true TRO'), (b) relative systematic uncertainties $100 \cdot u_{\alpha_s}^s / \alpha_s$, (c) relative random uncertainties $100 \cdot u_{\alpha_s}^r / \alpha_s$ for all scenarios ('HQ', 'SQ', 'LQ') and the relative error $100 \cdot \Delta\alpha_s / \alpha_s^{\text{true}}$ ('err'), (d) correlation functions $R_{\alpha_s, i}$ at 10, 30, 50 and 70 km, (e) correlation length L_{α_s} , and (f) observation-to-background weighting ratio r_{α}^{obw} for all scenarios ('HQ', 'SQ', 'LQ').

for the exponential variables α_s , N_r , and p_{dr} . Due to the division of p_{dr} by ρ_{dr} in the equation of state, their systematic uncertainties partly cancel out at altitudes where they have the same magnitude causing the systematic uncertainty of the dry temperature to be close to zero (at about 61 km in Figure 4.7b). The correlation functions inherit the spiky nature of N_r and the smoothed one from p_{dr} , which dominates towards lower altitudes (Figure 4.7d). Consequentially, the correlation length is relatively short at high altitudes, increases quickly at around 50 km, but then would exceed the range of the data-domain and hence stays roughly constant at low altitudes.

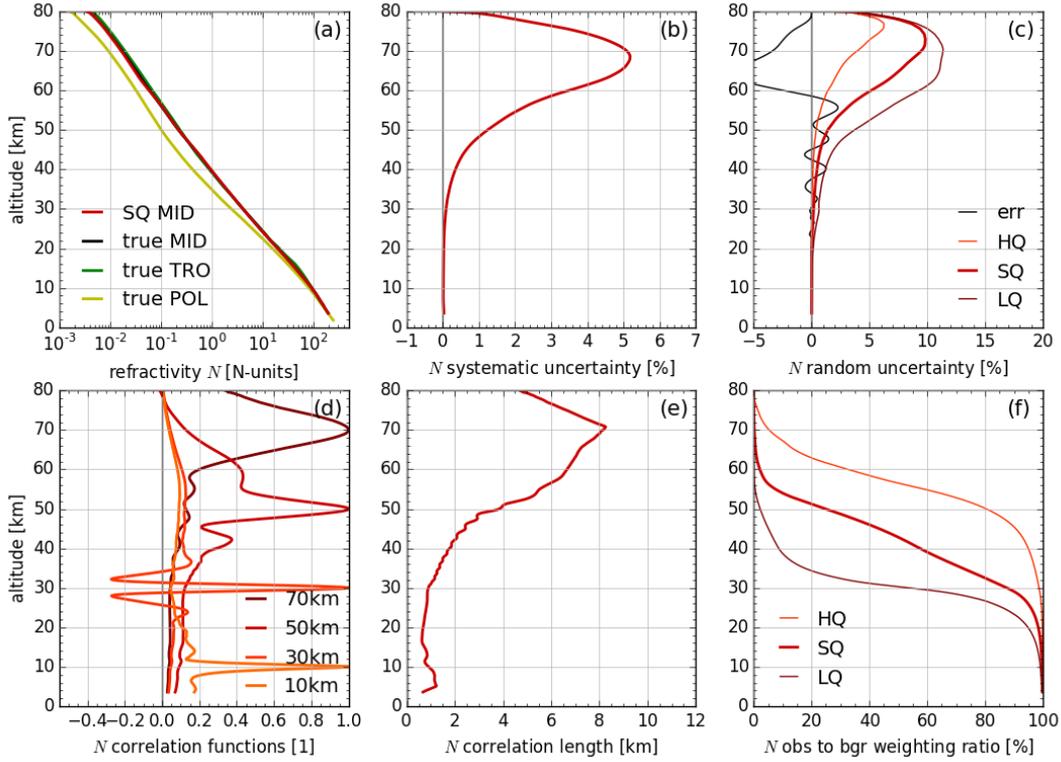


Figure 4.5: Results for retrieved refractivity N_r (case MID and scenario SQ, if not stated otherwise): (a) Refractivity N_r with error ('SQ MID') and without error for all cases ('true MID', 'true POL', 'true TRO'), (b) relative systematic uncertainties $100 \cdot u_{N_r}^s / N_r$, (c) relative random uncertainties $100 \cdot u_{N_r}^r / N_r$ for all scenarios ('HQ', 'SQ', 'LQ') and the relative error $100 \cdot \Delta N_r / N_r^{\text{true}}$ ('err'), (d) correlation functions $R_{N_r, i}$ at 10, 30, 50 and 70 km, (e) correlation length L_{N_r} , and (f) observation-to-background weighting ratio r_N^{obw} for all scenarios ('HQ', 'SQ', 'LQ').

4.5 Performance Demonstration

We demonstrate the correct performance of the implemented algorithm by performing a validation of the covariance propagation (CP) with the help of a Monte Carlo (MC) method (Section 4.5.1), retrieving individual simulated events for different geographic regions (Section 4.5.2), and retrieving test-day ensembles of RO events using real and simulated data (Section 4.5.3).

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

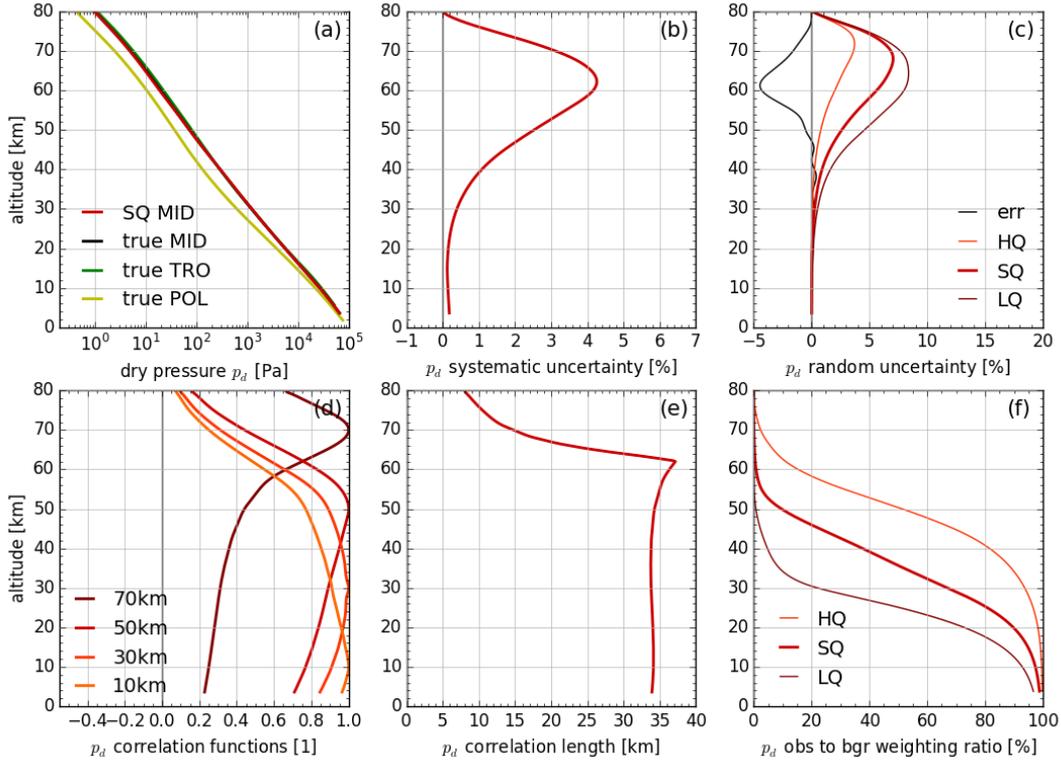


Figure 4.6: Results for retrieved dry pressure p_{dr} (case MID and scenario SQ, if not stated otherwise): (a) Dry pressure p_{dr} with error ('SQ MID') and without error for all cases ('true MID', 'true POL', 'true TRO'), (b) relative systematic uncertainties $100 \cdot u_{p_{dr}}^s/p_{dr}$, (c) relative random uncertainties $100 \cdot u_{p_{dr}}^r/p_{dr}$ for all scenarios ('HQ', 'SQ', 'LQ') and the relative error $100 \cdot \Delta p_{dr}/p_{dr}^{\text{true}}$ ('err'), (d) correlation functions $R_{p_{dr},i}$ at 10, 30, 50 and 70 km, (e) correlation length $L_{p_{dr}}$, and (f) observation-to-background weighting ratio r_{pd}^{obw} for all scenarios ('HQ', 'SQ', 'LQ').

4.5.1 Algorithm Validation Using a Monte Carlo Method

The GUM advises to use a MC method for uncertainty propagation, when the conditions for a GUM-type CP are not valid (e.g., input distributions are not normally distributed), or if their validity is difficult to determine. The crucial disadvantage of using a MC method for uncertainty propagation is its significantly lower processing speed.

As outlined in Section 4.3.1, the conditions for a GUM-type CP are met in our case. Consequentially the MC method is only used to validate the results of the CP, as recommended by JCGM (2011).

For the MC method, a large number m of draws from the distributions of the input

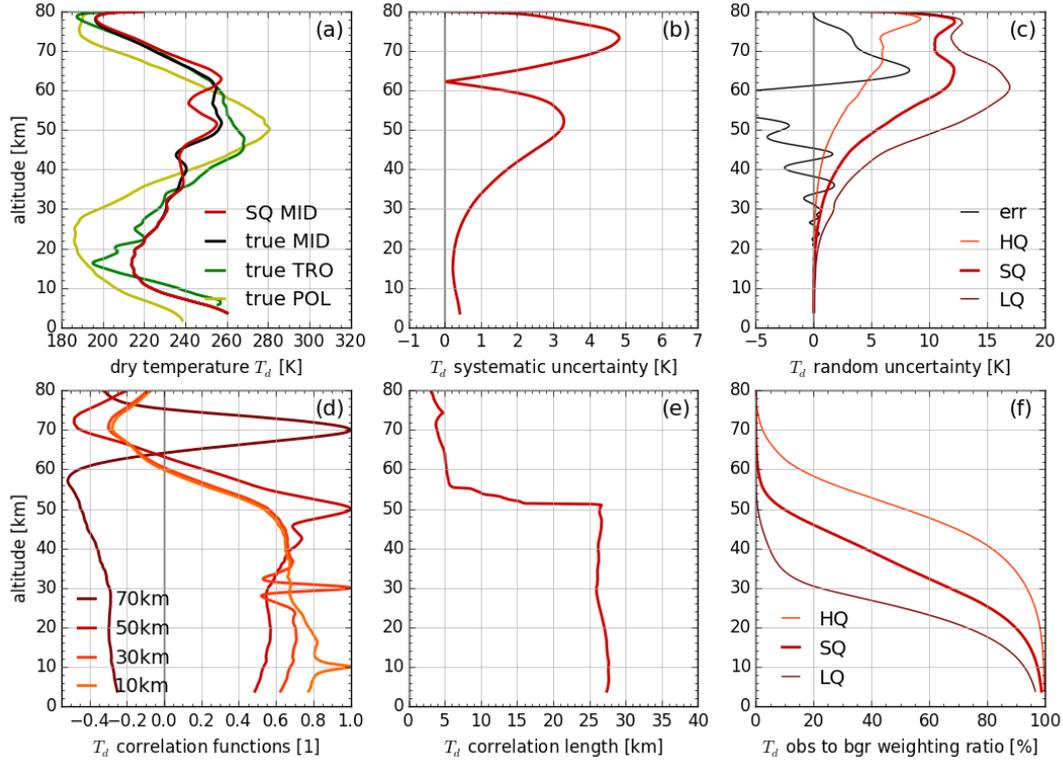


Figure 4.7: Results for retrieved dry temperature T_{dr} (case MID and scenario SQ, if not stated otherwise): (a) Dry temperature T_{dr} with error ('SQ MID') and without error for all cases ('true MID', 'true POL', 'true TRO'), (b) relative systematic uncertainties $100 \cdot u_{T_{dr}}^s/T_{dr}$, (c) relative random uncertainties $100 \cdot u_{T_{dr}}^r/T_{dr}$ for all scenarios ('HQ', 'SQ', 'LQ') and the relative error $100 \cdot \Delta T_{dr}/T_{dr}^{\text{true}}$ ('err'), (d), correlation functions $R_{T_{dr},i}$ at 10, 30, 50 and 70 km, (e) correlation length $L_{T_{dr}}$, and (f) observation-to-background weighting ratio $r_{T_d}^{\text{obw}}$ for all scenarios ('HQ', 'SQ', 'LQ').

variables are taken (described in Section 4.3.3), and for each of these draws the state retrieval is run through, to give m draws of the output variable y_k . For each output variable Y (with $Y \in \{\alpha_s, N_r, p_{dr}, T_{dr}\}$), these sample results y_k are recombined to calculate the distribution of Y (visualized by Items (b), (c), (d) and (e) in Figure 4.2). The distribution of Y is characterized by the mean profile

$$y^{\text{MC}} = \frac{1}{m}(y_1 + \dots + y_m), \quad (4.7)$$

and the covariance matrix

$$\mathbf{C}_Y^{\text{MC}} = \frac{1}{m-1} [(y_1 - y^{\text{MC}})(y_1 - y^{\text{MC}})^T + \dots + (y_m - y^{\text{MC}})(y_m - y^{\text{MC}})^T], \quad (4.8)$$

of the sample results y_k (each having rank 1). \mathbf{C}_Y^{MC} is then compared to the covariance propagated distribution characterized by \mathbf{C}_Y .

Additionally, if all operators are linear,

$$E[f(X)] = f(E[X]) \quad (4.9)$$

holds for a random variable X . Thus, by comparing y^{MC} with y^{true} (gained from a retrieval run using α_r^{true} and α_b with no errors superimposed as inputs) we can test the linearity (and adequacy) of the rOPS retrieval operators.

In order to validate the results we compared the covariance-propagated covariance matrices with the Monte-Carlo-propagated covariance matrices for variables α_s , N_r , p_{dr} and T_{dr} . The comparison is visualized for MID in Figure 4.8 (correlations), and for TRO, MID and POL in Figure 4.9 (random uncertainties).

Figure 4.8 shows that the error correlation matrices for each parameter are in agreement, with the limitation that the correlation matrices produced by the MC method are comparatively noisy. This is due to the finite number of draws from the distribution. Since the purpose of the MC propagation is only the basic validation of the correct implementation of the covariance propagation and is not needed for the uncertainty propagation, the noise is of no further concern.

Figure 4.9 (described in more detail in Section 4.5.2) shows that the covariance-propagated ('rand') and the MC-propagated ('MC') random uncertainties generally match very well for all three cases. There are slight differences at high altitudes for the POL case, visible, e.g., in Figure 4.9l.

4.5.2 Algorithm Performance for Individual Profiles

To verify whether the uncertainty propagation works independently of the event's geographic location, we show results for random and systematic uncertainties for the SQ scenario and all cases of Table 4.4, i.e., for the MID case, but also the tropic (TRO) and the polar (POL) case. For comparison we again calculated the random uncertainties using both, the CP and the MC method.

We start with the same input variables as described in Section 4.3.3 and already shown for the MID case in Section 4.4, i.e., α_r (α_r^{true} with error corresponding to random uncertainty of $2.5 \mu\text{rad}$ superimposed), α_b , $u_{\alpha_r}^r$, $u_{\alpha_b}^r$, R_{α_r} , R_{α_b} , $u_{\alpha_r}^s$, and $u_{\alpha_b}^s$. We retrieve all quantities as described in Section 4.4 and sketched in Figure 4.2, this time for each of the three events.

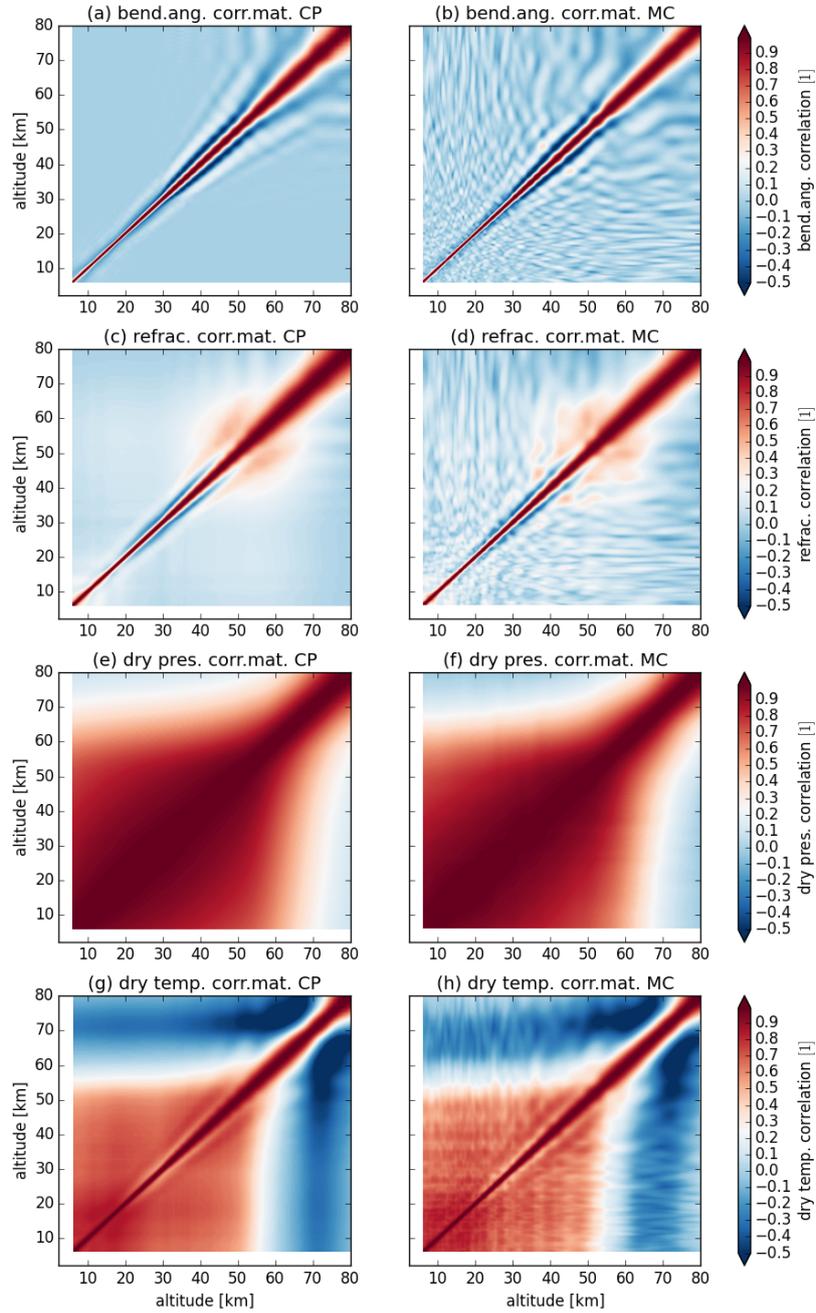


Figure 4.8: Error correlation matrices from CP and MC methods: (a) Covariance propagated R_{α_s} , and (b) Monte-Carlo propagated $R_{\alpha_s}^{MC}$ for statistically optimized bending angle, (c) propagated R_{N_r} and (d) Monte-Carlo $R_{N_r}^{MC}$ for retrieved refractivity, (e) propagated $R_{p_{dr}}$ and (f) Monte-Carlo $R_{p_{dr}}^{MC}$ for retrieved dry pressure, (g) propagated $R_{T_{dr}}$ and (h) Monte-Carlo $R_{T_{dr}}^{MC}$ for retrieved dry temperature.

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

Uncertainties for variables Y are calculated. Again $Y \in \{\alpha_s, N_r, p_{dr}, T_{dr}\}$. We calculate estimated systematic uncertainties u_Y^s ('syst'), estimated random uncertainties u_Y^r by using the CP method ('rand'), and the MC method ('MC'), one individual error realization superimposed ($y - y^{\text{true}}$) ('err'), and the combined uncertainty u_Y^c ('comb'),

$$(u_Y^c)^2 = (u_Y^s)^2 + (u_Y^r)^2. \quad (4.10)$$

In Figure 4.9 these quantities are visualized relative to the *true* state profiles y^{true} . The uncertainties of the dry temperature are plotted in absolute terms.

The vertically constant systematic uncertainty ($u_{\alpha_r}^s = 0.1 \mu\text{rad}$) is exactly (by construction), and $u_{\alpha_r}^r$ essentially the same for all three cases. There are small differences in the vertical variations of $u_{\alpha_r}^r$, but all three random uncertainty profiles are about $2.5 \mu\text{rad}$ in the vertical average.

Hence the main differences between the results of the three cases in Figure 4.9 arise from the different state profiles, i.e., the denominators in the relative uncertainty profiles, and from the background uncertainties, which vary along the geographical latitude. The MID and the TRO case are relatively similar, the POL is rather different from these two (as also visible in panels (a) of Figures 4.3 - 4.7).

In the troposphere $u_{p_{dr}}^s$ and $u_{T_{dr}}^s$ (Figures 4.9j - 4.9o)) are dominated by the contributions of the model error in the hydrostatic equation, and can thus reach 0.25 %, or 0.5 K respectively.

4.5.3 Statistical Performance Evaluation with Real Data

For a broader performance evaluation with real data we processed a test-day of real (CHAMP, COSMIC, MetOp) and simulated (simMetOp) data of GNSS RO satellite missions. Results are shown in Figure 4.10 and Figure 4.11 and compared to results from SK2005 and SP2011.

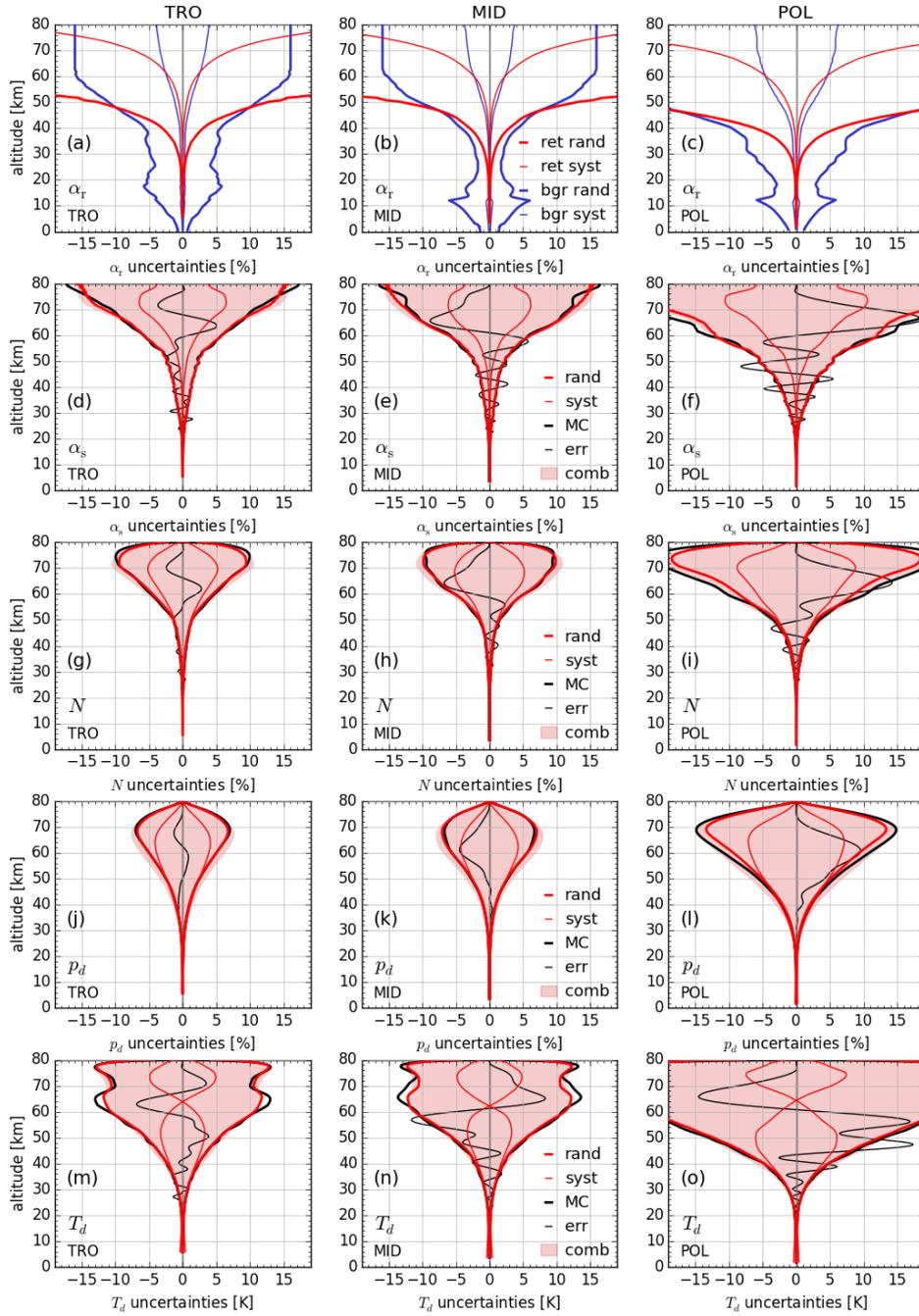


Figure 4.9: Results from uncertainty propagation for cases TRO, MID, POL: The input uncertainties $100 \cdot u_{\alpha_r}^r / \alpha_b$ ('ret rand') and $100 \cdot u_{\alpha_b}^r / \alpha_b$ ('bgr rand'), $100 \cdot u_{\alpha_r}^s / \alpha_b$ ('ret syst') and $100 \cdot u_{\alpha_b}^s / \alpha_b$ ('bgr syst') for cases MID (a), POL (b), and TRO (c), and the results for the random ('rand'), systematic ('syst'), and combined ('comb') uncertainties, the Monte-Carlo validation of the random uncertainties ('MC'), and one example error realization ('err') relative to the state profile for α_s (d, e, f), N_r (g, h, i), p_{dr} (j, k, l), and in absolute terms for T_{dr} (m, n, o). 107

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

Figure 4.10 shows profiles for the relative random ($100 \cdot u_{\alpha_s}^r/\alpha_s$) and systematic uncertainties ($100 \cdot u_{\alpha_s}^s/\alpha_s$) (left panels) and for r_{α}^{obw} (right panels) with respect to altitude. There is no *true* profile available for these real cases, so the retrieved state profiles are used as denominator. In order to get a sense for the spread of the single members of the test-day ensemble around their means, all single members (grey) and the means (color) are plotted, but some outliers ($< 5\%$ of the total number of profiles for each mission) have been disregarded.

For all ~ 100 CHAMP events, $100 \cdot u_{\alpha_s}^r/\alpha_s$ stays below 15% at altitudes < 50 km, and below 7% at altitudes < 40 km (Figure 4.10a). This is consistent with the findings of SK2005 and SP2011 down to 10 km (below which the random error strongly increases in the findings of these authors, due to representativity errors, as discussed below). As visible in Figure 4.10b, at 40 km the share of the observation in the optimized bending angle can range from as little as 3% for low- and up to 70% for high-quality measurements. Events with a low r^{obw} at 30 km exhibit a sudden increase of the random uncertainty at the transition altitude, because starting at 28 km the still relatively large $u_{\alpha_r}^r$ is weighted at 100%. Ongoing fine tuning of the statistical optimization will avoid this broadening in the future. $z_{\text{obw}50}$, the altitude at which r^{obw} of the mean is 50%, is about 33 km.

For COSMIC, Figure 4.10c, the mean of the relative estimated random uncertainties $100 \cdot u_{\alpha_s}^r/\alpha_s$ is smaller than for CHAMP over the entire altitude range. Due to the larger sample size however, some of the 1500 single members of the test-day ensemble show larger random uncertainty than those of the CHAMP ensemble. The range of the ensemble's r^{obw} profiles, visible in Figure 4.10d, is thus also larger for COSMIC.

Compared to CHAMP, measurement information prevails over background information already at higher altitudes on average, as $z_{\text{obw}50}$ is about 44 km for COSMIC.

Compared to the two other receivers, $u_{\alpha_r}^r$ of MetOp is significantly smaller and is comparable to $u_{\alpha_b}^r$ from around 60 to 40 km. $z_{\text{obw}50}$ consequentially also lies higher than for CHAMP and COSMIC, close to 50 km, as visible in Figure 4.10f. At 40 km the mean r^{obw} is already close to 85%, compared to about 60% for COSMIC. The real MetOp profiles still have a higher $u_{\alpha_r}^r$ than simMetOp, which is visible when comparing Figures 4.10e and 4.10g, particularly below 60 km, where $u_{\alpha_s}^r$ is essentially equal to $u_{\alpha_r}^r$.

For simMetOp, $u_{\alpha_r}^r$ is even smaller than for MetOp and roughly equal to $u_{\alpha_b}^r$ between 60 km and 70 km. Below 60 km the random uncertainty of the statistically optimized bending angle is thus already almost equal to the small measurement uncertainties and stays below 3% at altitudes < 50 km (for all ~ 700 profiles with one exception), and below 1.5% at altitudes < 40 km (Figure 4.10g). Consequentially, at 40 km the observation fully dominates the bending angle profile, with the mean r^{obw} at above 95%, as visible in Figure 4.10h. For simMetOp $z_{\text{obw}50}$ is about 66 km.

At high altitudes, the systematic uncertainty of simMetOp is larger than for COSMIC and CHAMP. This somewhat surprising result is caused by the effects of the statistical optimization: at > 60 km, $u_{\alpha_b}^s$ is generally smaller than $u_{\alpha_r}^s$ (depends somewhat on the geographic location of the event), as can be seen e.g., for the MID case in Figure 4.3b.

In the statistical optimization the ratio of the random uncertainties $u_{\alpha_r}^r$ to $u_{\alpha_b}^r$ influences the weighting. For simMetOp $u_{\alpha_r}^r$ is smaller than $u_{\alpha_b}^r$ above 60 km. This leads to an increased weighting of the bad $u_{\alpha_r}^s$ compared to the smaller $u_{\alpha_b}^s$ at high altitudes and therefore $u_{\alpha_s}^s$ is larger for simMetOp than for the other receivers.

The Figures 4.11a and 4.11b show the means of the bundles in Figure 4.10. The other panels in this figure show how these means propagate from one parameter to the next i.e., from α_s to N_r , further to p_{dr} , and finally to T_{dr} . Figures 4.11c, 4.11e, and 4.11g show that the relative random and systematic uncertainties keep their order of magnitude from one step to the next. The integral operators from α_s to N_r and from N_r to p_{dr} shift the maxima downwards.

Above 10 km, the results are in good agreement with the results of SK2005 and SP2011. The random uncertainties of SK2005 generally lie between ours for COSMIC and CHAMP, uncertainties of SP2011 are only slightly larger than our uncertainties for CHAMP. However, below 10 km the random uncertainties increase in SK2005 and SP2011, while ours decrease.

The most important reason for these differences lies in our assumed input uncertainties $u_{\alpha_r}^r$. We set them essentially vertically constant in this introductory study, but it is expected that uncertainties increase in the troposphere. Using bending angle uncertainties from the uncertainty propagation through the wave optics retrieval of the bending angle, as demonstrated by Gorbunov and Kirchengast (2015), is expected to help settle this issue. A second reason seems to be the large differences in the troposphere between the ECMWF and RO profiles found by SK2005 and SP2011. This difference might arise from slant RO events, which cause the RO samples to be relatively far off the given mean tangent-point (at which the vertical ECMWF profiles have been extracted) in the troposphere. Hence the compared profiles represent atmospheric states at very different locations. By contrast, our uncertainties are given with respect to the event trajectory.

The exact reasons for the differences to SK2005 and SP2011 will need further refined analysis, and simplifications in this implementation of the uncertainty propagation do not allow for an analysis now. However, using the completed rOPS system modeling approach, it will be possible to carefully quantify effects like these in the future (Kirchengast et al. 2016a; Kirchengast et al. 2016b).

Despite advantages from separating uncertainties according to random and systematic effects, some types of errors pose a challenge to this division. For example, the model errors, which can be random or systematic, and are counted as systematic uncertainties in our study, can be of significant size and might also contribute to the large random errors in SK2005 and SP2011.

The observation-to-background weighting ratio r^{obw} , plotted in Figures 4.11b, 4.11d, 4.11f and 4.11h, demonstrates how the background information is carried downwards towards lower altitudes from step to step along the retrieval chain, due to the integral operators.

4 Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

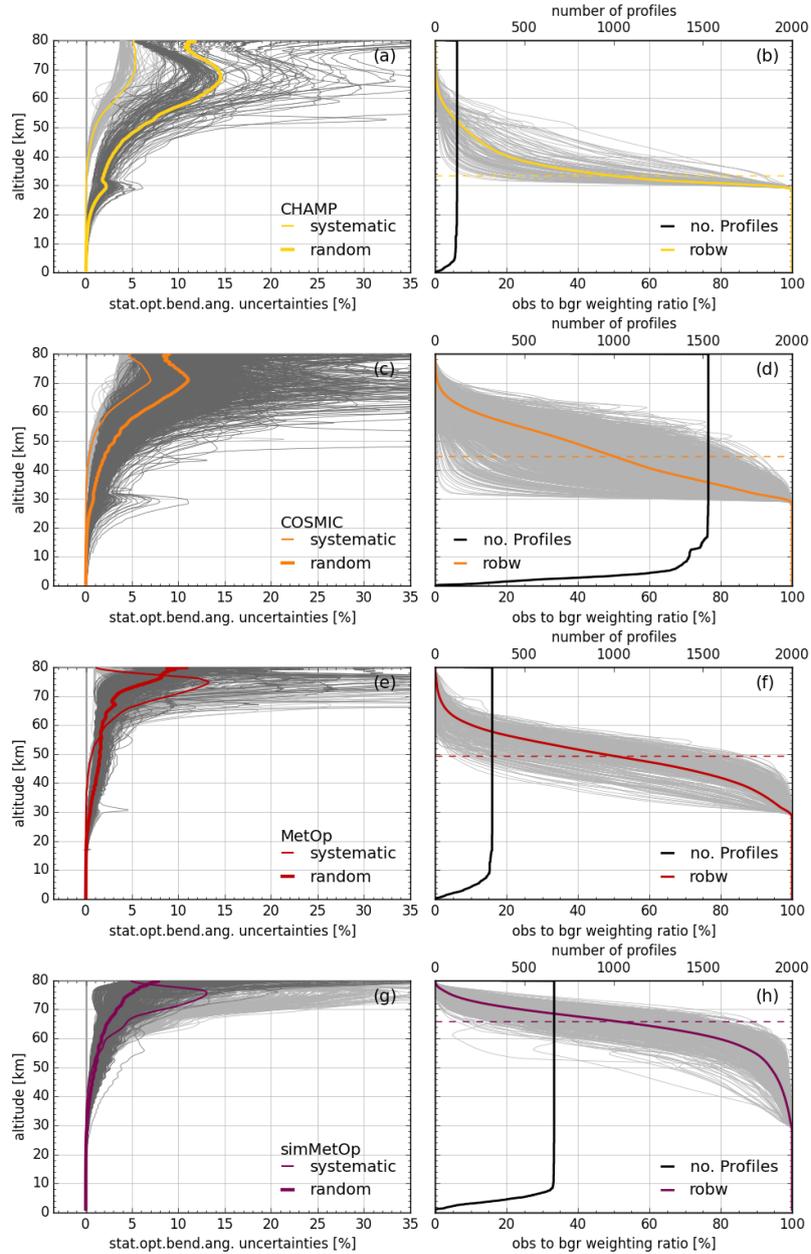


Figure 4.10: Results from uncertainty propagation for ensembles with real measurement data from July 15th 2008, for the statistically optimized bending angle α_s : Relative random $100 \cdot u_{\alpha_s}^r / \alpha_s$ ('random') and systematic $100 \cdot u_{\alpha_s}^s / \alpha_s$ ('systematic') uncertainties of each single event (gray), and their means (color) for CHAMP (a), COSMIC (c), MetOp (e) and simMetOp (g). Observation-to-background weighting ratio r_{α}^{obw} ('robw') of each individual event (gray) and their mean (color), the altitude $z_{\text{obw}50}$ (dashed color), and the number of events in the ensemble ('no. Profiles') for CHAMP (b), COSMIC (d), MetOp (f), and simMetOp (h).

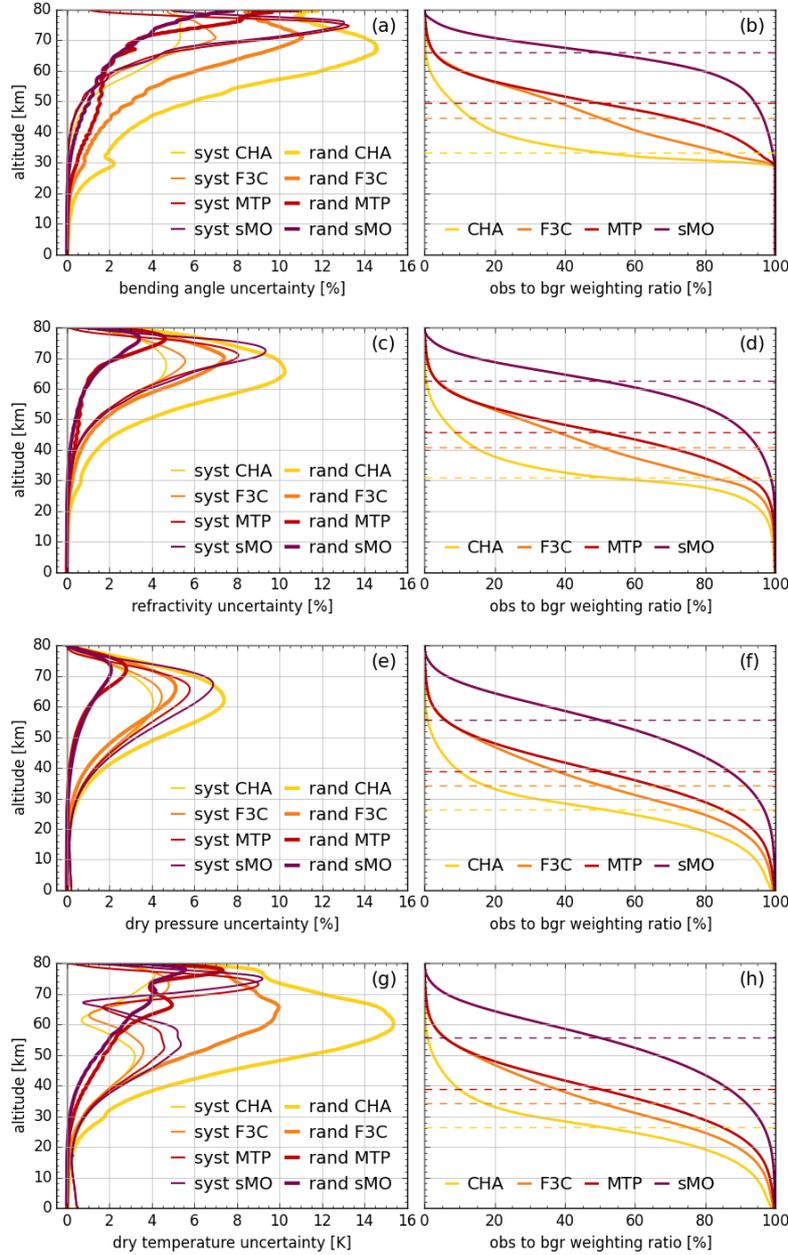


Figure 4.11: Results from uncertainty propagation for ensembles with real measurement data from July 15th 2008, for the principal output variables and satellites CHAMP ('CHA'), COSMIC ('F3C'), MetOp ('MTP'), and simMetOp ('sMO'): Means of relative random $100 \cdot u_{Y_r}^r / Y_r$ ('rand') and systematic $100 \cdot u_{Y_r}^s / Y_r$ ('syst') uncertainties in the left panels (for α_s in (a), for N_r in (c), for p_{dr} in (e) and absolute uncertainties in same notation for T_{dr} in (g)). Observation-to-background weighting ratio r_Y^{obw} and the altitude $z_{\text{obw}50}$ (dashed), in the right panels (α_s (b), N_r (d), p_{dr} (f), T_{dr} (h)).

4.6 Conclusions

Integrating uncertainty propagations in RO retrievals is essential for delivering climate benchmark datasets. We presented the uncertainty propagation algorithm from bending angle profiles via refractivity profiles to dry-air profiles ('L2a' processing), which we implemented in the rOPS at the WEGC. We validated the algorithm using a MC method and demonstrated the reliable performance of the uncertainty estimation for individual profiles of simulated events and for ensembles of real events.

Unique amongst all existing RO processing systems so far (Ho et al. 2012; Steiner et al. 2013), we provide estimates both of systematic and random uncertainty profiles, of error correlation matrices, and of observation-to-background weighting ratio profiles that make the influence of background information from high-altitude initialization transparent. We show that the amount of background information can vary considerably between individual profiles and strongly depends on the profile's observational uncertainty.

We find good agreement with authors of previous empirical error estimation studies above about 10 km, but remaining differences below 10 km will require the inclusion of wave-optics uncertainty propagation in the 'L1b' bending angle retrieval, which is currently on-going.

As next steps we will introduce the integrated uncertainty propagation for the L1b geometric-optics retrieval from excess phase profiles to bending angle profiles as well as the uncertainty propagation for the moist air retrieval (L2b) and for the orbit determination and excess phase processing (L1a). In this way the complete rOPS retrieval chain will function with integrated uncertainty estimation, and can be used for a new way of RO processing for the benefit of atmospheric and climate science and applications.

4.7 Acknowledgments

We thank UCAR/CDAAC Boulder for access to their RO excess phase and orbit data (available at <http://cdaac-www.cosmic.ucar.edu/>) as well as ECMWF Reading for access to their analysis and forecast data (available at <http://www.ecmwf.int/en/forecasts/datasets>). To access the relevant result files of the uncertainty propagation, please contact the corresponding author. The work was funded by the Austrian Aeronautics and Space Agency of the Austrian Research Promotion Agency (FFG-ALR) under projects OPSCLIMPROP and OPSCLIMTRACE and by the European Space Agency (ESA) under project MMValRO-E.

Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis

This chapter presents on-going work to be submitted to Remote Sensing of Environment (RSE) by Schwarz, Kirchengast, Schwärz, Proschek and Li (2018). The work in the context of this publication included design and implementation of uncertainty propagation in the rOPS retrieval, starting from dry-air variables and background information on temperature and humidity, through the direct temperature and humidity retrievals, the optimal estimation of temperature and humidity, and finally derivation of further thermodynamic variables.

My contribution to this study included the analytical derivation of the uncertainty propagation formulas, the conversion into numerical propagation algorithms, the implementation of the algorithms into the rOPS and the drafting of the manuscript for the publication. The work related to testing and validating the algorithms is in completion.

5.1 Abstract

The highly accurate, long-term stable data records from Global Navigation Satellite System (GNSS) radio occultation (RO) observations, are globally available as a continuous record since 2001 and provide Essential Climate Variables (ECVs) for the thermodynamic state of the free atmosphere, such as temperature and tropospheric water vapor profiles (involving background information). These data records would thus have the potential to serve as climate benchmark data, but in order to exploit this potential, atmospheric profile retrievals need to be very accurate and the remaining uncertainties quantified and traced throughout the retrieval chain.

The new Reference Occultation Processing System (rOPS) at the Wegener Center aims to deliver such an accurate retrieval chain with integrated uncertainty propagation. In this study we introduce and demonstrate the algorithms implemented for uncertainty propagation from dry-air to moist-air atmospheric profiles, for estimated systematic and random uncertainties, including error correlations, resolution profiles, and observation-to-background weighting profiles. We propagate random and systematic uncertainty profiles using linearized versions of the operators used for the basic profiles retrieval. The random uncertainty propagation was traced by a covariance propagation, validated using Monte-Carlo ensemble methods. We demonstrate the algorithm performance using test-day ensembles of real RO events.

Upon completion we will present the results of the Monte-Carlo validation, to show that our covariance propagation delivers correct results, and demonstrate that the new uncertainty estimation chain performs robustly, using the real event ensembles.

5.2 Introduction

With the help of background information, Essential Climate Variables (ECVs) like temperature, pressure or humidity in the free atmosphere can be derived from Global Navigation Satellite System (GNSS) Radio Occultation (RO) measurements, in an RO retrieval.

In order for GNSS RO retrieved ECV records to serve as climate benchmark data record, it must be ensured that the accuracy of the ECV record can be traced back to the SI-tied accuracy of the raw data. This in turn requires that 1. the RO retrieval is highly accurate and no undue amplification of uncertainties occurs along the retrieval chain occurs and that 2. the uncertainties (as the expression of the claimed accuracy) are propagated through the entire retrieval chain, from the raw data to ECVs. The Reference Occultation Processing System (rOPS) project (Kirchengast et al. 2017b) at the Wegener Center of the University of Graz (WEGC), aims to establish such a fully traceable processing for the first time (Kirchengast et al. 2016a; Kirchengast et al. 2016b).

The basic steps of an RO retrieval chain are outlined in Figure 5.1. These consist, first, of the precise orbit determination (POD) and excess phase processing (labeled L1a in Figure 5.1), second of the atmospheric bending angle retrieval (L1b), third, the refractivity and dry-air retrieval (L2a) and finally the moist air retrieval (L2b), to calculate thermodynamic ECVs of the free atmosphere, particularly temperature, pressure and humidity. Multiple studies described this retrieval chain in detail (e.g., Anthes 2011; Hajj et al. 2002; Kursinski et al. 1997; Steiner et al. 2011) and have shown that the accuracy of the RO ECVs, particularly in the upper troposphere and lower stratosphere (UTLS) region is high (e.g., Gobiet et al. 2007; Ho et al. 2012; Kuo et al. 2004; Kursinski et al. 1997; Scherllin-Pirscher et al. 2011a; Scherllin-Pirscher et al. 2011b; Steiner and Kirchengast 2005; Steiner et al. 2013).

The uncertainty propagation in the rOPS will eventually provide a traceable processing along the entire retrieval chain from the raw measurement data to the atmospheric ECVs, but this chain is implemented in sequential blocks. The L1a uncertainty propagation from raw data to excess phase profiles is part of on-going work and the associated manuscript is in preparation (Innerkofler et al. 2017, manuscript in preparation for Atmospheric Measurement Techniques; see also Innerkofler et al. (2016)). The L1b uncertainty propagation from excess phase to atmospheric bending angle profiles has already been introduced by Schwarz et al. (2018a) [SKS2018 hereafter], and the L2a propagation from atmospheric bending angle to dry-air profiles by Schwarz et al. (2017) [SKS2017 hereafter]. Uncertainty propagation for the WO bending angle retrieval has been demonstrated for simulated and real events by Gorbunov and Kirchengast (2015) and Gorbunov and Kirchengast (2018), but completion in the rOPS is still on-going.

This study is a direct continuation of the work presented by SKS2018 and SKS2017, and propagation and validation methods applied are as in these previous studies. The rOPS L2b processor, the moist-air retrieval uses the dry-air parameters of the L2a processor and background information on temperature and humidity serve as input, from which, first, in the so-called *direct method*, temperature with humidity prescribed (labelled as (1a) in Figure 5.1) and then humidity with temperature prescribed (1b), are retrieved. The retrieved temperature and humidity profiles are then used together with the background information to derive temperature and humidity profiles in an *optimal estimation* (2). Finally, complementary thermodynamic variables are derived from these optimally estimated temperature and humidity profiles (3). These three main steps of the L2b processor, shown in Figure 5.1, are expanded in Figure 5.2. The L2b uncertainty propagation also follows this chain-like architecture of the L2b processor, using propagated L2a output uncertainties and background uncertainties as input.

The new rOPS moist-air algorithm in its basic form has already been introduced by Li et al. (2018), including the state retrieval and a simplified, variance-based, random uncertainty propagation. The general retrieval accuracy has been demonstrated by comparison to retrieval results of other RO data centers, particularly the COSMIC Data Analysis and Archive Center (CDAAC) from the University Corporation for Atmospheric Research (UCAR), Boulder, and the Radio Occultation Meteorology Satellite Application Facilities (ROM-SAF) from the Danish Meteorological Institute, Copenhagen (Li et al. 2018). This paper advances this basic uncertainty propagation by the following elements: First, as in SKS2018 and SKS2017, covariance propagation is used to propagate *estimated random uncertainties*, and Monte-Carlo ensemble methods (MC) are used to validate these results. Figure 5.2 illustrates the propagation steps of the covariance propagation alongside the basic state retrieval, and also indicates how the MC validation is executed. The covariance propagation also allowed to implement a full covariance weighted (rather than just variance weighted) optimal estimation.

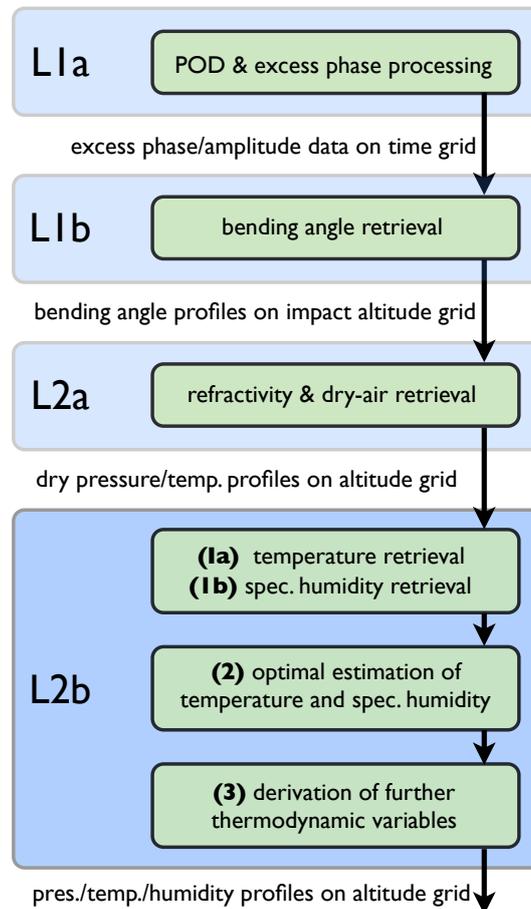


Figure 5.1: Schematic view of the main processors of the retrieval chain in the rOPS (L1a, L1b, L2a, L2b) and the main operators of the L2b processor (1a, 1b, 2, 3), which are in the focus of this study.

Secondly, as in the L2a and L1b processors, we also propagate *estimated basic and apparent systematic uncertainties* independent from the random uncertainties, and additionally provide estimates for *correlation length* profiles, *resolution* profiles, and *observation-to-background weighting* profiles.

The flow of the systematic uncertainties through the moist-air retrieval is not shown in Figure 5.2, because it largely follows the retrieval chain for the state retrieval. A complete (and detailed) description of the implemented uncertainty propagation algorithms—the main innovation of this study—can be found in C.1.

This paper is structured into the following sections. The applied uncertainty estimation, propagation and validation methods, and the data sources and data preparation are

presented in Section 5.3. Section 5.4 introduces the sequence of steps of the uncertainty propagation through the rOPS L2b processor, illustrating this by the retrieval steps for an exemplary real RO event. The results from the MC validation of the propagation algorithm are presented in Section 5.5. The performance of the algorithm is demonstrated using test-day ensembles of real data from the RO missions CHALLENGING Minisatellite Payload (CHAMP) (Wickert et al. 2001) and FORMOSAT-3 Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) (Anthes et al. 2008), and of simulated data approximating characteristics of the Meteorological Operational Satellite A (Luntama et al. 2008) [simMetOp data hereafter]. The respective results are presented and discussed in Section 5.6, Section 5.7 provides a summary and conclusions. The L2b uncertainty propagation algorithms are introduced in a detailed, step-by-step description in C.1.

5.3 Methods and Data

5.3.1 Uncertainty Propagation Methods

We follow the *Guide to the Expression of Uncertainty in Measurement* (JCGM 2008a; JCGM 2008b; JCGM 2011) [GUM hereafter], and aim to follow terminology as provided by the International Vocabulary of Metrology (JCGM 2012), a terminology also adopted by the GUM.

SKS2017 provides an introduction to the methods used in the rOPS uncertainty framework, including the motivation for using the respective uncertainty estimation, propagation and validation methods.

We categorize uncertainties into *estimated random uncertainties* and *estimated systematic uncertainties*. Effects of unpredictable or stochastic temporal and spatial variations in repeated observations can be estimated based on a single RO event—due to oversampling in the RO raw data—and are included in the estimated random uncertainties.

Systematic effects, which cannot be quantified using statistical data analysis based on just one single RO profile, are corrected when known, as recommended by the GUM. The uncertainty of the corrections and potentially remaining residual systematic effects are estimated using GUM Type B uncertainty methods, and added to the estimated systematic uncertainty. These residual systematic uncertainties can either be event-by-event systematic, or event-by-event random. In the latter case these systematic uncertainties will appear as random uncertainties in larger ensembles of RO events (as e.g., in climatologies). We term systematic uncertainty which are fundamentally systematic across different RO events, *estimated basic systematic uncertainties*, and those which appear systematic just for individual RO events *estimated apparent systematic uncertainties*.

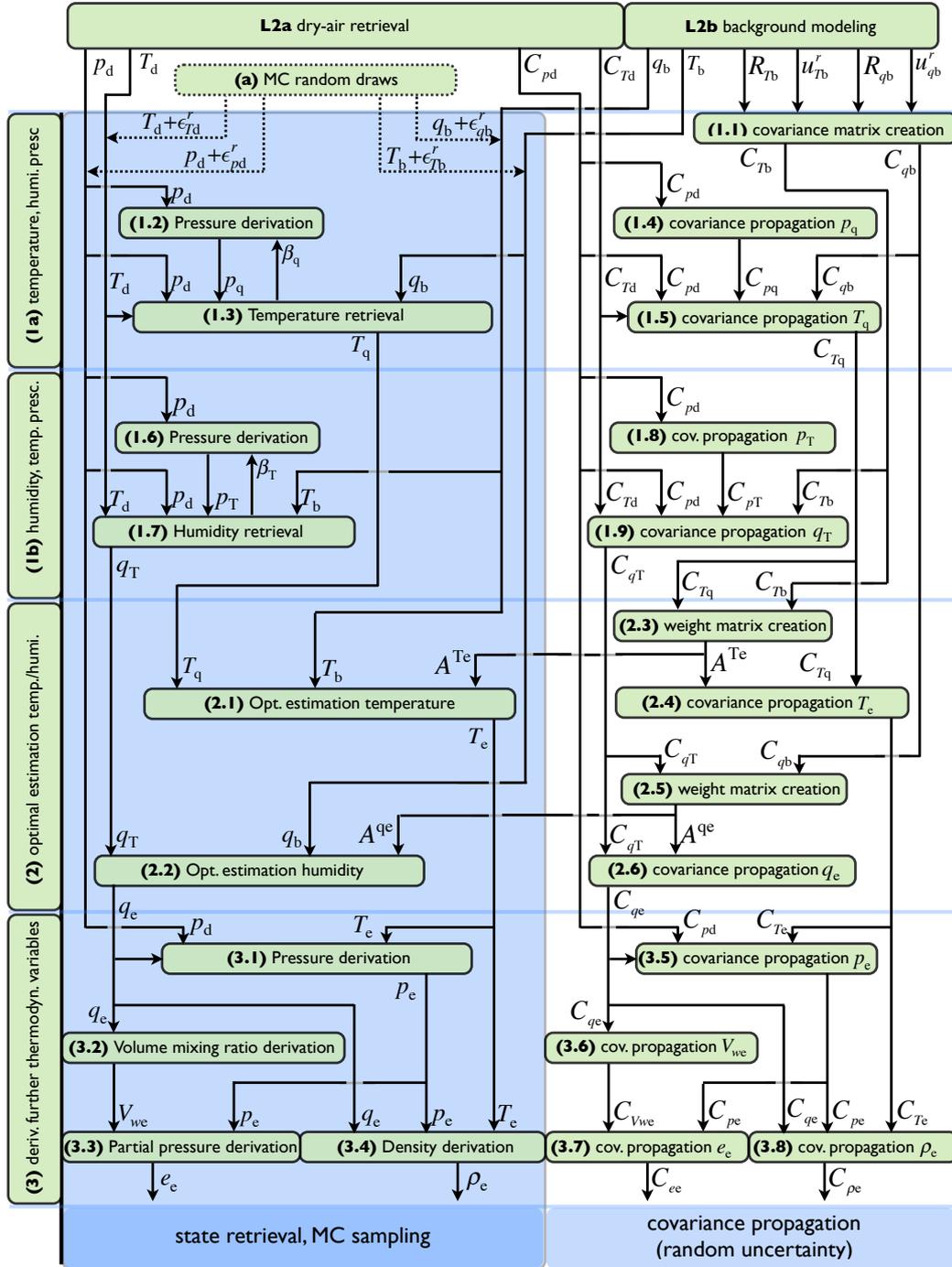


Figure 5.2: Detailed workflow for state retrieval, uncertainty propagation and propagation of the observation-to-background weighting ratio for the main L2b operators from dry-air to moist-air atmospheric profiles. The mathematical notation, including all symbols, is introduced in Table 5.1.

Table 5.1: Principal variables for the L2b state retrieval and uncertainty propagation.

Variable	Unit	Description
X_d	U	state profile of retrieved dry pressure/dry temperature, retrieved in the dry-air retrieval and output of the L2a processor, input profiles of the L2b moist-air retrieval with $X_d \in \{p_d(z), T_d(z)\}$, and unit $U \in \{\text{Pa}, \text{K}\}$, comprising elements $X_{d,i}$.
X_b	U	state profile of background temperature/specific humidity, derived from ECMWF short-range forecast fields, input profiles of the L2b moist-air retrieval with $X_b \in \{T_b(z), q_b(z)\}$, and unit $U \in \{\text{T}, \text{kg/kg}\}$, comprising elements $X_{b,i}$.
X_q	U	state profile of retrieved pressure/temperature with specific humidity prescribed, with $X_q \in \{p_q(z), T_q(z)\}$, and unit $U \in \{\text{Pa}, \text{K}\}$, comprising elements $X_{q,i}$.
X_T	U	state profile of retrieved pressure/specific humidity with temperature prescribed, with $X_T \in \{p_T(z), q_T(z)\}$, and unit $U \in \{\text{Pa}, \text{kg/kg}\}$, comprising elements $X_{T,i}$.
X_e	U	state profile of optimally estimated temperature/specific humidity/water vapor volume mixing ratio/-pressure/partial pressure of water vapor/density, retrieved in the L2b moist-air retrieval, with $X_e \in \{T_e(z), q_e(z), V_{we}(z), p_e(z), e_e(z), \rho_e(z)\}$, and unit $U \in \{\text{K}, \text{kg/kg}, 1, \text{Pa}, \text{Pa}, \text{kg/m}^3\}$, comprising elements $X_{e,i}$.
u_X^s	U	estimated systematic uncertainty profile of X (with X and U as defined above), comprising elements $u_{X,i}^s$.
u_X^r	U	estimated random uncertainty profile of X (with X and U as defined above), comprising elements $u_{X,i}^r$.
\mathbf{R}_X	1	error correlation matrix of X (see above), with elements $R_{X,ij}$.
\mathbf{C}_X	U^2	error covariance matrix of X (with X and U as defined above), comprising elements $C_{X,ij} = u_{X,i}^r \cdot u_{X,j}^r \cdot R_{X,ij}$.
l_X	m	correlation length profile of X (with X as defined above), comprising elements $l_{X,i}$.
w_X	m	resolution profile of X (with X as defined above), in altitude domain comprising elements $w_{X,i}$.

In RO, the prime variables in use are quasi-vertical profiles, which are numerically expressed in the rOPS by rank-1 vectors of N elements. We call these variables *state*

variables. Along the retrieval chain, *explicit, linear, multivariate* operators, $\mathbf{A}^{\mathbf{xy}}$, transform vectors of *correlated input* variables \mathbf{X} of length N (input profiles) to vectors of correlated output variables \mathbf{Y} of length N (output profiles), in generally linear transformations

$$\mathbf{y} = \mathbf{A}^{\mathbf{xy}} \mathbf{x}, \quad (5.1)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_i, \dots, x_N)^T$ and $\mathbf{y} = (y_1, y_2, \dots, y_i, \dots, y_N)^T$ are estimates (or realizations) of the input and output random variables \mathbf{X} and \mathbf{Y} .

According to the GUM, uncertainty propagation through these linear operators can be done using

$$\mathbf{C}_y = \mathbf{A}^{\mathbf{xy}} \cdot \mathbf{C}_x \cdot (\mathbf{A}^{\mathbf{xy}})^T, \quad (5.2)$$

when the input uncertainties are normally distributed, which—as argued in SKS2018 and SKS2017—is true for the estimated random uncertainty as propagated from the L1b and L2a processors. Random uncertainty estimates of the input profile \mathbf{x} are expressed by the covariance matrix, \mathbf{C}_x , those of output profile \mathbf{y} by covariance matrix \mathbf{C}_y .

Different to the L1b and L2a processors, many operators $\mathbf{J}^{\mathbf{xy}}$ of the L2b processor (i.e., the Items 1.2, 1.3, 1.6, 1.7, 3.1, 3.2, 3.3, 3.4 in Figure 5.2) do not qualify as *linear* measurement models as defined in the GUM, being of the form

$$\mathbf{y} = \mathbf{f}^{\mathbf{xy}}(\mathbf{x}), \quad (5.3)$$

but can be linearized by Taylor expansion to first order—at least for the purpose of the uncertainty propagation—and thus approximated by

$$d\mathbf{y} := \mathbf{y} - \mathbf{y}_0 \approx \mathbf{J}^{\mathbf{xy}} \cdot (\mathbf{x} - \mathbf{x}_0) = \mathbf{J}^{\mathbf{xy}} \cdot d\mathbf{x}, \quad (5.4)$$

where small variations $d\mathbf{x}$ of the input quantity \mathbf{x} around a zero-order model state \mathbf{x}_0 transfer to small variations $d\mathbf{y}$ of the output quantity \mathbf{y} around the zero-order model state \mathbf{y}_0 . Here $\mathbf{J}^{\mathbf{xy}}$ corresponds to the Jacobi-Matrix composed of elements

$$J_{ij}^{\mathbf{xy}} = \frac{\partial y_i}{\partial x_j}. \quad (5.5)$$

In these cases, Equation 5.2 sufficiently approximates the uncertainty propagation through the non-linear operator $\mathbf{f}^{\mathbf{xy}}(\mathbf{x})$ in the neighborhood of $\mathbf{f}^{\mathbf{xy}}(\mathbf{x}_0)$. We actually use the retrieved state profiles \mathbf{x} and \mathbf{y} as the zero order model states \mathbf{x}_0 and \mathbf{y}_0

Often output profiles in the L2b processor depend on $P > 1$ input profiles. Then the input vector \mathbf{x} corresponds to multiple profiles \mathbf{x}_μ stacked upon one another, each of length N ,

$$\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_\mu^T, \dots, \mathbf{x}_P^T), \quad (5.6)$$

where

$$\mathbf{x}_\mu = (x_{\mu 1}, x_{\mu 2}, \dots, x_{\mu i}, \dots, x_{\mu N})^T. \quad (5.7)$$

\mathbf{x} is thus of length $P \cdot N$ (with index $\mu = 1, 2, \dots, P$ and $i = 1, 2, \dots, N$). Essentially all non-linear L2b operators work level-by-level and no cross-correlations between the different \mathbf{x}_μ are assumed, i.e. only the elements $x_{\mu i}$ within \mathbf{x}_μ are correlated. Practically this means that inter-variable correlations (such as between background temperature and background humidity) are disregarded, which we found justified from experimental check (see below). Equation 5.2 can then be simplified, and each element of \mathbf{C}_y determined by

$$C_{y,ij} = \sum_{\mu=1}^P \frac{\partial y_i}{\partial x_{\mu i}} \frac{\partial y_j}{\partial x_{\mu j}} C_{x_{\mu i} x_{\mu j}}. \quad (5.8)$$

C.2 provides a proof for this simplification.

Output random uncertainty profiles \mathbf{u}_y^r and correlation matrices \mathbf{R}_y can be derived from \mathbf{C}_y , using (in index notation)

$$u_{y,i}^r = \sqrt{C_{y,ii}} \quad (5.9)$$

and

$$R_{y,ij} = \frac{C_{y,ij}}{u_{y,i}^r u_{y,j}^r}. \quad (5.10)$$

On the other hand, correlation matrix and random uncertainties can be combined to the covariance matrix using

$$C_{y,ij} = u_{y,i}^r \cdot u_{y,j}^r \cdot R_{y,ij}. \quad (5.11)$$

Correlation length profiles \mathbf{l}_y are determined from \mathbf{R}_y . The used algorithm estimates the correlation length by searching for the distances downward and upward of the correlation functions' asymmetric peak at which the correlation function has dropped to $1/e$. The returned correlation length is the arithmetic mean of these two distances. Additionally the correlation length is constrained by the data domain, i.e., the correlation length can never be larger than the profile length.

Because in the L1b and L2a processing, estimated *systematic uncertainties* were assumed to be highly correlated along the RO profiles over the active ranges of the operators, they could be propagated by application of the state retrieval operator, \mathbf{A}^{xy} (or the linearized version of the operator, \mathbf{J}^{xy}), using

$$\mathbf{u}_y^s = \mathbf{A}^{\text{xy}} \cdot \mathbf{u}_x^s, \quad (5.12)$$

where \mathbf{u}_x^s and \mathbf{u}_y^s are the rank-1 systematic uncertainty profiles of the input and output variables. In the operators of the L2b processor, the systematic uncertainties are still highly correlated along *one* input profile \mathbf{x}_μ , however, as with the random uncertainties, the correlations *between* the P input profiles are small and can therefore be disregarded. Moreover, for the systematic uncertainty estimation, we are only interested in the diagonal elements, and thus, because operators again work level-by-level, using

$j = i$ and $(u_{x,i}^s)^2 := C_{x,ii}$, Equation 5.8 can be further simplified to a univariate problem for systematic uncertainties,

$$(u_{y,i}^s)^2 := u_{y,ii}^s = \sum_{\mu=1}^P \left(\frac{\partial y_i}{\partial x_{\mu i}} \right)^2 (u_{x,i}^s)^2. \quad (5.13)$$

In addition to random uncertainties, systematic uncertainties and error correlation matrices, upon completion the algorithm will also estimate *resolution* profiles \mathbf{w}_y as context information along the provided random uncertainty profiles, and observation-to-background weighting ratio profiles $\mathbf{r}_y^{\text{obw}}$, for quantifying the relative contributions of observational vs. background information in the output profiles.

To ensure the validity of Equation 5.8, we empirically analyzed inter-variable correlations, particularly between the input variables background temperature and background specific humidity on the one hand, and dry temperature and dry pressure on the other hand. We found the inter-variable correlations to be small to negligible for background temperature and specific humidity, where they are up to ~ 0.2 at same vertical level, quickly decreasing for cross-levels, as illustrated in the Appendix, Figure C.1.

For dry temperature and dry pressure we provide an estimation of the potential scale of the effect of neglecting inter-variable correlations in Section 5.4.1, for the example case of T_q , the retrieved temperature with humidity prescribed.

We use profiles $x(z)$ (with argument) and vectors \mathbf{x} (bold) interchangeably, since in the rOPS algorithm vectors are the numerical representations of the vertical profiles.

5.3.2 Data Sources and Preparation

Figure 5.2 provides an overview on the main used variables in the L2b processor and illustrates how the input variables are processed. The corresponding definitions of the variable symbols can be found in Table 5.1.

The main input for the L2b processor are dry-air profiles of retrieved RO events, provided by the L2a processor of the rOPS, i.e. dry temperature $T_d(z)$ (shorthand for the SKS2017 notation $T_{dr}(z)$), and dry pressure $p_d(z)$, plus their associated systematic uncertainty profiles $u_{T_d}^s(z)$, $u_{p_d}^s(z)$, and covariance matrices C_{T_d} , C_{p_d} .

The retrieved RO events for this study were obtained using a test-day ensemble of excess phase profiles from the 15th of July 2008. For the CHAMP and COSMIC ensembles, orbit state and excess phase profiles were provided by CDAAC/UCAR. For the simMetOp ensemble, the End-to-End GNSS Occultation Performance Simulation and Processing System (EGOPS) (Fritzer et al. 2009) was used for generating the simulated MetOp orbit state and excess phase profiles with realistic receiver noise. These excess phase profiles were then processed with the rOPS L1b and L2a processors to obtain the dry-air profiles for each event.

The respective estimated basic and apparent systematic uncertainties, represented by the systematic uncertainty profiles $u_{T_d}^s(z)$ and $u_{p_d}^s(z)$, and the estimated random

uncertainties, represented by the covariance matrices \mathbf{C}_{Td} and \mathbf{C}_{pd} , are provided automatically along with the L2a dry-air retrieval output due to the integrated uncertainty propagation (SKS2018, SKS2017). From the test-day ensemble of the 15th July 2008, one representative COSMIC event with a near-surface penetration depth was chosen as example case to illustrate the algorithm sequence.

Due to under-determination, the ECVs pressure, temperature and humidity can not all be independently calculated from RO data alone, and therefore background information on temperature and humidity is needed. Dry-air profiles allow to calculate pressure and temperature profiles without the use of background information, in regions where the magnitude of the atmospheric moisture is negligible. For the calculation of dry-air profiles, the 'wet term' in the refractivity equation (Smith and Weintraub 1953), and thus atmospheric moisture, is disregarded, irrespective of the actual humidity of the atmosphere. The differences between dry-air and physical pressure and temperature profiles are elaborated by Scherllin-Pirscher et al. (2011b) and Scherllin-Pirscher et al. (2017).

In the rOPS moist-air retrieval, background information on temperature and specific humidity is used to calculate the ECVs also in regions of non-negligible atmospheric moisture. To generate these background profiles temperature $T_b(z)$, and specific humidity $q_b(z)$, plus their associated systematic uncertainty profiles $u_{T_b}^s(z)$, $u_{q_b}^s(z)$, and covariance matrices \mathbf{C}_{T_b} and \mathbf{C}_{q_b} , we used European Centre for Medium-Range Weather Forecasts (ECMWF) forecast and analysis fields. The latitude, longitude and time of the retrieved RO event were used to create co-located background temperature and specific humidity profiles, $T_b(z)$, and $q_b(z)$, from ECMWF short-range (24 h) forecast fields, chosen for their high quality (Bauer et al. 2015; Untch et al. 2006). The used ECMWF fields were available at 91 vertical levels (L91), have a horizontal resolution of about 300 km (triangular truncation T42), corresponding to the approximate horizontal resolution of RO profiles, and were calculated for the 15th of July 2008.

We took the systematic uncertainty of the background temperature $u_{T_b}^s(z)$, as in Li et al. (2013), Equation (14) (a model slightly simplified from Scherllin-Pirscher et al. (2011b)), essentially a vertical constant ~ 0.5 K. The systematic uncertainty of the background specific humidity is about $u_{q_b}^s = 0.05 \cdot q_b$ kg/kg, in line with parameter b_a in Li et al. (2018) (which is defined relative to the mean analysis profile, not the forecast profile, however).

Global temperature and specific humidity uncertainty fields, from ECMWF short-range forecast and analysis fields, calculated according to Li et al. (2013), are used as basis for the estimated random uncertainty profiles $u_{T_b}^r(z)$ and $u_{q_b}^r(z)$. Vertical profiles are at the location of the retrieved RO events were drawn from the appropriate horizontal bins of the global fields. Similar to the background uncertainties, background correlation matrices \mathbf{R}_{T_b} and \mathbf{R}_{q_b} are calculated according to Li et al. (2013). The covariance matrices \mathbf{C}_{T_b} and \mathbf{C}_{q_b} are then calculated using Equation 5.11 (Item (1.1) in Figure 5.2).

To create the random draws for the MC validation (see Section 5.5), representative

covariance matrices $\mathbf{C}_X^{\text{STD}}$ (for $X \in \{T_d, p_d, T_b, q_b\}$) are selected, a large number K of error realization profiles $\epsilon_{X,k}^r$ created from them and superimposed onto simulated “true” profiles T_d^T , p_d^T , and T_b^T and q_b^T . The “true” dry-air profiles are obtained from end-to-end simulated ‘error-free’ EGOPS simulations (i.e., no receiver system errors superimposed), and the *true* background profiles were constructed from the ECMWF analysis fields. The MC validation method is described in Section 5.5, and in more detail in SKS2017, where the same method is applied to the rOPS L2a uncertainty propagation.

5.4 Algorithm Sequence and Example Results

In this Section the L2b uncertainty propagation algorithm sequence is introduced. We illustrate the effects of the algorithm on the main uncertainty variables with the help of the chosen COSMIC example case. The L2b input variables, calculated as described in Section 5.3.2, are illustrated in Figures 5.3 to 5.6.

For each of the L2b retrieval steps, i.e., segments (1a), (1b), (2), and (3) in Figure 5.2, the results for the output variables are shown in Figures 5.7 to 5.12. These variables are the state profile $Y(z)$ (with $Y \in \{T_q, q_T, T_e, q_e, p_e, \rho_e\}$) in panels (a), the estimated systematic uncertainty profile $u_Y^s(z)$ and the estimated random uncertainty profile $u_Y^r(z)$ in (b), representative correlation functions $R_{Y,i}$ (with i such that $z_i \in \{3 \text{ km}, 7 \text{ km}, 12 \text{ km}, 17 \text{ km}\}$) in (c), and the correlation length profiles $l_Y(z)$ and resolution profiles $w_Y(z)$ in (d).

The full rOPS moist-air retrieval algorithm is only performed up to a dry-moist transition altitude, z^{DM} plus a transition half-width Δz^{DM} . Above $z^{\text{DM}} + \Delta z^{\text{DM}}$ (17 km), the moist thermodynamic profiles are approximated using the dry-air L2a profiles and the specific humidity, $q_b(z)$, leaving uncertainties smaller than the accuracy target in the rOPS (Li et al. 2018). The Figures 5.3 to 5.12 show the domain below z^{DM} , set to 20 km for this illustrative example case. Operationally, the transition altitude is set to 16 km in the rOPS, with a transition half-width Δz^{DM} of 1 km (transition range 15 km to 17 km).

The detailed step-by-step presentation of the algorithm in C.1 is structured into the same subsections as this section, to facilitate convenient comparison to the more thorough algorithm description in the appendix.

Figure 5.3 shows the input profiles for dry pressure, as provided by the L2a processor. Because of the integration in the pressure integral, the random uncertainty $u_{p_d}^r(z)$ is relatively small compared to the systematic uncertainty $u_{p_d}^s(z)$, which includes structural uncertainty stemming from the hydrostatic equilibrium assumption (Figure 5.3b). The correlation functions $R_{p_d,i}$ in Figure 5.3c have a relatively pronounced peak compared to the example event shown in SKS2017, Figure 6d. This is due to the differences between the simulated event shown in SKS2017 and the real event here. Some recent improvements in the statistical optimization further increased this effect. The correlation length $l_{p_d}(z)$, shown in Figure 5.3d increases strongly towards 20 km altitude (the correlation length determination is explained in Section 5.3.1, and in more detail in SKS2017).

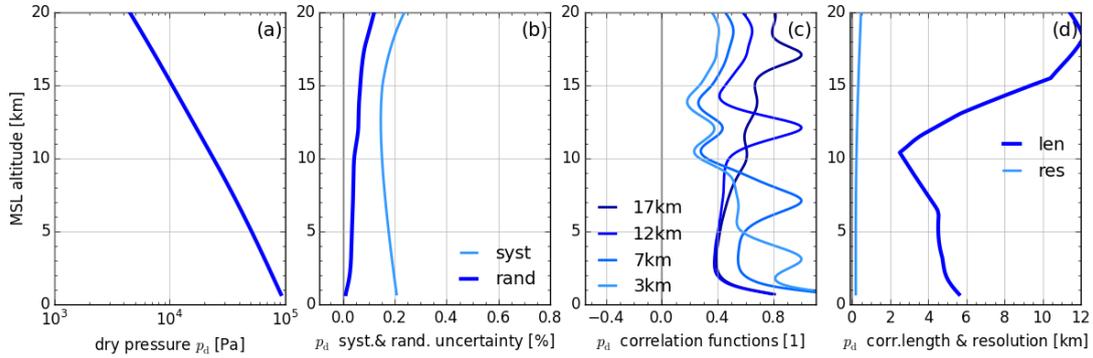


Figure 5.3: Input profiles retrieved dry pressure p_d in (a), relative systematic uncertainty profiles $100 \cdot u_{p_d}^s/p_d$ and relative random uncertainty profiles $100 \cdot u_{p_d}^r/p_d$ in (b), representative correlation functions $R_{p_d,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{p_d} and resolution w_{p_d} in (d).

Figure 5.4 illustrates the input profiles for the dry temperature. Again the peaks of the correlation functions are more pronounced than in SKS2017, for the same reasons as for the dry pressure.

For the background temperature, Figure 5.5a shows the state profile $T_b(z)$, and as context information also the dry temperature profile $T_d(z)$, to reveal the impact of the humidity (which is ignored in $T_d(z)$). $T_b(z)$ also has somewhat broader correlations than $T_d(z)$.

The specific humidity of the chosen example case is relatively small (maximum magnitude of ~ 1.1 g/kg), but these dry conditions allowed the near-surface penetration depth of this event, down to about 400 m above sea level.

The slight jitter visible in correlation functions of $T_b(z)$ and $q_b(z)$, compared to those shown in Li et al. (2018), is due to the use of zonally dissolved correlation matrices here, rather than the global matrix as in Li et al. (2018), causing the number of profiles used to generate the matrix to be an order of magnitude smaller.

5.4.1 Direct Retrievals of Temperature and Specific Humidity

In these first two direct retrieval steps (Segments 1a and 1b in Figure 5.2), physical temperature $T_q(z)$ and pressure $p_q(z)$ profiles (using background information on specific humidity, $q_b(z)$) and specific humidity $q_T(z)$ and pressure $p_T(z)$ profiles (using background information on temperature, $T_b(z)$) are calculated based on the dry-air variables $T_d(z)$ and $p_d(z)$.

5 Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis

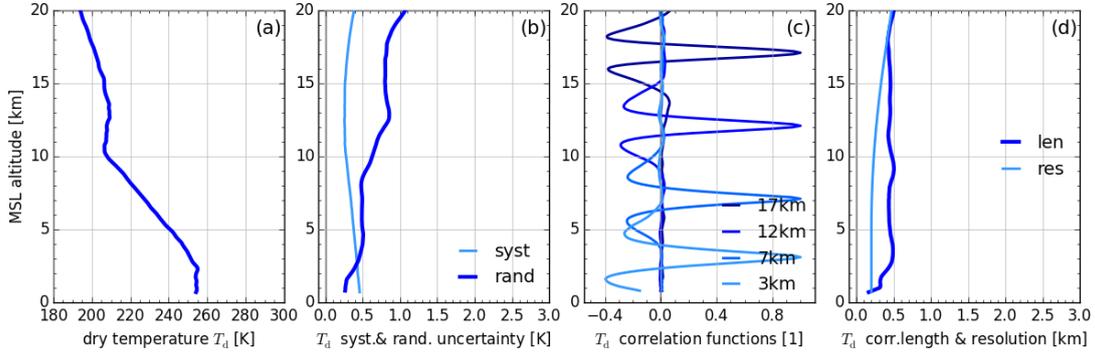


Figure 5.4: Input profiles retrieved dry temperature T_d in (a), systematic uncertainty profiles $u_{T_d}^s$ and random uncertainty profiles $u_{T_d}^r$ in (b), representative correlation functions $R_{T_d,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{T_d} and resolution w_{T_d} in (d).

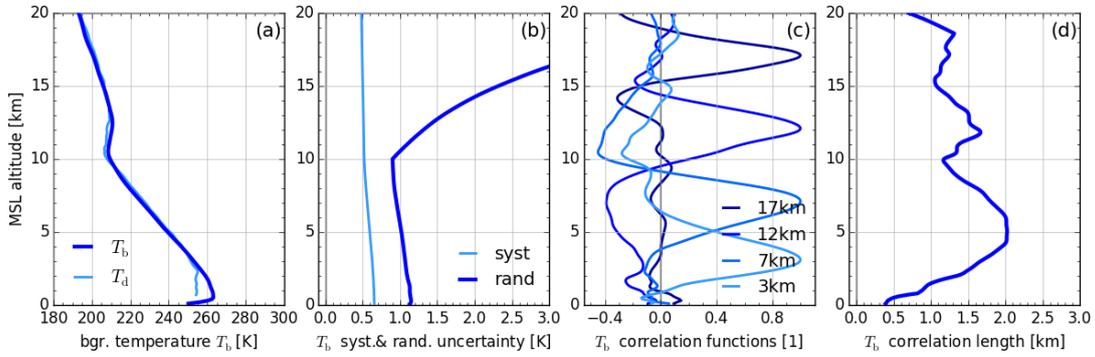


Figure 5.5: Input profiles background temperature T_b (with T_d for comparison) in (a), systematic uncertainty profiles $u_{T_b}^s$ and random uncertainty profiles $u_{T_b}^r$ in (b), representative correlation functions $R_{T_b,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{T_b} in (d).

Retrieval of Temperature with Specific Humidity Prescribed

In this direct retrieval, $T_q(z)$ and $p_q(z)$ are solved iteratively using $p_d(z)$, $T_d(z)$, and $q_b(z)$ as input (Items 1.2 and 1.3 in Figure 5.2), but the uncertainty propagation (shown for the covariance propagation of the random uncertainty in Items 1.4 and 1.5 in Figure 5.2), can be determined step-by-step, by linearization of the implicit state retrieval equations. The results for $p_q(z)$ are not shown, but Figure 5.7 depicts all relevant variables for $T_q(z)$. Although $q_b(z)$ has a longer correlation length than $T_d(z)$ over all altitudes, the random uncertainty $u_{T_q}^r(z)$ and the correlation length $l_{T_q}(z)$ increase in the lower troposphere. This is because only in the lower troposphere $q_b(z)$ and its random uncertainty are dominating in terms of magnitude compared to $T_d(z)$ (while above the lower troposphere its the other way round).

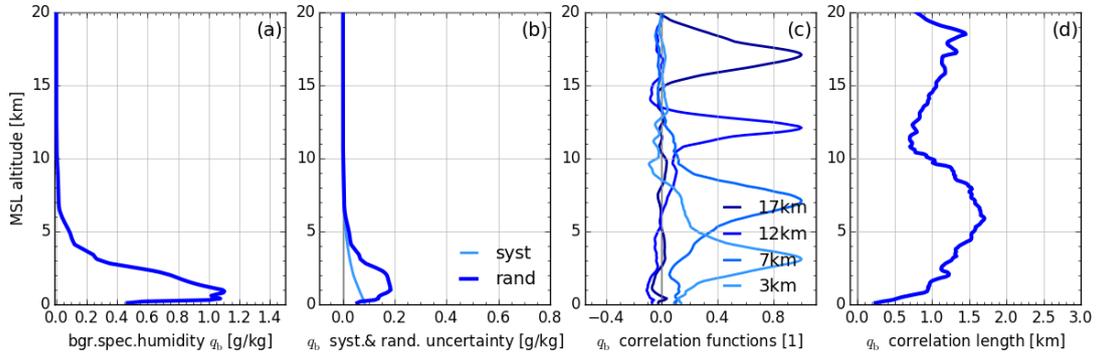


Figure 5.6: Input profiles background specific humidity q_b in (a), systematic uncertainty profiles $u_{q_b}^s$ and random uncertainty profiles $u_{q_b}^r$ in (b), representative correlation functions $R_{q_b,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{q_b} in (d).

In the propagation of random uncertainty to $u_{T_q}^r(z)$, we did not consider inter-variable correlations, because we assumed that the dry parameters $T_d(z)$ and $p_d(z)$ are essentially uncorrelated. This is a reasonable assumption (even though correlations are found somewhat larger than for the inter-variable correlation between background quantities $T_b(z)$ and $q_b(z)$), given that here and also in the derivation of $q_T(z)$ (Section 5.4.1), the resulting uncertainty is strongly dominated by the contribution of one of the input variables, and thus the consideration of the inter-variable correlation between $T_d(z)$ and $p_d(z)$ would have a very small impact on the estimated uncertainty of the calculated output variables.

For example, in our representative COSMIC case, when calculating variance $(u_{T_q,i}^r)^2 := C_{T_q,ii}$ (Equation C.22, Item 1.5 in Figure 5.2), above ~ 5 km, the contributions stemming from the variances of the dry temperature, $C_{T_d,ii}$, are almost two orders of magnitude larger (ca. 0.25 K^2 to 0.9 K^2) than those of the dry and the retrieved pressure, $C_{p_d,ii}$, and $C_{p_q,ii}$ (ca. 0.001 K^2 to 0.007 K^2), thus even in the worst case scenario, i.e. 100% correlation between $T_d(z)$ and $p_d(z)$, their arithmetically summed-up terms (rather than by sum-of-squares), would still be essentially equal to the $T_d(z)$ term. Below 5 km, the specific humidity term is two orders of magnitude larger than all others.

Retrieval of Specific Humidity with Temperature Prescribed

The directly retrieved specific humidity profile $q_T(z)$ is also iteratively calculated together with $p_T(z)$, using prescribed $T_b(z)$ and the dry parameters as input (Item 1.6 and 1.7 in Figure 5.2). The uncertainty propagation is done by linearization of the retrieval operators (Item 1.8 and 1.9 in Figure 5.2). The state profile $q_T(z)$ (Figure 5.8a) resembles $q_b(z)$, but has a “blob” near 10 km. As can be seen in Figure 5.8b, this “blob” is of the same order of magnitude as the uncertainty and thus just a retrieval artefact within

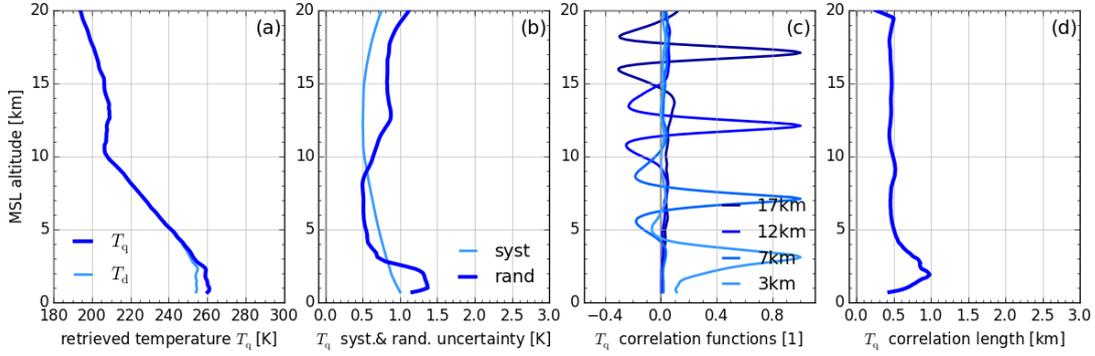


Figure 5.7: Results for retrieved temperature profiles T_q (with T_d for comparison) in (a), systematic uncertainty profiles $u_{T_q}^s$ and random uncertainty profiles $u_{T_q}^r$ in (b), representative correlation functions $R_{T_q,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{T_q} in (d).

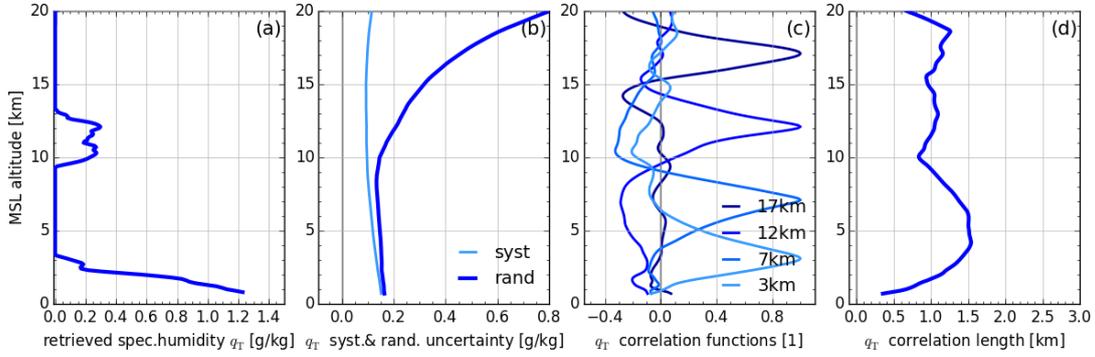


Figure 5.8: Results for retrieved specific humidity profiles q_T in (a), systematic uncertainty profiles $u_{q_T}^s$ and random uncertainty profiles $u_{q_T}^r$ in (b), representative correlation functions $R_{q_T,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{q_T} in (d).

noise level. When compared to Figure 5.5, the shapes of the random uncertainty, the correlation functions and the correlation length profile indicate that $q_T(z)$ is strongly influenced by background temperature $T_b(z)$.

5.4.2 Optimal Estimation of Temperature and Specific Humidity

In this second step (Segment 2 in Figure 5.2), the temperature $T_q(z)$ from the direct retrieval is combined with background temperature $T_b(z)$ in an optimal estimation (Item 2.1 in Figure 5.2). Also specific humidity $q_T(z)$ and $q_b(z)$ are combined in an optimal way (Item 2.2 in Figure 5.2).

In Li et al. (2018) the optimal estimation is based on an inverse variance weighting of retrieved and background uncertainties. Due to the implemented covariance propagation

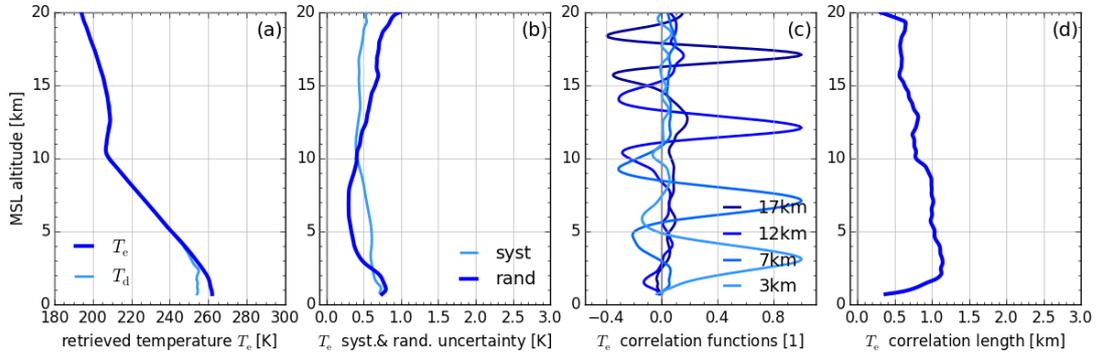


Figure 5.9: Results for optimal estimation temperature profiles T_e (with T_d for comparison) in (a), systematic uncertainty profiles $u_{T_e}^s$ and random uncertainty profiles $u_{T_e}^r$ in (b), representative correlation functions $R_{T_e,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{T_e} in (d).

in the context of this study, the weighting matrix of the optimal estimation can now be calculated from the propagated covariance matrices (Item 2.3 and 2.5 in Figure 5.2) and differences in correlation lengths thus taken into account.

Figure 5.9a shows the results for optimal estimation temperature $T_e(z)$ (again compared to $T_d(z)$). Due to the larger random uncertainty $u_{T_b}^r(z)$ of the background temperature $T_b(z)$ (Figure 5.9b), compared to those of the retrieved temperature $T_q(z)$, the optimal estimation gives more weight to $T_q(z)$, thus also the correlation functions rather resemble those of $T_q(z)$ (Figure 5.9c).

The optimal estimation specific humidity $q_e(z)$ on the other hand (Figure 5.10) is dominated by the background humidity $q_b(z)$, particularly above 3 km, where the background uncertainty $u_{q_b}^r(z)$ is much smaller than the retrieved uncertainty $u_{q_T}^r(z)$. For more moist tropospheric conditions, such as in the tropics, the background humidity plays much less role.

5.4.3 Optimal Estimation of other Thermodynamic Variables

From the temperature and specific humidity profiles, determined by the optimal estimation, other thermodynamic variables like the optimal pressure profile $p_e(z)$ (Item 3.1 in Figure 5.2), the volume mixing ratio of water vapour $V_{we}(z)$ (Item 3.2 in Figure 5.2), the partial pressure of water $e_e(z)$ (Item 3.3 in Figure 5.2) and the density profile ρ_e (Item 3.4 in Figure 5.2) can be calculated.

The results for $p_e(z)$ of the COSMIC example case are shown in Figure 5.11. The pressure uncertainties and correlations are very similar to those of the dry pressure, shown in Figure 5.3, also because the atmospheric moisture in the example case is fairly small.

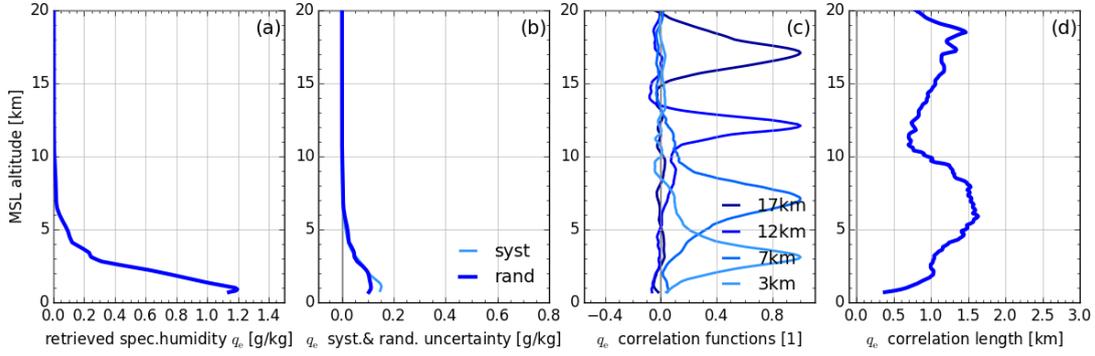


Figure 5.10: Results for optimal estimation specific humidity profiles q_e in (a), systematic uncertainty profiles $u_{q_e}^s$ and random uncertainty profiles $u_{q_e}^r$ in (b), representative correlation functions $R_{q_e,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{q_e} in (d).

The results for $\rho_e(z)$ (Figure 5.12) show that its random uncertainties and correlation functions are mostly influenced by the temperature profile $T_e(z)$, because they dominate over those of the pressure and the specific humidity profile.

5.5 Algorithm Validation

The GUM advises to use a Monte-Carlo (MC) method for uncertainty propagation if the retrieval operators do not fulfill the criteria for a GUM-type covariance propagation (CP). In our case the MC method is only used to validate the results of the CP, as recommended by the GUM (JCGM 2011).

For the MC validation of the CP, a large number M of error profile realizations $\epsilon_{X,m}^r(z)$ (with $m = 1, \dots, M$) needs to be superimposed onto a 'true' state profile $x^T(z)$, creating M draws, $x_m^{\text{MC}}(z)$, representative for the distribution of random variable $X(z)$, each of which is then ran through the state retrieval to create sample outputs $y_m^{\text{MC}}(z)$ at every step along the retrieval chain. These samples are then recombined to output covariance matrices, \mathbf{C}_y^{MC} , using the sum over all outer products between the output profiles $y_m^{\text{MC}}(z)$, i.e.,

$$\mathbf{C}_{y,ij}^{\text{MC}} = \frac{1}{M-1} \sum_{m=1}^M (y_{i,m} - \bar{y}_i^{\text{MC}})(y_{j,m} - \bar{y}_j^{\text{MC}}), \quad (5.14)$$

for the MC covariance matrix of output $Y(z)$, where \bar{y}^{MC} is the mean profile of the draws, N is the length of the rank-1 profiles, and $i, j = 1, \dots, N$. For each output variable, \mathbf{C}_y^{MC} is compared to the propagated covariance matrices \mathbf{C}_y^{CP} . To be able to attribute potential changes between CP and MC covariance matrices more accurately, we use Equations 5.9 and 5.10 to get $u_y^r(z)$ and \mathbf{R}_y , and compare them separately.

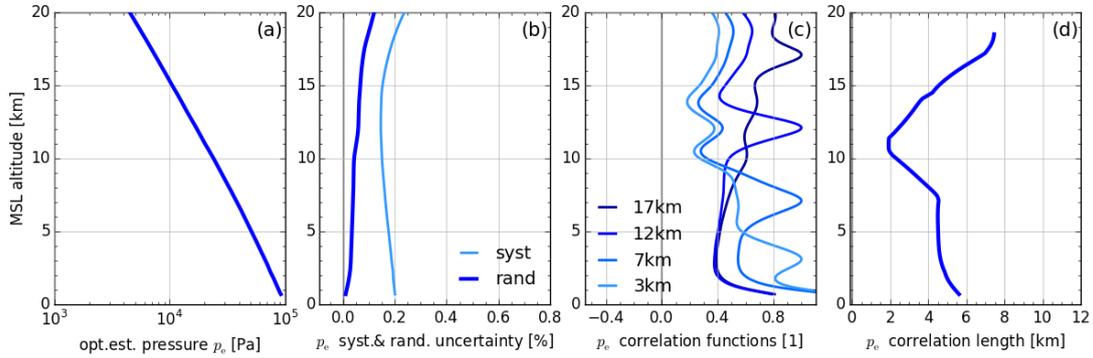


Figure 5.11: Results for optimal estimation pressure profiles p_e in (a), relative systematic uncertainty profiles $100 \cdot u_{pe}^s / p_e$ and relative random uncertainty profiles $100 \cdot u_{pe}^r / p_e$ in (b), representative correlation functions $R_{pe,i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{pe} in (d).

In the moist-air retrieval, there are four input quantities $T_d(z)$, $p_d(z)$, $T_b(z)$, and $q_b(z)$, of which the full input distribution, given by the stacked means-vector and the block-covariance matrices containing all covariances and cross-covariances, would need to be sampled. However, since the covariance propagation is based on the assumption of zero inter-variable correlation, the consistent MC input distribution can be represented by the covariance matrices \mathbf{C}_{T_d} , \mathbf{C}_{p_d} , \mathbf{C}_{T_b} and \mathbf{C}_{q_b} alone, and all potential cross-covariances between these variables are assumed to be zero. Consequentially the four random variables can also be sampled independently (through independent seeds for the random draws). The sampling of the four input variables is illustrated by Item (a) in Figure 5.2.

Upon completion of the L2b implementation, the corresponding random uncertainty profile $u_y^r(z)$ and correlation matrix \mathbf{R}_y will be calculated, for each of the output variables $Y(z)$ (with $Y \in \{T_q, T_e, q_T, q_e, p_e, \rho_e\}$) and compared to its CP equivalent in a validation result figure.

5.6 Performance Demonstration

We aim to evaluate the performance of the implemented uncertainty propagation algorithm, using a test-day of real (CHAMP, COSMIC) and simulated (simMetOp) data of GNSS RO satellite missions.

The related work is on-going. Upon completion, results for systematic and random uncertainty profiles, as well as correlation length, resolution and observation-to-background weighting profiles, will be shown for input variables $T_d(z)$, $p_d(z)$, $T_b(z)$, and $q_b(z)$, and output variables $T_e(z)$, $q_e(z)$, $p_e(z)$, and $\rho_e(z)$, in two performance demonstration result figures.

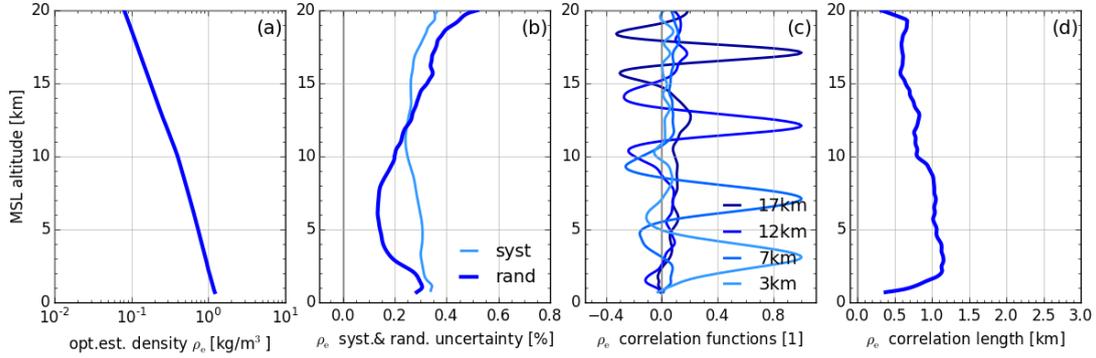


Figure 5.12: Results for optimal estimation density profiles ρ_e in (a), relative systematic uncertainty profiles $100 \cdot u_{\rho_e}^s / \rho_e$ and relative random uncertainty profiles $100 \cdot u_{\rho_e}^r / \rho_e$ in (b), representative correlation functions $R_{\rho_e, i}$ at 3, 7, 12 and 17 km in (c), and correlation length l_{ρ_e} in (d).

5.7 Conclusions

In order to deliver ECV data records of climate benchmark quality using GNSS radio occultation, uncertainties need to be propagated from the raw data through the entire retrieval chain to ECVs. In the rOPS, developed at the WEGC, we implemented integrated uncertainty propagation to deliver estimates for systematic and random uncertainties, error correlations, vertical resolution profiles, and the observation-to-background ratios, along with the basic profiles retrieval. The rOPS is the first processing system to do so among all existing RO processing systems (Ho et al. 2012; Steiner et al. 2013).

We validated the implemented moist-air algorithm via comparison to Monte-Carlo sample propagation results and demonstrated the performance of the algorithm using real data ensembles. These results will soon be completed and from the preliminary results we find close agreement between the implemented covariance propagation of random uncertainties and the Monte-Carlo validation runs, verifying the correctness of the implemented algorithm. The test-day ensembles for three different missions show reliable, robust and consistent results with only a small number of failed profiles and outliers.

Together with the integration of the uncertainty propagation algorithm from excess phase to atmospheric bending angle (L1b) presented by SKS2018 and from atmospheric bending angle profiles to dry-air profiles (L2a) presented by SKS2017, the rOPS can now provide these estimates from excess phase to moist-air atmospheric profiles.

Once the last link through the L1a processor from raw occultation and orbit data is completed (Innerkofler et al. 2016; Innerkofler et al. 2017a), the rOPS can run with integrated uncertainty estimation through the entire retrieval chain, a major step towards climate benchmark data provision, and beneficial for atmospheric and climate science

and applications.

5.8 Acknowledgements

We thank UCAR/CDAAC Boulder for access to their RO excess phase and orbit data (available at <http://cdaac-www.cosmic.ucar.edu/>) as well as ECMWF Reading for access to their analysis and forecast data (available at <http://www.ecmwf.int/en/forecasts/datasets>). To access the relevant result files of the uncertainty propagation, please contact the corresponding author. The developed and implemented algorithm is provided in the appendix. The work was funded by the Austrian Aeronautics and Space Agency of the Austrian Research Promotion Agency (FFG-ALR) under projects OPSCLIMTRACE and by the European Space Agency (ESA) under project MMValRO-E.

Summary and Conclusions

TO contribute to the solution of the global climate monitoring problem, the Reference Occultation Processing System (rOPS) project at the Wegener Center for Climate and Global Change (WEGC) aims to establish a fully traceable radio occultation retrieval, to which this thesis has been contributing.

Current observational records of the Earth's free atmosphere have shortcomings from a climate perspective, because the corresponding observation systems were designed for purposes other than climate monitoring, and thus the accuracy, long-term stability, and data homogeneity and coverage in the existing observational records were insufficient for climate purposes.

The establishment of observational records, which are of global coverage, of high accuracy and long-term stability, tied to irrefutable standards, and measure Essential Climate Variables (ECVs), so-called *climate benchmark observations*, is of crucial importance for the climate sciences.

Observations based on the Global Navigation Satellite System (GNSS) radio occultation (RO) technique have the potential to deliver these climate benchmark observations for thermodynamic ECVs in the free atmosphere, if the claimed accuracy of the ECVs can be traced to the raw data, the accuracy of which is tied to absolute SI standards. To establish this trace, uncertainties need to be propagated from the raw data through the entire retrieval chain to the ECVs. The goal of this thesis was the development of this uncertainty propagation algorithm and its implementation into the rOPS processing routines, from excess phase profiles to ECVs.

In the context of this thesis, integrated uncertainty propagation to deliver estimates for systematic and random uncertainties, error correlations, vertical resolution profiles, and observation-to-background ratios, along with the basic profiles retrieval were implemented into the L1b, L2a and L2b processors of the rOPS. The uncertainty propagation chain in the rOPS is thus complete from excess phase profiles via bending angle and refractivity profiles to the thermodynamic ECVs. Among the existing RO processing systems, the rOPS at the WEGC is the first to provide these uncertainty quantities.

Three core publications present the developed uncertainty propagation algorithms, the results of the algorithm validation and the demonstration of the robust performance of these algorithms. Schwarz et al. (2018a) introduced the uncertainty propagation from

excess phase profiles to atmospheric bending angle profiles (L1b processing), Schwarz et al. (2017) the uncertainty propagation from atmospheric bending angle profiles to dry-air atmospheric profiles (L2a processing), and Schwarz et al. (2018b) from dry-air profiles to moist-air profiles (L2b processing), which are the final thermodynamic ECV profiles.

The validation of the covariance based propagation was performed based on Monte Carlo (MC) methods, and the results showed a high level of agreement of the outcomes of these two independent uncertainty propagation methods. The robust performance of the algorithms was demonstrated using test-day ensembles of different Radio Occultation (RO) satellite missions, and good consistency of the results with previous empirical error estimation studies was found.

On-going work concerns the uncertainty propagation from raw orbit and occultation data to the excess phase profiles (L1a processing), the uncertainty propagation through the wave-optics bending angle retrieval (part of the L1b processing) and the finalization of the uncertainty propagation through the moist-air retrieval (L2b processing), where MC validation and performance demonstration are in completion.

For some of the (particularly systematic) uncertainty contributions along the L1b, L2a and L2b propagation chain the evaluation of the uncertainty magnitudes relied on best guess estimates. These could be further improved if the magnitudes could be quantitatively constrained by a more thorough analysis of the effects at the root of these uncertainties. On-going work in a new project at the WEGC in the context of the rOPS development aims to deliver such improvements.

Upon completion of the full uncertainty chain, and thus the realization of the fully traceable retrieval, the WEGC can serve as global distribution center for RO data of benchmark quality, using data from past, present, and various future satellite missions such as the Constellation Observing System for Meteorology, Ionosphere, and Climate-2 (COSMIC-2) and the Feng Yun series-3 satellites (FY-3)/GNSS radio-occultation sounder (GNOS). rOPS data products can then provide a new reference standard for thermodynamic ECVs for atmosphere and climate research and applications.

**Integrating uncertainty propagation in GNSS radio
occultation retrieval: from excess phase to atmospheric
bending angle profiles**

A.1 Algorithm Description

In this appendix the rOPS L1b uncertainty propagation algorithm is introduced, following the L1b retrieval chain (Figure 3.2; Sect. 3.4) step by step, starting with excess phase profile L_r as input and proceeding to L_F , D_r , α_G , α_M , α_F and finally α_r . The relevant variable definitions and symbol explanations are summarized in Tables 3.1 and 3.2. A fully detailed algorithmic description is provided by Kirchengast et al. (2017b).

If not stated otherwise, elements of the vector-type vertical profiles are addressed using subscript i (with $i \in \{1, 2, \dots, N\}$), and optionally j (with $j \in \{1, 2, \dots, N\}$), running from top downward towards the bottom of the profile, where N is the number of vertical grid levels. Until the interpolation of all quantities to the common monotonic impact altitude grid z_a , all quantities are provided on an equidistant 50 Hz time grid t with grid points t_i .

All steps in Sects. A.1.1 and A.1.2 are applied to each of the GNSS transmitter channels' carrier frequencies f_{Tk} , as also indicated by the index k in Figure 3.2. In the notation of these sections we therefore suppress the index k for the convenience of simplified readability. Also for conciseness we write the estimated systematic uncertainty equations only for the total systematic uncertainties u^s and briefly address the type of the relevant components (whether basic systematic uncertainty u^b or apparent systematic uncertainty u^a) in the surrounding text.

A.1.1 Doppler shift retrieval

Basic low-pass filtering

The Doppler differentiation (item 1.4 in Figure 3.2) would potentially amplify high-frequency noise in the excess phase profiles. To avoid this amplification, a BWS low-pass filter (e.g., Smith 1999) is applied onto the excess phase profiles first (item 1.2 in Figure 3.2).

For this basic filtering the relative cutoff frequency f_c/f_s is set to 0.05, equivalent to $f_c = 2.5$ Hz, 21 grid points, or a cutoff period $\tau_c = 1/f_c = 0.4$ s, for the standard sampling rate f_s of 50 Hz used for all RO profiles in the L1b processor of the rOPS. The corresponding sample width of the Blackman window \tilde{M} (with samples $m \in \{0, \dots, M\}$) is set to $\tilde{M} = 2 \cdot f_s/f_c$, yielding 41 grid points. This ensures a reliable filter performance, also allowing the vertical resolution of the filtered data to be robustly quantified.

With such a design, the BWS low-pass filter combines efficient removal of high-frequency noise with a narrow smoothing window. The BWS filter thus achieves a better smoothing effect, while keeping a w_{LF} of higher resolution than a simple moving-average BC filter. Based on a time segment of a few seconds of the excess phase delta profile of the COSMIC example event (also used for Figs. 3.3 to 3.8), Figure A.1 illustrates how the BWS filter compares to boxcar filters of 11 and 21 grid points. The corresponding filter functions are displayed in Figure A.1a, while Figure A.1b compares the filter results.

It is clearly seen that the smoothing window width of the BWS filter best corresponds to an 11-point boxcar filter (confirmed numerically by minimization of the sum of squared differences between boxcar and BWS filter result), while giving considerably better filtering results (as for example visible between 31.5 and 32.0 s, where the 11-point boxcar filter zigzags around the BWS result). The effective filter width of the BWS filter, which we also term “boxcar-equivalent width”, is therefore its full width at half maximum (see Figure A.1a), corresponding to $\tilde{M}/4 + 1$ samples with our design.

The actually used sample width M of the BWS filter is equal to \tilde{M} , except that it decreases at the top and bottom of the profile such that it does not reach beyond the first/last element of the vector to be filtered. At the i th grid point (with $i \in \{1, 2, \dots, N\}$, and N being the profile length in grid points), the filter width M is thus

$$M = \begin{cases} \tilde{M} & \text{for } \tilde{M}/2 < i < N - \tilde{M}/2 \\ 2i - 1 & \text{for } 1 \leq i \leq \tilde{M}/2 \\ 2(N - i) + 1 & \text{for } N - \tilde{M}/2 < i < N \end{cases}. \quad (\text{A.1})$$

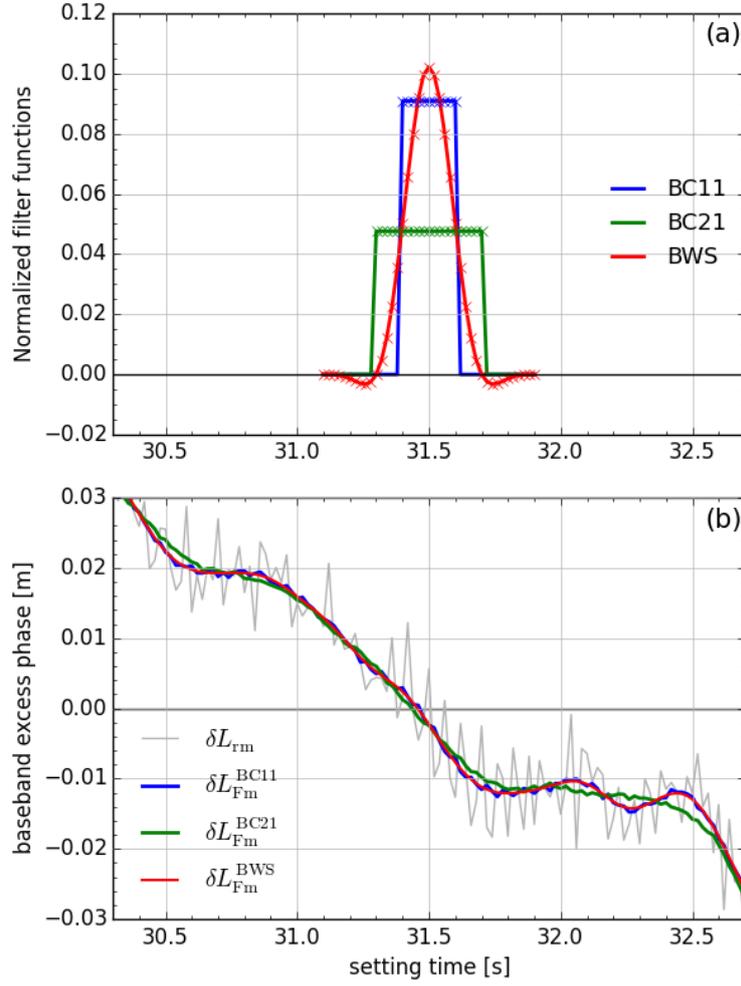


Figure A.1: Comparison of the Blackman windowed sinc (BWS) low-pass filter and boxcar (BC) filters based on a representative segment (between 30.3 and 32.7 s) of the excess phase profile L_{r1} of the COSMIC example event. (a) Filter functions for the BWS filter with $f_c = 2.5$ Hz and $M = 41$ points (“BWS”, red) and boxcar filters with $M = 21$ points (“BC21”, green) and with $M = 11$ points (“BC11”, blue), around the central value of the segment (31.5 s). (b) Filter effects on the excess phase profile L_{r1} from running the filters over the segment. Shown are the unfiltered excess phase delta profile (“ δL_{r_m} ”, light gray), the BWS filtered profile with $f_c = 2.5$ Hz and $M = 41$ points (“ $\delta L_{F_m}^{BWS}$ ”, red), and the boxcar filtered profiles with $M = 21$ points (“ $\delta L_{F_m}^{BC21}$ ”, green) and $M = 11$ points (“ $\delta L_{F_m}^{BC11}$ ”, blue).

The *state* profile of the filtered phase L_F is obtained using the “baseband approach” (Kirchengast et al. 2016a), i.e., by first subtracting a zero-order model profile L_m and

A Integrating uncertainty propagation in GNSS radio occultation retrieval: from excess phase to atmospheric bending angle profiles

applying the filter only to the delta profile $\delta L_{\text{rm}} = L_{\text{r}} - L_{\text{m}}$ (with the model profile being adequately smooth over the scale of the filter window width). This approach efficiently mitigates residual numerical biases. After the application of the BWS filter, the model profile is added back again. We express the BWS filter as a linear matrix operator \mathbf{A}^{BWS} and get (item 1.2 in Figure 3.2)

$$L_{\text{Fi}} = L_{\text{mi}} + \sum_{j=0}^N A_{ij}^{\text{BWS}} \cdot \delta L_{\text{rm}j} \quad (\text{A.2})$$

for the filtered excess phase, where $j \in \{1, 2, \dots, N\}$. The band matrix operator \mathbf{A}^{BWS} has elements

$$A_{ij}^{\text{BWS}} = \begin{cases} 0 & \text{for } j < i - M/2 \text{ and} \\ & \text{for } j > i + M/2 \\ w_{j-i+M/2} & \text{for } i - M/2 < j < i + M/2 \end{cases} . \quad (\text{A.3})$$

The central filter weight $w_{0+M/2}$ at $j = i$ is the $(M/2)$ th filter element (according to the definition of the BWS weights below); therefore its index is $M/2$. With $m = j - i + M/2$ (and therefore $0 \leq m \leq M$), each single BWS weight is calculated using

$$w_m = \frac{w_{\text{raw},m}}{\sum_{m=0}^M w_{\text{raw},m}} \quad (\text{A.4})$$

and

$$w_{\text{raw},m} = \begin{cases} \frac{\sin(2\pi f_c/f_s(m-M/2))}{m-M/2} \left[0.42 - 0.5 \cos \left(2\pi \frac{m}{M} \right) + 0.08 \cos \left(4\pi \frac{m}{M} \right) \right] & \text{for } m \neq M/2 \\ 2\pi f_c/f_s & \text{for } m = M/2 \end{cases} . \quad (\text{A.5})$$

The *estimated random uncertainty* is then propagated by covariance propagation (item 1.3 in Figure 3.2),

$$\mathbf{C}_{LF} = \mathbf{A}^{\text{BWS}} \cdot \mathbf{C}_{Lr} \cdot (\mathbf{A}^{\text{BWS}})^{\text{T}} . \quad (\text{A.6})$$

The random uncertainty profile u_{LF}^r and the error correlation matrix \mathbf{R}_{LF} are not needed for the subsequent random uncertainty propagation but are calculated from \mathbf{C}_{LF} for being available for the L1b output, using

$$u_{LF,i}^r = \sqrt{C_{LF,ii}} \quad (\text{A.7})$$

and

$$R_{LF,ij} = \frac{C_{LF,ij}}{u_{LF,i}^r u_{LF,j}^r}. \quad (\text{A.8})$$

The *correlation length* profile l_{Lr} has elements

$$l_{Lr,i} = \left. \frac{dz}{dt} \right|_i \cdot |t_i - t(R_{LF,ij} = 1/e)| \quad (\text{A.9})$$

computed upward and downward from the main peak of the correlation function and then averaged. Here dz/dt is the scan velocity profile, obtained from using the MSL altitude grid z_t calculated as part of the forward modeling towards L_m at the corresponding time grid t (cf. Table 3.2).

We note that after the L2a refractivity retrieval also the MSL altitude grid consistent with the retrieved refractivity profile could be used (as described by SKS2017, Appendix A therein), from a repeated forward modeling. The difference for the scan velocity estimate is found to be very small, however, since the forward-modeled z_t based on collocated refractivity profiles from ECMWF short-range forecast fields is already sufficiently reliable, and this also keeps the L1b processor as a decoupled predecessor of the L2a processor.

For the *estimated systematic uncertainty*, interpreted as a basic systematic uncertainty (Sect. 3.3.2), we apply the same low-pass filter as used for the *state* profile (item 1.2 in Figure 3.2), but with no zero-order profile subtracted, i.e.,

$$u_{LFi}^s = \sum_{j=0}^N A_{ij}^{\text{BWS}} \cdot u_{Lrj}^s. \quad (\text{A.10})$$

The *resolution* in time of L_F and its uncertainties, τ_{BW} , is the boxcar-equivalent width (cf. Figure A.1a) determined by the cutoff frequency f_c of the BWS filter,

$$\tau_{LF} \approx \frac{1}{f_c + \Delta f_c/2} \approx \frac{1}{2f_c}, \quad (\text{A.11})$$

with our design choice $\tilde{M} = 2(f_s/f_c)$ and with the BWS filter stopband-to-passband transition width being (Smith 1999)

$$\Delta f_c \approx \frac{4f_s}{\tilde{M}}. \quad (\text{A.12})$$

Given $f_c = 2.5$ Hz, this results in an effective resolution $\tau_{LF} = 0.2$ s and corresponds to the resolution obtained when applying a 11-point boxcar filter as explained at the beginning of this section above. The filter window intercomparison in Figure A.1a also illustrates this, because the full width at half maximum of the 2.5 Hz 41-point BWS filter is 11 points.

This resolution in time can finally be converted to the vertical (MSL altitude) resolution in space:

$$w_{LF,i} = \left. \frac{dz}{dt} \right|_i \cdot \tau_{LF}, \quad (\text{A.13})$$

where, as for the correlation length estimation (Equation A.9), the scan velocity profile is employed to convert from the time domain to MSL altitude domain.

Doppler shift derivation

After the application of the BWS filter to the excess phase profiles L_r (for both carrier frequencies of the given GNSS system), the *state* profile of the Doppler is derived from the filtered phase profiles L_F (item 1.4 in Figure 3.2). To minimize systematic errors from the numerical differentiation to negligible magnitude, the model profile L_m is again subtracted from the filtered phase profile,

$$\delta L_{Fm} = L_F - L_m, \quad (\text{A.14})$$

and the resulting delta profile δL_{Fm} is then differentiated. After the derivative, the zero-order Doppler shift model profile D_m is added (the latter also available from the forward modeling, in a form strictly consistent with the excess phase model profile L_m).

Based on careful tests of different formulations, we use a five-point derivative scheme. The discretization of this five-point derivative $\delta D_{rm,i}$ is given by

$$\begin{aligned} \delta D_{rm,i} &= \left. \frac{d\delta L_{Fm}(t)}{dt} \right|_i \\ &= \frac{-\delta L_{Fm,i-2} + 8\delta L_{Fm,i-1} - 8\delta L_{Fm,i+1} + \delta L_{Fm,i+2}}{-t_{i-2} + 8t_{i-1} - 8t_{i+1} + t_{i+2}} \end{aligned} \quad (\text{A.15})$$

for each of the frequencies (e.g., Syndergaard 1999). This can be expressed in matrix form as

$$D_{r,i} = D_{m,i} + \delta D_{rm,i} = D_{m,i} + \sum_{j=1}^N A_{ij}^{\text{L2D}} \cdot \delta L_{Fm,j}, \quad (\text{A.16})$$

using matrix operator \mathbf{A}^{L2D} with

$$\mathbf{A}^{\text{L2D}} = \frac{1}{12\Delta t} \begin{bmatrix} -18 & 24 & -6 & 0 & 0 & 0 & 0 & 0 & 0 \\ -6 & 0 & -6 & 0 & 0 & 0 & \ddots & 0 & 0 \\ -1 & 8 & 0 & -8 & 1 & 0 & & 0 & 0 \\ 0 & -1 & 8 & 0 & -8 & 1 & \ddots & 0 & 0 \\ & \ddots & & \ddots & & \ddots & & \ddots & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & -8 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 & & 24 & -18 \end{bmatrix}, \quad (\text{A.17})$$

where $\Delta t = t_{i+1} - t_i$, being 0.02 s in our case of $f_s = 50$ Hz.

The *estimated random uncertainty* can then be propagated (item 1.5 in Figure 3.2) using

$$\mathbf{C}_D = \mathbf{A}^{\text{L2D}} \cdot \mathbf{C}_{LF} \cdot \left(\mathbf{A}^{\text{L2D}}\right)^T. \quad (\text{A.18})$$

The covariance matrix is again (cf. Equations A.7 and A.8) decomposed into estimated random uncertainties and error correlation matrix (item 2.2 in Figure 3.2) using

$$u_{Dr,i}^r = \sqrt{C_{Dr,ii}} \quad (\text{A.19})$$

and

$$R_{Dr,ij} = \frac{C_{Dr,ij}}{u_{Dr,i}^r u_{Dr,j}^r}. \quad (\text{A.20})$$

For the *estimated systematic uncertainty*, further on interpreted as basic systematic uncertainty (cf. Equation A.10), we apply the derivative operator (item 1.4 in Figure 3.2) to the systematic uncertainties, with no zero-order profile subtracted, i.e.,

$$u_{Dr,i}^s = \sum_{j=1}^N A_{ij}^{\text{L2D}} \cdot u_{LF,j}^s. \quad (\text{A.21})$$

The *resolution* remains unaffected by the Doppler shift derivation, since the five-point sample width of the derivative operator is fully within the 11-point effective filter width (stopband) of the BWS filter applied before, so that $\tau_{Dr} = \tau_{LF}$ and $w_{Dr} = w_{LF}$.

A.1.2 Bending angle retrieval

GO bending angle retrieval

From the Doppler shift *state* profile D_r (again for both frequencies of the given GNSS system) we can derive the impact parameter profile a_t and GO bending angle profile α_G (item 2.1 in Figure 3.2) using first the geometric relation

$$D_{r,i} = [v_{R,i} \cos(\phi_{R,i}) - v_{T,i} \cos(\phi_{T,i})] - \dot{r}_{RT,i}, \quad (\text{A.22})$$

where

$$\phi_{R,i} = \eta_{R,i} - \arcsin\left(\frac{a_{t,i}}{r_{R,i}}\right) \quad (\text{A.23})$$

and

$$\phi_{T,i} = (\pi - \eta_{T,i}) - \arcsin\left(\frac{a_{t,i}}{r_{T,i}}\right) \quad (\text{A.24})$$

for each individual level of the time grid t_i , in order to determine a_t from sequential application to all levels (Kursinski et al. 1997; Syndergaard 1999). Here $v_{R,i} = |\vec{v}_{R,i}|$ is

the receiver velocity; $r_{R,i} = |\vec{r}_{R,i}|$ the receiver radial position; $\eta_{R,i}$ the angle between the receiver velocity and position vectors; $\phi_{R,i}$ then the angle between the receiver velocity and ray path vectors (and all these equivalently for the transmitter); and $\dot{r}_{RT,i} = \left| \frac{d(\vec{r}_T - \vec{r}_R)}{dt} \right|_i$ is the time derivative of the distance between the transmitter and the receiver at time t_i , i.e., the “kinematic straight-line Doppler shift” to be subtracted in Equation (A.22) to match the (excess) Doppler shift $D_{r,i}$ induced by the atmosphere (and ionosphere).

Based on a_t , the elements of the GO bending angle profile α_G are subsequently calculated using another geometrical relation:

$$\alpha_{G,i} = \theta_{RT,i} - \arccos\left(\frac{a_{t,i}}{r_{R,i}}\right) - \arccos\left(\frac{a_{t,i}}{r_{T,i}}\right), \quad (\text{A.25})$$

where $\theta_{RT,i}$ is the opening angle between the transmitter and receiver position vectors. Syndergaard (1999, Figure 1.5 therein) provides an illustration of the relevant geometry.

All the variables in Equations (A.22)–(A.25) are defined in the occultation plane spanned by the receiver and transmitter position vectors after oblateness correction (Syndergaard 1998), i.e., after they have been transformed to originate in the Earth ellipsoid’s center of local curvature in the occultation plane at the mean tangent point (MTP) location of the RO event.

The MTP location is defined as the geodetic (geographic) location on the WGS84 ellipsoid, where the straight-line path between transmitter and receiver touches this ellipsoid, i.e., where the straight-line tangent height is zero. This can be computed with very high accuracy at the sub-meter level (see (Scherllin-Pirscher et al. 2017), for more details on the geolocation accuracy of RO). Using the MTP location’s center of local curvature rather than the Earth’s center of mass as the origin is essential to ensure that the assumption of spherical symmetry, implicit in Equations (A.22) to (A.25), is accurately valid geometrically.

The impact parameter retrieval is solved iteratively, because it is impossible to rearrange Equations (A.22) to (A.24) into an explicit expression for the retrieval of the impact parameter, but it is mildly nonlinear and converges fast, in particular if the initial guess for $a_{t,i}$ is estimated from the previous level (starting at the top level with the straight-line impact parameter).

After the GO bending angle retrieval, the bending angles of all GNSS frequencies are interpolated to a common monotonic impact altitude grid z_a (item 2.6 in Figure 3.2), based on the monotonically sorted impact parameter grid of the leading channel, a_{t1} (i.e., $k = 1$).

For each element of z_a we get (item 2.3 in Figure 3.2)

$$z_{a,i} = a_{t,j1} - h_G - R_C, \quad (\text{A.26})$$

where j is the index of the elements of the sorted impact parameter grid a_{t1} . h_G is the geoid undulation (see (Scherllin-Pirscher et al. 2017), for a detailed discussion of its use in RO analysis), and R_C is the local radius of curvature of the RO event.

Because the impact parameter is only implicitly expressed in Equations (A.22)–(A.24), but GUM-type uncertainty propagation along Equations (3.2) and (3.4) requires an explicit measurement model, we make use of a linearization of the bending angle retrieval. We use the approach described by Melbourne et al. (1994), and applied to uncertainty propagation by Syndergaard (1999), for the propagation of the *estimated random uncertainty* from Doppler shift D_r to GO bending angle α_G (item 2.5 in Figure 3.2).

This linearization establishes a direct relation between random uncertainties of the Doppler shift $u_{D_r}^r$ and the uncertainties of the bending angle $u_{\alpha_G}^r$, using

$$u_{\alpha_G(t),i}^r \approx - \left(\frac{da_{\text{SL}}}{dt} \right)^{-1} \Big|_i u_{D_r(t),i}^r, \quad (\text{A.27})$$

where a_{SL} is the straight-line impact parameter. These bending angle uncertainties $u_{\alpha_G}^r$ are relative to the time grid as independent coordinate. To get the desired uncertainties with respect to the impact altitude grid z_a (introduced in Equation A.26), the uncertainties of the impact altitude z_a need to be transferred to the bending angle, so that the z_a grid can subsequently be considered free of error. Syndergaard (1999) showed that this can be done by replacing Equation (A.27) with

$$u_{\alpha_G(z_a),i}^r \approx - \left(\frac{da_t^T}{dt} \right)^{-1} \Big|_i u_{D_r(t),i}^r, \quad (\text{A.28})$$

where a_t^T is the “true” impact parameter. We use the forward-modeled impact parameter a_{tm} instead (i.e., adopt $a_{tm} = a_t^T$) and accept the additional error thus incurred, assuming it is smaller than the 2% relative error due to the linearization estimated by Melbourne et al. (1994). This is a reasonable assumption given the high quality of our forward-modeled profiles derived from ECMWF short-range forecast refractivity fields.

As a consequence we have to accept that the overall inaccuracy of our random uncertainty estimate cannot be brought below 2%. Therefore, to ensure that our simplified estimate does not underestimate the real uncertainty, we account for the linearization error by multiplying a factor $f_{u\alpha\text{lin}} = 1.02$ to the uncertainty of the retrieved GO bending angle

$$u_{\alpha_G,i}^r = f_{u\alpha\text{lin}} \cdot u_{\alpha_G(z_a),i}^r. \quad (\text{A.29})$$

In this way we acknowledge that although the calculation of the *state* of the bending angle does not make use of the linearization, and therefore the linearization does not increase the uncertainty of the *state* profile, it may increase the error in the uncertainty estimate itself.

Finally, the $u_{\alpha_G}^r$ profile is also interpolated to the common monotonic impact altitude grid z_a .

In the GO approximation, the bending angle values at each grid point only depend on the Doppler shift values of the same grid points; i.e., the existing correlations between

the errors at different levels are left unchanged: i.e., $\mathbf{R}_{\alpha G} = \mathbf{R}_{D_r}$. The covariance matrix can hence be calculated by recombining the Doppler shift correlation matrix with the propagated uncertainties (item 2.8 in Figure 3.2),

$$C_{\alpha G,ij} = u_{\alpha G,i}^r \cdot u_{\alpha G,j}^r \cdot R_{D_r,ij}. \quad (\text{A.30})$$

For the propagation of the *estimated systematic uncertainty* (item 2.4 in Figure 3.2) three types of potential systematic errors adding to the impact parameter uncertainty u_{at}^s , and consequentially the bending angle uncertainty $u_{\alpha G}^s$, are taken into account: systematic errors in the Doppler shift, i.e., $u_{D_r}^s$; systematic errors in the velocities of the satellites, i.e., $u_{v_R}^s$ and $u_{v_T}^s$; and systematic errors in the positions of the satellites, i.e., $u_{r_R}^s$ and $u_{r_T}^s$. The latter two orbit-borne types are interpreted as apparent systematic uncertainties (Sect. 3.3.2), while the excess phase-borne uncertainty $u_{D_r}^s$ is a basic systematic uncertainty.

For the propagation of these estimated systematic uncertainties to $u_{\alpha G}^s$, Equations (A.22)–(A.24) are linearized around the retrieved state quantities (serving as zero-order state), and no terms higher than first-order are kept. Then ϕ_R and ϕ_T in Equation (A.22) are substituted by the linearized versions of Equations (A.23) and (A.24), and the resulting equation is solved (level by level) for the impact parameter $a_{t,i} = f(D_{r,i}, r_{R,i}, r_{T,i}, v_{R,i}, v_{T,i})$, with $i = 1, 2, \dots, N$. Adopting the first-order deviations to represent the estimated systematic uncertainties, we obtain

$$u_{at,i}^s = \frac{1}{k_{at,i}} \sqrt{(u_{D_r,i}^s)^2 + (k_{v_R,i} \cdot u_{v_R,i}^s)^2 + (k_{r_R,i} \cdot u_{r_R,i}^s)^2 + (k_{v_T,i} \cdot u_{v_T,i}^s)^2 + (k_{r_T,i} \cdot u_{r_T,i}^s)^2}, \quad (\text{A.31})$$

where

$$\begin{aligned} k_{at,i} &= \left. \frac{\partial D_r}{\partial \phi_R} \right|_i \cdot \left. \frac{\partial \phi_R}{\partial a_t} \right|_i + \left. \frac{\partial D_r}{\partial \phi_T} \right|_i \cdot \left. \frac{\partial \phi_T}{\partial a_t} \right|_i = -v_{R,i} \\ &\quad \cdot \sin \phi_{Ri} \cdot \frac{1}{\sqrt{r_{R,i}^2 - a_{t,i}^2}} - v_{T,i} \cdot \sin \phi_{Ti} \cdot \frac{1}{\sqrt{r_{T,i}^2 - a_{t,i}^2}}, \\ k_{v_R,i} &= -\left. \frac{\partial D_r}{\partial v_R} \right|_i = -\cos \phi_{Ri}, \\ k_{v_T,i} &= -\left. \frac{\partial D_r}{\partial v_T} \right|_i = -\cos \phi_{Ti}, \\ k_{r_R,i} &= \left. \frac{\partial D_r}{\partial \phi_R} \right|_i \cdot \left. \frac{\partial \phi_R}{\partial r_R} \right|_i = \frac{v_{R,i} \cdot \sin \phi_{Ri} \cdot a_{t,i}}{r_{R,i} \sqrt{r_{R,i}^2 - a_{t,i}^2}}, \\ k_{r_T,i} &= \left. \frac{\partial D_r}{\partial \phi_T} \right|_i \cdot \left. \frac{\partial \phi_T}{\partial r_T} \right|_i = \frac{v_{T,i} \cdot \sin \phi_{Ti} \cdot a_{t,i}}{r_{T,i} \sqrt{r_{T,i}^2 - a_{t,i}^2}}. \end{aligned} \quad (\text{A.32})$$

A number of simplifications have been made to arrive at this result. First, the last term in Equation A.22 is disregarded since errors in the positions are assumed to be constant with respect to the short time duration of an RO event; remaining errors $\Delta\dot{r}_{\text{RT}}$ after taking the derivative are therefore of higher order. Next, orbit position and velocity uncertainties are both assumed to be constant within the short duration of an event, and the velocity uncertainties obtained are interpreted as uncertainties along the direction of the velocity vector. Consequentially, the uncertainty is also projected along with the vector into the ray path direction. A more conservative estimation (which we consider overly conservative in context) would interpret the uncertainties as ellipsoids at the velocity vectors' heads and would hence take the full magnitude of the uncertainties along the ray path direction (without projection).

Furthermore, since all error sources (the processing of the occultation tracking data and the POD for transmitter and receiver) are essentially independent from each other, the different input uncertainties are assumed to be uncorrelated. Finally, we reasonably assumed the errors of the angle between the position and velocity vectors (η) to be negligible ($u_\eta^s \approx 0$) for the purpose here, for both the transmitter and receiver.

In order to finally derive the systematic uncertainty of the bending angle from the impact parameter's uncertainty, we continue with a linearization of Equation (A.25) and arrive at

$$u_{\alpha\text{G},i}^s = \sqrt{\left(u_{\theta\text{RT},i}^s\right)^2 + \left(k_{at,i} \cdot u_{at,i}^s\right)^2 + \left(k_{r\text{R},i} \cdot u_{r\text{R},i}^s\right)^2 + \left(k_{r\text{T},i} \cdot u_{r\text{T},i}^s\right)^2}, \quad (\text{A.33})$$

where

$$\begin{aligned} k_{at,i} &= \left. \frac{\partial\alpha}{\partial a_t} \right|_i = \frac{1}{\sqrt{r_{\text{R},i}^2 + a_{t,i}^2}} + \frac{1}{\sqrt{r_{\text{T},i}^2 + a_{t,i}^2}}, \\ k_{r\text{R},i} &= \left. \frac{\partial\alpha}{\partial r_{\text{R}}} \right|_i = -\frac{a_{t,i}}{r_{\text{R},i}\sqrt{r_{\text{R},i}^2 - a_{t,i}^2}}, \\ k_{r\text{T},i} &= \left. \frac{\partial\alpha}{\partial r_{\text{T}}} \right|_i = -\frac{a_{t,i}}{r_{\text{T},i}\sqrt{r_{\text{T},i}^2 - a_{t,i}^2}}. \end{aligned} \quad (\text{A.34})$$

In practice we separately calculate the basic and apparent systematic uncertainty estimates ($u_{\alpha\text{G}}^b$ from the first RHS terms in Equations A.31 and A.33, $u_{\alpha\text{G}}^a$ from the orbit-borne terms) and afterwards obtain $u_{\alpha\text{G}}^s$ as a combined result, in order to enable separate propagation in subsequent processing steps.

The *resolution* profile remains unaffected by the bending angle retrieval, since the level-by-level approach of the algorithm does not create extra correlation and further vertical smoothing, so that $\tau_{\alpha\text{G}} = \tau_{D_r}$ and $w_{\alpha\text{G}} = w_{D_r}$.

WO bending angle retrieval

After the GO bending angle, the WO bending angle *state* profile $\alpha_W(z_a)$ is retrieved (item 2.7 in Figure 3.2) from excess phase profile $L_r(t)$ (and its uncertainties) and the amplitude profile $A_r(t)$ (and uncertainties) in a WO retrieval following Gorbunov and Kirchengast (2015) and Gorbunov and Kirchengast (2018). Along with the state profile, the *systematic uncertainty* profile $u_{\alpha_W}^s$, the covariance matrix \mathbf{C}_{α_W} , and the *resolution* profile w_{α_W} are derived.

The covariance matrix \mathbf{C}_{α_W} is then decomposed to *random uncertainty* profile $u_{\alpha_W}^r$ and correlation matrix \mathbf{R}_{α_W} in the same form as done above for \mathbf{C}_{D_r} (Equations A.19 and A.20) and \mathbf{C}_{L_F} (Equations A.7 and A.8). The estimated systematic uncertainty $u_{\alpha_W}^s$ is composed of a basic systematic uncertainty $u_{\alpha_W}^b$, propagated through the wave-optical retrieval from the excess phase uncertainty $u_{L_r}^s$, and an apparent systematic uncertainty $u_{\alpha_W}^a$, estimated in the lower troposphere as residual bias uncertainty of a regression-based boundary layer bias correction (Gorbunov and Kirchengast 2018).

The WO bending angle retrieval algorithm and the associated uncertainty propagation algorithm are not explicitly described here; the reader is referred to Gorbunov and Kirchengast (2015) and Gorbunov and Kirchengast (2018). However, we have prepared the merging with the WO bending angle variables (they will be actually merged in when the WO tests within the rOPS are complete), which is described next.

Merging of GO and WO bending angle profiles

The α_W profile, prepared on the common grid z_a , and the α_G profile are merged over an upper-tropospheric transition range (item 2.9 in Figure 3.2). The gradual transition, weighted by a symmetric half-sine function, has a defined impact altitude transition of half-width $\Delta z_a^{\text{GW}} = 2$ km around transition altitude z_a^{GW} , allowed within 9 km to 14 km, estimated from α_G data quality. The resulting merged bending angle profile α_M is

$$\alpha_{M,i} = \gamma_i \cdot \alpha_{G,i} + (1 - \gamma_i) \cdot \alpha_{W,i}, \quad (\text{A.35})$$

where the weighting profile γ is formulated as

$$\gamma_i = \begin{cases} 1 & \text{for } z_{a,i} \geq z_a^{\text{GW}} + \Delta z_a^{\text{GW}} \\ 0.5 \cdot \left[\sin \left(\frac{\pi}{2} \cdot \frac{z_{a,i} - z_a^{\text{GW}}}{\Delta z_a^{\text{GW}}} \right) + 1 \right] & \text{for } |z_{a,i} - z_a^{\text{GW}}| < \Delta z_a^{\text{GW}} \\ 0 & \text{for } z_{a,i} \leq z_a^{\text{GW}} - \Delta z_a^{\text{GW}}. \end{cases} \quad (\text{A.36})$$

To determine the random uncertainties for the merged GO–WO input bending angle, we need to merge the covariance matrices of both bending angles.

We can assume both incoming covariance matrices \mathbf{C}_{α_G} and \mathbf{C}_{α_W} are provided on the common monotonic target grid z_a (i.e., also the WO uncertainties and correlations are

interpolated to this common grid before the merger). We further can reasonably assume that there are no cross-correlations between GO and WO errors, given the very different retrieval schemes. Based on this we can compose the covariance matrix of the merged bending angle profile, $\mathbf{C}_{\alpha M}$ (item 2.10 in Figure 3.2) as follows. Outside the merging zone (i.e., outside of $z_a^{\text{GW}} \pm \Delta z_a^{\text{GW}}$) we can assign

$$C_{\alpha M, ij} = \begin{cases} C_{\alpha G, ij} & \text{for } z_{a\text{Top}} > z_{a, i} > z_a^{\text{GW}} + \Delta z_a^{\text{GW}} \\ & \text{and } z_{a\text{Top}} > z_{a, j} > z_a^{\text{GW}} + \Delta z_a^{\text{GW}} \\ C_{\alpha W, ij} & \text{for } z_a^{\text{GW}} - \Delta z_a^{\text{GW}} > z_{a, i} > z_{a\text{Bot}} \\ & \text{and } z_a^{\text{GW}} - \Delta z_a^{\text{GW}} > z_{a, j} > z_{a\text{Bot}} \\ 0 & \text{for } z_{a\text{Top}} > z_{a, i} > z_a^{\text{GW}} + \Delta z_a^{\text{GW}} \\ & \text{and } z_a^{\text{GW}} - \Delta z_a^{\text{GW}} > z_{a, j} > z_{a\text{Bot}} \\ 0 & \text{for } z_a^{\text{GW}} - \Delta z_a^{\text{GW}} > z_{a, i} > z_{a\text{Bot}} \\ & \text{and } z_{a\text{Top}} > z_{a, j} > z_a^{\text{GW}} + \Delta z_a^{\text{GW}} \end{cases}, \quad (\text{A.37})$$

while within the merging zone we can assign

$$C_{\alpha M, ij} = \gamma_i \gamma_j C_{\alpha G, ij} + (1 - \gamma_i)(1 - \gamma_j) C_{\alpha W, ij}, \quad (\text{A.38})$$

wherein i is understood such that $z_a^{\text{GW}} + \Delta z_a^{\text{GW}} > z_{a, i} > z_a^{\text{GW}} - \Delta z_a^{\text{GW}}$ and j such that $z_{a\text{Top}} > z_{a, j} > z_{a\text{Bot}}$.

Because of the symmetry of the covariance matrix, the covariance elements in the merging zone orthogonal to the one above, i.e., for $z_a^{\text{GW}} + \Delta z_a^{\text{GW}} > z_{a, j} > z_a^{\text{GW}} - \Delta z_a^{\text{GW}}$ and $z_{a\text{Top}} > z_{a, i} > z_{a\text{Bot}}$, are calculated according to the same formula.

Due to the linear relation between α_M , α_G , and α_W , expressed by Equation (A.35), a bias $u_{\alpha G}^s$ in the GO bending angle and a bias $u_{\alpha W}^s$ in the WO bending angle can be combined linearly as well, and we can compute the *estimated systematic uncertainty* of the merged bending angle $u_{\alpha M}^s$ according to (item 2.9 in Figure 3.2)

$$u_{\alpha M, i}^s = \gamma_i \cdot u_{\alpha G, i}^s + (1 - \gamma_i) \cdot u_{\alpha W, i}^s. \quad (\text{A.39})$$

In practice this formulation is again applied separately for the basic and apparent systematic uncertainty estimates, afterwards obtaining the $u_{\alpha M}^s$ profile as a combined result, in order to allow separate propagation in subsequent processing steps.

The *resolution* profile of the bending angle, $w_{\alpha M}$, is equal to the GO resolution $w_{\alpha G}$ above $z_a^{\text{GW}} + \Delta z_a^{\text{GW}}$, equal to the WO resolution $w_{\alpha W}$ below $z_a^{\text{GW}} - \Delta z_a^{\text{GW}}$ and has a transition with transition weight γ_i in between, again following the linear formulation such as in Equations (A.35) and (A.39).

Because the integration and testing of the uncertainty propagation through the rOPS WO bending angle retrieval are currently still ongoing, as noted in Sect. A.1.2 above, the examples shown in this study are all GO-only; i.e., only the GO retrieval is performed. The merging algorithm as described is ready to include the WO bending angles, however.

A.1.3 Atmospheric bending angle derivation

In order to retrieve the atmospheric bending angle profile α_r , ionospheric effects need to be corrected for, using the retrieved bending angles from each transmitter frequency channel. Since the only GNSS constellation currently used for RO is the GPS – except for recent initial data from the Chinese GNOS instrument using BeiDou signals (Bai et al. 2018; Liao et al. 2016) – the data characteristics of the GPS case (with $k \in \{1, 2\}$, $f_{T1} = 1.57542$ GHz, and $f_{T2} = 1.22760$ GHz) are in the prime focus of this section.

This concerns in particular special provisions for the minor (L2) channel noise filtering and its tropospheric extrapolation. In general the algorithms are applicable for any of the available GNSS systems, however; if the minor channel (f_{T2}) delivers similar data quality to the major one (f_{T1}), the special provisions for the former will practically have no effect.

Adaptive low-pass filtering and minor-channel extrapolation

Before applying the dual-frequency ionospheric correction, the merged bending angle *state* profiles $\alpha_{M,k}(z_a)$ at the common z_a grid are filtered with further BWS low-pass filter operations, and the minor channel is extrapolated.

For α_{M1} the filter is set to the same cutoff frequency as the basic BWS filter preceding the Doppler derivation (i.e., $f_{c1} = 2.5$ Hz), ensuring a reliable reference resolution and basic smoothness of the whole merged profile. For filtering of α_{M2} a (GPS L2) noise-minimization algorithm is used, following the approach of Sokolovskiy et al. (2009) for optimal filtering for ionospheric correction. We search for minimized noise employing a flexible cutoff frequency $f_{c2} \in \{2.5, 2, 10/7, 1, 5/7, 0.5 \text{ Hz}\}$, corresponding to using cutoff periods τ_{c2} from 0.4 to 2 s and sample widths of $M = 40$ to $M = 200$ (for BWS filter design details see Sect. A.1.1).

We adopt the cutoff frequency f_{c2} for α_{M2} filtering that minimizes the noise fluctuations of the ionosphere-corrected atmospheric bending angle delta profile $\delta\alpha_{rm}^{fc2}(z_a) = \alpha_r^{fc2}(z_a) - \alpha_m(z_a)$ when evaluated over the mesospheric altitude range between 50 and 70 km (similar to the functional minimization of (Sokolovskiy et al. 2009); Equation 4 therein). At these high altitudes the residual atmospheric mean signal after subtraction of the forward-modeled signal $\alpha_m(z_a)$ is very small ($< 0.03\text{--}0.3 \mu\text{rad}$), and therefore the noise level representative for the given RO event is well quantifiable.

The weight matrix of the BWS filter, $\mathbf{A}_k^{\text{BWS}}$, is determined for both frequencies analogously to Equations (A.3) to (A.5). When using the baseband approach with model profile α_m to create the delta profile $\delta\alpha_{Mm}$ with elements

$$\delta\alpha_{Mmi,k} = \alpha_{Mi,k} - \alpha_{mi}, \quad (\text{A.40})$$

the filtered bending angle is then (item 3.1 in Figure 3.2)

$$\alpha_{Fi,k} = \alpha_{mi} + \sum_{j=0}^N A_{ij,k}^{\text{BWS}} \cdot \delta\alpha_{Mmj,k}, \quad (\text{A.41})$$

where $i, j \in \{1, 2, \dots, N\}$ and $k \in \{1, 2\}$.

Due to the stronger power of the L1 signal for (most of) the GPS satellites, the GPS signals of both frequencies are not of the same quality, and the L2 data (for those satellites where encrypted and hence power-degraded L2 signals are transmitted) do not reach down as far as the L1 data (i.e., $z_{\text{aBot2}} > z_{\text{aBot1}}$). If due to this reason α_{F2} does not reach down as far as α_{F1} and $z_{\text{aBot2}} \leq z_{\text{aBot2Max}}$ (with z_{aBot2Max} currently set to 15 km), a *tropospheric bending angle extrapolation* (TBAE) is applied in order to artificially extend α_{M2} to also reach down to z_{aBot1} (item 3.3 in Figure 3.2).

Briefly summarized, this TBAE is currently implemented as follows. A linear gradient profile for the difference profile between the two bending angles, $\alpha_{F12} = (\alpha_{F1} - \alpha_{F2})$, is estimated by a least squares fit over a sufficiently wide impact altitude range from z_{aBot2} upward (as wide as the extrapolation range, at least 10 km). This gradient profile is then linearly extended down to z_{aBot1} and subtracted from α_{F1} , to obtain the extrapolated part of α_{F2} from z_{aBot2} to z_{aBot1} . If $z_{\text{aBot2}} > z_{\text{aBot2Max}}$, then no TBAE is performed since the extrapolation range is considered too large. Details are provided by Kirchengast et al. (2017b), where the most recent version of the atmospheric bending angle derivation is described that includes this α_{F12} extrapolation in a further advanced form.

For the propagation of the *estimated random uncertainty* we get (item 3.2 in Figure 3.2)

$$\mathbf{C}_{\alpha F,k} = \mathbf{A}_k^{\text{BWS}} \cdot \mathbf{C}_{\alpha M,k} \cdot \left(\mathbf{A}_k^{\text{BWS}}\right)^{\text{T}} \quad (\text{A.42})$$

for the bending angle error covariance matrices of the leading ($k = 1$) and minor ($k = 2$) channel.

If a TBAE is applied to α_{F2} , the random uncertainty of α_{F2} below z_{aBot2} is equal to that of α_{F1} , because the noise is “copied” from α_{F1} since the linear gradient profile from fitting α_{F12} is noise-free. As a consequence, in these cases, we set the matrix elements of $\mathbf{C}_{\alpha F2}$ to (item 3.4 in Figure 3.2)

$$C_{\alpha F2,ij} = \begin{cases} \mathbf{C}_{\alpha F2,ij} & \text{for } z_{\text{aTop}} > z_{\text{a},i} > z_{\text{aBot2}} \\ & \text{and } z_{\text{aTop}} > z_{\text{a},j} > z_{\text{aBot2}} \\ \mathbf{C}_{\alpha F1,ij} & \text{for } z_{\text{aBot2}} > z_{\text{a},i} > z_{\text{aBot1}} \\ & \text{and } z_{\text{aBot2}} > z_{\text{a},j} > z_{\text{aBot1}} \\ 0 & \text{for } z_{\text{aBot2}} > z_{\text{a},i} > z_{\text{aBot1}} \\ & \text{and } z_{\text{aTop}} > z_{\text{a},j} > z_{\text{aBot2}} \\ 0 & \text{and } z_{\text{aTop}} > z_{\text{a},i} > z_{\text{aBot2}} \\ & \text{for } z_{\text{aBot2}} > z_{\text{a},j} > z_{\text{aBot1}} \end{cases}. \quad (\text{A.43})$$

$\mathbf{C}_{\alpha F1}$ and $\mathbf{C}_{\alpha F2}$ can then be decomposed as needed into $u_{\alpha F1}^r$, $\mathbf{R}_{\alpha F1}$, and $u_{\alpha F2}^r$, $\mathbf{R}_{\alpha F2}$, respectively. Kirchengast et al. (2017b) describe the most recent version consistent with a further advanced form of the TBAE, where the separate assignments according to Equation (A.43) are no longer needed.

The *estimated systematic uncertainties* $u_{\alpha M,k}^s$ (in practice the basic and apparent systematic uncertainty estimates separately) are filtered with the same filter settings as for the state profiles (item 3.1 in Figure 3.2) and are thus obtained in the form

$$u_{\alpha Fi,k}^s = \sum_{j=0}^N A_{ij,k}^{\text{BWS}} \cdot u_{\alpha Mj,k}^s. \quad (\text{A.44})$$

Since these are smooth profiles, they are marginally changed by this low-pass filtering. The systematic uncertainty component contributed by the TBAE to the estimated systematic uncertainty is added after the ionospheric correction (see next subsection).

As for the basic low-pass filtering of excess phases (Sect. A.1.1), the *resolution* profiles of the filtered bending angles $w_{\alpha F1}$ and $w_{\alpha F2}$ are determined by the cutoff frequencies f_{c1} and f_{c2} of the BWS filters, following Equations (A.11) and (A.13).

Ionospheric correction

Based on the filtered and sometimes extrapolated *state* profiles α_{F1} and α_{F2} , the ionospheric refractive effects are corrected for by the standard dual-frequency correction of bending angles (Vorob'ev and Krasil'nikova 1994) used in the f_{T1} - f_{T2} difference profile form (Sokolovskiy et al. 2009) (item 3.5 in Figure 3.2). For the elements of the retrieved atmospheric bending angle profile α_r we thus get

$$\alpha_{r,i} = \alpha_{F1,i} + \gamma_{f_{T12}} \cdot \delta\alpha_{F12,i}, \quad (\text{A.45})$$

where

$$\delta\alpha_{F12,i} = \alpha_{F1,i} - \alpha_{F2,i} \quad (\text{A.46})$$

and

$$\gamma_{f_{T12}} = \frac{f_{T2}^2}{f_{T1}^2 - f_{T2}^2}. \quad (\text{A.47})$$

Propagated through the operator of the ionospheric correction (Equation A.45, currently used here in the classical form with f_{T1} and f_{T2} terms), the *estimated random uncertainty* of the resulting atmospheric bending angle, expressed by the error covariance matrix $\mathbf{C}_{\alpha r}$ (item 3.6 in Figure 3.2), is obtained as

$$\mathbf{C}_{\alpha r} = (1 + \gamma_{f_{T12}})^2 \mathbf{C}_{\alpha F1} + \gamma_{f_{T12}}^2 \mathbf{C}_{\alpha F2}. \quad (\text{A.48})$$

$\mathbf{C}_{\alpha r}$ can then also be decomposed into $u_{\alpha r}^r$ and $\mathbf{R}_{\alpha r}$ with the usual equations (cf., e.g., Equations A.19 and A.20).

Equation (A.45) is as well applied to propagate the *estimated systematic uncertainty* (in practice the basic and apparent systematic uncertainty estimates separately) through the ionospheric correction using (item 3.5 in Figure 3.2)

$$u_{\alpha_r,i}^s = u_{\alpha_{F1},i}^s + \gamma_{fT12} \cdot (u_{\alpha_{F1},i}^s - u_{\alpha_{F2},i}^s), \quad (\text{A.49})$$

where it is assumed that the systematic errors in α_{F1} and α_{F2} are positively correlated, i.e., have the same sign, and the associated uncertainty estimates are hence subtracted from one another (as the bending angles are in Equation A.46). This assumption is reasonable, since the same sources of non-ionospheric systematic effects apply to both frequency channels (Doppler shift, orbit velocity, and orbit position uncertainties).

In the case of TBAE, Equation (A.49) needs to be supplemented below z_{aBot2} , since additional uncertainties $u_{\alpha_{2\text{TE}}}^s$ arise from the errors made in the fitting parameters and in the extrapolation model (linear extrapolation) of the TBAE. Hence, for the range $z_{\text{aBot2}} > z_{\text{a},i} \geq z_{\text{aBot1}}$,

$$u_{\alpha_r,i}^s = u_{\alpha_r}^s(z_{\text{aBot2}}) + u_{\alpha_{2\text{TE},i}}^s, \quad (\text{A.50})$$

with $u_{\alpha_{2\text{TE}}}^s$ being the conservative estimate for additional (apparent) systematic uncertainty within the extrapolated impact altitude range. We set $u_{\alpha_{2\text{TE}}}^s$ to zero at z_{aBot2} and linearly increase it from there downwards with a gradient of $1 \mu\text{rad}$ per 10 km (an experience-based best guess; cf. (Scherllin-Pirscher et al. 2011a; Scherllin-Pirscher et al. 2011b), who also address aspects of such tropospheric extrapolation in their discussions of error sources). It is interpreted as an apparent systematic uncertainty estimate, since due to the linear fit-based TBAE construction its event-to-event bias character will be essentially random (Scherllin-Pirscher et al. 2011a).

Also, the ionospheric correction currently applied in the rOPS is just a first-order correction, which will leave higher-order residual ionospheric errors in α_r (e.g., Danzer et al. 2013; Healy and Culverwell 2015; Liu et al. 2013; Liu et al. 2015; Syndergaard 2000). The uncertainty from higher-order *residual ionospheric biases* (RIBs), u_{RIB}^s , is therefore added to the propagated (basic) systematic uncertainty. u_{RIB}^s is interpreted as basic systematic uncertainty, since the higher-order ionospheric residuals may not vanish in ensemble-of-events averaging. The other non-ionospheric sources of systematic errors and the RIBs can be reasonably assumed to be uncorrelated. The total estimated systematic uncertainty of the retrieved atmospheric bending angle α_r hence is

$$u_{\alpha_r,i}^s = \sqrt{(u_{\alpha_r,i}^s)^2 + (u_{\text{RIB}}^s)^2}. \quad (\text{A.51})$$

Based on previous studies (e.g., Danzer et al. 2013; Danzer et al. 2015; Liu et al. 2013; Liu et al. 2015), u_{RIB}^s is taken to be constant along the entire profile and is estimated to amount to $0.05 \mu\text{rad}$. These last two components, $u_{\alpha_{2\text{TE}}}^s$ and u_{RIB}^s , are indicated as item 3.7 in Figure 3.2. It is clear that this initial systematic uncertainty estimation can

be significantly improved by future dedicated work on better quantifying and (if suitable) correcting for the systematic uncertainty components.

The *resolution* of the retrieved bending angle, w_{α_r} , essentially corresponds to the higher resolution of the two bending angle profiles α_{F1} and α_{F2} , and thus generally closely matches $w_{\alpha_{F1}}$ in most cases. As a simple but robust and suitable estimate, assuming that the resolutions of α_r and α_{F1} scale in the same way as the correlation lengths l_{α_r} and $l_{\alpha_{F1}}$ (derived from \mathbf{R}_{α_r} and $\mathbf{R}_{\alpha_{F1}}$ as described for \mathbf{R}_{LF} in Equation A.9), we compute w_{α_r} as

$$w_{\alpha_r,i} = \frac{l_{\alpha_r,i}}{l_{\alpha_{F1},i}} \cdot w_{\alpha_{F1},i}. \quad (\text{A.52})$$

In concluding, we note that the atmospheric bending angle derivation algorithms used in this study – i.e., the adaptive filtering, TBAE, and ionospheric correction parts as described in this section – have recently received further advancement towards a form fully based on the combination of α_{F1} and the difference profile α_{F12} (rather than of α_{F1} and α_{F2}), more aligned with the concept of Sokolovskiy et al. (2009). A detailed description of this most recent version is found in Kirchengast et al. (2017b).

A.2 Variance propagation for comparison

The full covariance propagation applied to propagate random uncertainties requires numerically “expensive” matrix operations, and therefore considerable efforts were made to seize opportunities for reducing the number of numerical operations (e.g., by only calculating with those elements of the band matrix \mathbf{A}^{BWS} which lie within the width of the filter window).

However, as demonstrated in Sect. 3.5, simplification to a mere variance propagation (i.e., only considering the diagonal elements of the covariance matrices) is not reasonably possible because it leads to an unacceptable overestimation of random uncertainties. This overestimation occurs since the influence of the covariance elements – and thus for example the partially compensating impact of the negative side peaks in the correlation functions – is disregarded.

Here we state the two equations used to obtain the variances-only propagation results shown for comparison purposes in Figure 3.9: the estimated random uncertainty was propagated through the BWS filter using

$$(u_{\alpha_{Fi},k}^r)^2 = \sum_{j=0}^N (A_{ij,k}^{\text{BWS}})^2 \cdot (u_{\alpha_{Mj},k}^r)^2, \quad (\text{A.53})$$

and subsequently through the ionospheric correction using

$$(u_{\alpha_r,i}^r)^2 = (1 + \gamma_{fT12})^2 \cdot (u_{\alpha_{F1},i}^r)^2 + \gamma_{fT12}^2 \cdot (u_{\alpha_{F2},i}^r)^2. \quad (\text{A.54})$$

**Integrating uncertainty propagation in GNSS radio
occultation retrieval: from bending angle to dry-air
atmospheric profiles**

B.1 Algorithm Description

In this Appendix the L2a uncertainty propagation algorithm implemented in the rOPS is introduced, following the retrieval chain step by step, starting with α_r and α_b as input and proceeding to α_s , N , p_d and finally T_d . A fully detailed description is provided by Kirchengast et al. (2017b).

If not stated otherwise, elements of the vector-type profiles are addressed using subscript i (with $i = 1, \dots, n$), and optionally j (with $j = 1, \dots, n$), running from top downward towards the bottom of the profile, where n is the total number of vertical grid levels.

B.1.1 High Altitude Initialization

In order to perform the statistical optimization we first calculate the weighting matrix \mathbf{A}^{SO} from $z_{a\text{TopMax}}$ to $z_{a\text{ST}} - \Delta z_{a\text{ST}}$ (Item 1.3 in Figure 4.2), following Rodgers (2000, Equation 2.31 therein) (see also Healy (2001) for an early application in RO statistical optimization),

$$\mathbf{A}^{\text{SO}} = \mathbf{C}_{\alpha_b}(\mathbf{C}_{\alpha_b} + \mathbf{C}_{\alpha_r})^{-1}. \quad (\text{B.1})$$

Then the **state retrieval** is performed (Item 1.5 in Figure 4.2), using

$$\delta\alpha_{\text{sb}} = \mathbf{A}^{\text{SO}} \delta\alpha_{\text{rb}}, \quad (\text{B.2})$$

where $\delta\alpha_{\text{rb}} = \alpha_r - \alpha_b$, following the reasoning provided by Li et al. (2013) and Li et al. (2015). The bending angle is only statistically optimized down to $z_{a\text{ST}}$ and set to α_r

below. A linear combination is used for transition around this merging point. Thus,

$$\alpha_s = \begin{cases} \alpha_b + \delta\alpha_{sb} & \text{for } z_{a\text{TopMax}} \geq z_a > z_{a\text{ST}} + \Delta z_{a\text{ST}} \\ \gamma \cdot (\alpha_b + \delta\alpha_{sb}) + (1 - \gamma) \cdot \alpha_r & \text{for } z_{a\text{ST}} + \Delta z_{a\text{ST}} \geq z_a \geq z_{a\text{ST}} - \Delta z_{a\text{ST}} \\ \alpha_r & \text{for } z_{a\text{ST}} - \Delta z_{a\text{ST}} > z_a \geq z_{\text{BotM}} \end{cases} \quad (\text{B.3})$$

The weight $\gamma(z_a)$ of the applied continuous and differentiable half-sine-weighted transition is defined as (Li et al. 2015)

$$\gamma(z_a) = 0.5 \cdot \left[\sin \left(\frac{\pi}{2} \cdot \frac{z_a - z_{a\text{ST}}}{\Delta z_{a\text{ST}}} \right) + 1 \right]. \quad (\text{B.4})$$

The **random uncertainty** is calculated (Item 1.6 in Figure 4.2) using

$$\mathbf{C}_{\alpha_s} = \mathbf{A}^{\text{SO}} \cdot \mathbf{C}_{\alpha_r} \quad (\text{B.5})$$

for $z_{a\text{TopMax}} \geq z_a > z_{a\text{ST}} + \Delta z_{a\text{ST}}$ and

$$\mathbf{C}_{\alpha_s} = \mathbf{C}_{\alpha_r} \quad (\text{B.6})$$

below $z_{a\text{ST}} - \Delta z_{a\text{ST}}$. For $z_{a\text{BotS}} + \Delta z_{a\text{ST}} \geq z_{a,i} \geq z_{a\text{BotS}} - \Delta z_{a\text{ST}}$ we again apply the half-sine-weighted transition, and the elements of the covariance matrix of α_s are

$$C'_{\alpha_s, i, j < i} = \gamma_i \cdot C_{\alpha_s, i, j < i} + (1 - \gamma_i) \cdot C_{\alpha_r, i, j < i}. \quad (\text{B.7})$$

Making use of the symmetry of the matrix, we apply the same formula for elements $C_{\alpha_s, i < j, j}$ for $z_{a\text{BotS}} + \Delta z_{a\text{ST}} \geq z_{a,j} \geq z_{a\text{ST}} - \Delta z_{a\text{ST}}$ and with i and j exchanged.

For writing out, $C_{\alpha_s, ij}$ is split up into uncertainty profile and error correlation matrix,

$$u_{\alpha_s, i}^r = \sqrt{C_{\alpha_s, ii}}, \quad (\text{B.8})$$

$$R_{\alpha_s, ij} = \frac{C_{\alpha_s, ij}}{u_{\alpha_s, i}^r \cdot u_{\alpha_s, j}^r}. \quad (\text{B.9})$$

The **systematic uncertainty** $u_{\alpha_s}^s(z_a)$ is calculated by applying the state-retrieval, i.e. Equations B.2 and B.3, but using $\delta u_{\alpha_r b}^s = u_{\alpha_r}^s - u_{\alpha_b}^s$ instead of $\delta\alpha_{rb}$ as input. Below $z_{a\text{ST}} - \Delta z_{a\text{ST}}$, $u_{\alpha_s}^s(z_a)$ is set to $u_{\alpha_r}^s(z_a)$.

According to Li et al. (2015), the **observation-to-background weighting ratio** of α_s is defined as

$$r_{\alpha}^{\text{obw}} = \frac{(u_{\alpha_b}^r)^2}{(u_{\alpha_r}^r)^2 + (u_{\alpha_b}^r)^2}, \quad (\text{B.10})$$

(Item 1.4 in Figure 4.2) and provides insight into whether the observation or the background information dominates α_s . It can be interpreted as weighting function of a linear combination of the observed and background bending angle with elements $\alpha_{s,i} = r_{\alpha,i}^{\text{obw}} \alpha_{r,i} + (1 - r_{\alpha,i}^{\text{obw}}) \alpha_{b,i}$, if the statistical optimization were performed based solely on the diagonal elements of the covariance matrices (i.e. ignoring correlations).

Below $z_{a\text{ST}} - \Delta z_{a\text{ST}}$, r_{α}^{obw} is set to 1. Around the merging point, again a half-sine-weighted transition is applied. Upon output it is multiplied by 100 and provided in %.

B.1.2 Refractivity Retrieval

The next step in the **state retrieval** is the refractivity retrieval, where refractivity $N_r(z)$ is derived from $\alpha_s(z_a)$, by solving an Abel integral (Item 2.1 in Figure 4.2). α_s is given with respect to impact altitude z_a and thus impact parameter a_z (the elements $a_{z,i}$ are provided on the same grid points as the $z_{a,i}$). Elements $N_{r,i}$ are first calculated at grid points $a_{z,i}$ (and thus also $z_{r,i}$) and then interpolated to z_i (see Table 4.3 for the definition of the vertical grids).

A major innovation in the rOPS is the so-called 'baseband approach' (Kirchengast et al. 2016a), in which a zero-order profile α_m is subtracted from α_s , and only the delta-profile $\delta\alpha_{sm}$ is used as input for the numerical Abel operator. α_m is chosen such that its Abel integral can be solved analytically for the logarithm of the refractive index $(\ln n)_m$ (Kirchengast et al. 2017b).

We calculate the transform matrix of the Abel integral \mathbf{A}^{Abl} using the first-order solution of the powerful piecewise analytical rule recently developed by Syndergaard and Kirchengast (2016),

$$\mathbf{A}_{ij}^{\text{Abl}} = \begin{cases} 0 & \text{for } i < j \\ \frac{2}{3\pi} p_{k-1} & \text{for } i = j \\ \frac{2}{3\pi} \left(\frac{p_{k-1}^2}{p_{k-1} + p_k} - \frac{p_{k+1}^2}{p_{k+1} + p_k} \right) & \text{for } i > j \end{cases}, \quad (\text{B.11})$$

where $p_k = \text{arccosh}(a_{z,i+k}/a_{z,i})$ for a given i , and $k = j - i$.

The refractivity profile can then be derived by applying

$$\delta(\ln n)_{\text{rm}} = \mathbf{A}^{\text{Abl}} \cdot \delta\alpha_{sm}, \quad (\text{B.12})$$

where $\delta(\ln n)_{\text{rm}}$ is the delta-profile of the logarithm of the refractive index. The analytically calculated zero-order profile of the logarithm of the refractive index $(\ln n)_m$ is added again, and the refractive index n_r derived by exponentiation,

$$n_r = \exp[(\ln n)_r] = \exp[(\ln n)_m + \delta(\ln n)_{\text{rm}}], \quad (\text{B.13})$$

and refractivity N_r follows as

$$N_r = 10^6(n_r - 1). \quad (\text{B.14})$$

Then, using n_r , the corresponding altitude z_r can be calculated for each vertical grid point of z_a (Item 2.6 in Figure 4.2), using

$$z_r = \frac{z_a + h_G + R_C}{n_r} - (h_G + R_C), \quad (\text{B.15})$$

with geoid undulation, h_G , and the local radius of curvature of the occultation event, R_C .

As addressed in Table 4.3, z_a is calculated from z and N_b (from the ECMWF forecast fields). Since N_b is close to N_r , the grid points of z_r are almost identical to those of z . Using z_r all retrieved state and uncertainty quantities (including the bending angle profiles) can be interpolated to the equidistant preset 100 m grid z (Item 2.7 in Figure 4.2).

The zero-order profile is chosen as a smooth profile independently from the measurement, and thus the **random uncertainty** is assumed to remain available on the delta-profile. The covariance matrix of $(\ln n)_r$ is therefore

$$\mathbf{C}_{nr} = \mathbf{A}^{\text{Abl}} \mathbf{C}_{\text{os}} (\mathbf{A}^{\text{Abl}})^T. \quad (\text{B.16})$$

Using $(\ln n)_r \approx (n_r - 1)$ given $|n_r - 1| \ll 1$ in Equation B.14, shows that $N_r \approx 10^6(\ln n)_r$. Hence the random uncertainty of $(\ln n)_r$ can be propagated to N_r , by simply using

$$\mathbf{C}_{N_r} = (10^6)^2 \cdot \mathbf{C}_{nr}. \quad (\text{B.17})$$

These two propagation steps are represented by Item 2.2 in Figure 4.2. The random uncertainty profile $u_{N_r}^r$ is derived from \mathbf{C}_{N_r} by applying the equivalent of Equation B.8. $u_{N_r}^r$ is still given with respect to the z_a grid. In order to know the uncertainties with respect to the MSL altitude grid z_r , the uncertainties of z_r with respect to z_a , i.e., $u_{z_r(z_a)}^r$ are evaluated and then transferred to the uncertainties of N_r , to get $u_{N_r(z_r)}^r$ (Item 2.5 in Figure 4.2). Based on Equation B.15 and the reasoning provided by Syndergaard (1999, Equations (3.23) or (3.36) therein), and after some approximations, we arrive at

$$u_{N_r(z_r),i}^r = \sqrt{(u_{N_r(z_a),i}^r)^2 \cdot (1 + f_{z_r,i})}, \quad (\text{B.18})$$

with

$$f_{z_r,i} = \left(\left. \frac{\partial N_r}{\partial z_r} \right|_i \cdot a_i \cdot 10^{-6} \right)^2. \quad (\text{B.19})$$

$f_{z_r,i}$ is the factor by which the square of $u_{N_r(z_a),i}^r$ (i.e., the uncertainty of the refractivity w.r.t. the impact altitude grid) is increased due to uncertainties in the vertical grid levels $z_{r,i}$. At the lowest altitudes where the refractivity gradient can reach -0.05 m^{-1} and

$a_z \approx R_C$, f_{zr} can be up to 0.1 and consequentially has to be accounted for. Finally $u_{N_r(zr)}^r$ is also interpolated to z (Item 2.7 in Figure 4.2).

It is assumed that correlations R_{N_r} are not affected by the vertical grid changes, because the small grid uncertainties will have little effect on the relatively long-range variation of correlations $R_{N_r,ij}$ with respect to altitude. Because of the proximity of z_i to $z_{r,i}$ it is also assumed that $R_{N_r(z)} \approx R_{N_r(zr)}$, and so R_{N_r} is not interpolated to z , but assumed to be applicable on the target grid.

Using Equation 4.5, the covariance matrix of $N_r(z)$, \mathbf{C}_{N_r} , is finally composed (Item 2.8 in Figure 4.2).

The **systematic uncertainty** is propagated using the state retrieval,

$$u_{N_r}^s = 10^6 (\exp(\mathbf{A}^{\text{AbI}} \cdot u_{\alpha s}^s) - 1). \quad (\text{B.20})$$

To account for uncertainties introduced by assuming *spherical symmetry* in the refractivity retrieval, an uncertainty term $f_{u\text{SphSym}} \cdot N_r$ is added to $u_{N_r}^s$ (Item 2.4 in Figure 4.2). Major contributions to this uncertainty come from horizontal variations of atmospheric variables, as they occur, e.g., due to frontal systems. The propagated and the model-uncertainties together are

$$u_{N_r}^s := (u_{N_r}^s)' = \sqrt{(u_{N_r}^s)^2 + (f_{u\text{SphSym}} \cdot N_r)^2}. \quad (\text{B.21})$$

Based on error estimates of Foelsche and Kirchengast (2004), the relative uncertainty profile $f_{u\text{SphSym}}$ is set to 0.01 % for altitudes down to 7 km, and then linearly increasing to reach 0.05 % at 0 km.

Before the interpolation, z_r is calculated using Equation B.15. Consequentially, systematic uncertainty in the geoid undulation u_{hG}^s needs to be considered here too. The estimated systematic uncertainty of the refractivity on the altitude grid is hence

$$u_{N_r(zr),i}^s = \sqrt{(u_{N_r(zr),i}^s)^2 + \left(\frac{\partial N_r}{\partial z_a} \Big|_i \right)^2 \cdot (u_{hG}^s)^2}. \quad (\text{B.22})$$

Systematic effects from the interpolation itself are assumed to stay below the relative error threshold of 10^{-4} , and will thus not introduce further systematic errors, but the uncertainties $u_{N_r(zr),i}^s$ also need to be interpolated to target grid z .

Due to the integral from $z_{a\text{TopMax}}$ to the level of interest $z_{a,i}$, the levels above $z_{a,i}$ with their higher share of background information cause the **observation-to-background weighting ratio** $r_{N,i}^{\text{obw}}$ to be smaller than $r_{\alpha,i}^{\text{obw}}$. r_N^{obw} can be determined applying

$$r_N^{\text{obw}} = \tilde{\mathbf{A}}^{\alpha 2N} \cdot r_{\alpha}^{\text{obw}}, \quad (\text{B.23})$$

illustrated by Item 2.3 in Figure 4.2, where $\tilde{\mathbf{A}}^{\alpha 2N}$ is the row-normalized kernel of the Abel integral

$$\tilde{A}_{ij}^{\alpha 2N} = \frac{A_{ij}^{\text{Abl}} \cdot \alpha_{s,j}}{\sum_{j=1}^i A_{ij}^{\text{Abl}} \cdot \alpha_{s,j}}. \quad (\text{B.24})$$

Before writing out, r_N^{obw} is also interpolated from the z_r grid to the z grid.

B.1.3 Dry-Air Retrieval

In the dry-air **state retrieval**, the so-called *dry-air variables* are calculated from N_r , using the Smith-Weintraub equation (Kursinski et al. 1997; Smith and Weintraub 1953), the hydrostatic integral and the equation of state. The first step is to calculate the retrieved dry density profile ρ_{dr} from N_r (Item 3.1 in Figure 4.2),

$$\rho_{dr} = \frac{\bar{M}}{c_1 R^*} \cdot N_r, \quad (\text{B.25})$$

with mean molar mass of dry-air $\bar{M} = 28.964 \text{ kg kmol}^{-1}$, the universal molar gas constant $R^* = 8314.45 \text{ J K}^{-1} \text{ kmol}^{-1}$, and the Smith-Weintraub constant $c_1 = 0.7760 \text{ K Pa}^{-1}$.

The next step is to calculate—under the hydrostatic assumption—the retrieved pressure gradient $\eta_{dr} = \rho_{dr} \cdot g$ from ρ_{dr} and the acceleration of gravity g , and integrate the result vertically to obtain the retrieved dry pressure profile p_{dr} . Similar to the ‘baseband approach’ used to solve the Abel integral, also for the hydrostatic integral we subtract a zero-order model profile η_{dm} from the dry pressure gradient η_{dr} (Kirchengast et al. 2017b) and get

$$\delta\eta_{drm} = \eta_{dr} - \eta_{dm} = \rho_{dr} \cdot g - \eta_{dm}. \quad (\text{B.26})$$

The zero-order model part is again solved analytically and the remainder is solved for numerically (Item 3.5 in Figure 4.2). Thus,

$$p_{dr} = p_{dm} + \mathbf{A}^{\text{Hyd}} \cdot \delta\eta_{drm}, \quad (\text{B.27})$$

where the hydrostatic transform matrix is simply an integral discretization,

$$A_{ij}^{\text{Hyd}} = \begin{cases} \Delta z_{j+1} & \text{for } j = 1 \\ \Delta z_j & \text{for } j \leq i \text{ and } i, j > 1 \\ 0 & \text{for } j > i \end{cases}. \quad (\text{B.28})$$

Here Δz_j stands for the difference between two grid-points of the altitude grid $\Delta z_j = z_j - z_{j-1}$.

The state retrieval of the dry temperature applies the equation of state (Item 3.9 in Figure 4.2),

$$T_{dr} = \frac{\overline{M}}{R^*} \frac{p_{dr}}{\rho_{dr}}, \quad (\text{B.29})$$

is the only non-linear operator in the L2a processor and requires a linearization to allow a GUM-type uncertainty propagation.

Calculating the **random uncertainty** of ρ_{dr} (Item 3.2 in Figure 4.2),

$$\mathbf{C}_{\rho_{dr}} = \left(\frac{\overline{M}}{c_1 R^*} \right)^2 \cdot \mathbf{C}_{N_r}, \quad (\text{B.30})$$

and of η_{dr} (3.4 in Figure 4.2),

$$C_{\eta_{dr},ij} = g_i \cdot g_j \cdot C_{\rho_{dr},ij}, \quad (\text{B.31})$$

prepares for the covariances of the retrieved dry pressure profile p_{dr} (Item 3.6 in Figure 4.2). They are calculated using

$$\mathbf{C}_{p_{dr}} = \mathbf{A}^{\text{Hyd}} \cdot \mathbf{C}_{\eta_{dr}} \cdot (\mathbf{A}^{\text{Hyd}})^T. \quad (\text{B.32})$$

According to Syndergaard (1999) and Rieder and Kirchengast (2001b), random uncertainty propagation to dry temperature can be done by expressing p_{dr} and ρ_{dr} as a function of N_r (using Equations B.25, B.26, B.27) and linearizing the resulting equation by a Taylor series expansion about a reasonable model profile T_{dm} . A relation between first-order changes in refractivity δN_{rm} and dry temperature changes δT_{drm} is established:

$$\delta T_{drm} = \mathbf{A}^{\text{N2T}} \cdot \delta N_{rm}. \quad (\text{B.33})$$

Here $\delta N_{rm} = N_r - N_m$, $\delta T_{drm} = T_{dr} - T_{dm}$ and \mathbf{A}^{N2T} has the elements

$$A_{ij}^{\text{N2T}} = \begin{cases} \frac{g_j \Delta z_j \overline{M}}{N_{m,i} R^*} & \text{for } j < i \\ \frac{g_i \Delta z_i \overline{M}}{N_{m,i} R^*} - \frac{T_{dm,i}}{N_{m,i}} & \text{for } j = i \\ 0 & \text{for } j > i \end{cases} \cdot \quad (\text{B.34})$$

The error covariance matrix of the dry temperature is then (Item 3.10 in Figure 4.2):

$$\mathbf{C}_{T_{dr}} = \mathbf{A}^{\text{N2T}} \cdot \mathbf{C}_{N_r} \cdot (\mathbf{A}^{\text{N2T}})^T. \quad (\text{B.35})$$

It is assumed that the use of model profiles N_m and T_{dm} (here taken from ECMWF short-range forecast fields) does not introduce uncertainty to the resulting dry temperature profile, because their exact values do not affect the accuracy of the baseband dry temperature retrieval (but the accuracy of the forward model process connecting T_{dm} and N_m does).

The **systematic uncertainty** is again propagated by the state operators expressed in Equations B.25 to B.29.

B Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles

First, the uncertainty profile of the dry density $u_{\rho_{dr}}^s$ is calculated from $u_{N_r}^s$ by means of Equation B.25. The use of the Smith-Weintraub equation and the equation of state therein increases the estimated uncertainty due to potential inaccuracy of the first Smith-Weintraub constant c_1 and due to the non-ideal gas character of the atmosphere. These two uncertainties are added to the overall estimated systematic uncertainty for each element of ρ_{dr} , using

$$u_{\rho_{dr},i}^s := (u_{\rho_{dr},i}^s)' = \sqrt{(u_{\rho_{dr},i}^s)^2 + \left(-\frac{\overline{M}N_{r,i}}{R^*c_1^2}\right)^2 \cdot (f_{uc1} \cdot c_1)^2 + (f_{uIdlGas,i} \cdot \rho_{dr,i})^2}, \quad (\text{B.36})$$

where the fraction in the second term under the square root is the derivative of Equation B.25 with respect to c_1 . We set the relative uncertainty f_{uc1} to 0.2 % (Rüeger 2002) and $f_{uIdlGas}$ to 0.1 % at $z = 0$ km and exponentially decreasing with altitude by the standard scale height of 7 km, roughly following Kursinski et al. (1997).

Then the systematic uncertainty profile of the dry pressure $u_{p_{dr}}^s$ is derived using Equations B.26 and B.27 (but without subtracting a model profile),

$$u_{p_{dr}}^s = \mathbf{A}^{\text{Hyd}} \cdot (g \cdot u_{\rho_{dr}}^s). \quad (\text{B.37})$$

The integration of the pressure gradient sums up all density contributions of height levels above the one for which the pressure is calculated. In GNSS RO events, the integrated individual density quantities actually stem from locations along the trajectory of the occultation event—which might be rather slant (Foelsche et al. 2011)—but by applying the hydrostatic equilibrium equation it is assumed that these individual pressure gradients all lie vertically above the height level of interest at the bottom of the upward integration (Engeln et al. 2003). This may introduce a systematic error to the resulting pressure profile, to be added to the estimated systematic uncertainty of the dry pressure profile $u_{p_{dr}}^s$ (Item 3.8 in Figure 4.2),

$$u_{p_{dr},i}^s := (u_{p_{dr},i}^s)' = \sqrt{(u_{p_{dr},i}^s)^2 + (f_{uHydEqu,i} p_{dr,i})^2}. \quad (\text{B.38})$$

Following estimates by Engeln et al. (2003, Fig.5 therein), we set the relative uncertainty profile $f_{uHydEqu}$ to 0.01 % at $z = 60$ km, linearly increasing to 0.1 % at $z = 15$ km, and further to 0.2 % at $z = 0$ km.

In contrast to the previous operators, the non-linear character of Equation B.29 renders the propagation of the conservative bound estimates for the systematic uncertainty through the state operator (c.f. Equation 4.4) impossible. Consequentially propagation is done following Equation 4.3, i.e., by propagating the uncertainty added to the *true* state,

$$u_{T_{dr}}^s = \frac{\overline{M}}{R^*} \frac{(p_{dr} + u_{p_{dr}}^s)}{(\rho_{dr} + u_{\rho_{dr}}^s)} - T_{dr}. \quad (\text{B.39})$$

The subtraction of T_{dr} allows to calculate $u_{T_{dr}}^s$.

In the equation of state (i.e. Equation B.29) the non-ideal character of the atmospheric gas is neglected, and a model error must be added to the uncertainty $u_{T_{dr}}^s$ (Item 3.11 in Figure 4.2):

$$u_{T_{dr},i}^s := (u_{T_{dr},i}^s)' = \sqrt{(u_{T_{dr},i}^s)^2 + (f_{u\text{IdlGas},i} T_{dr,i})^2} \quad (\text{B.40})$$

We use the same relative uncertainty profile $f_{u\text{IdlGas}}$ as for ρ_{dr} . We note that later refined estimations of this type of error sources can further improve upon the current simplistic estimates.

From N_r to ρ_{dr} the **observation-to-background weighting ratio** does not change, but analogous to the Abel integral, the pressure integral causes r_{pd}^{obw} to be typically smaller than $r_{\rho d}^{\text{obw}}$ over all altitude levels. r_{pd}^{obw} is calculated using

$$r_{pd}^{\text{obw}} = \tilde{\mathbf{A}}^{\rho 2p} \cdot r_{\rho d}^{\text{obw}}, \quad (\text{B.41})$$

where $\tilde{\mathbf{A}}^{\rho 2p}$ is the row-normalized kernel of the pressure integral,

$$\tilde{A}_{ij}^{\rho 2p} = \frac{A_{ij}^{\text{Hyd}} \cdot \eta_{dr,j}}{\sum_{j=1}^i A_{ij}^{\text{Hyd}} \cdot \eta_{dr,j}} = \frac{A_{ij}^{\text{Hyd}} \cdot \rho_{dr,j} \cdot g_j}{\sum_{j=1}^i A_{ij}^{\text{Hyd}} \cdot \rho_{dr,j} \cdot g_j}. \quad (\text{B.42})$$

To simplify the non-trivial propagation of the observation-to-background weighting ratio to T_{dr} , r_{pd}^{obw} is set as conservative lower bound for $r_{T_d}^{\text{obw}}$. Conservative, because without further integration, the background info is definitely not smeared further down to lower altitudes in $r_{T_d}^{\text{obw}}$ compared to r_{pd}^{obw} .

Integrating uncertainty propagation in GNSS radio occultation retrieval: advanced moist-air algorithm and performance analysis

C.1 Algorithm Description

In this Appendix the L2b uncertainty propagation algorithm implemented in the rOPS is described, following the retrieval chain step by step, starting with L2a output profiles $T_d(z)$ and $p_d(z)$, and background profiles $T_b(z)$ and $q_b(z)$ as input and proceeding to $T_q(z)$, $p_q(z)$, $q_T(z)$ and $p_T(z)$, then to $T_e(z)$ and $q_e(z)$, and finally to $V_{we}(z)$, $p_e(z)$, $e_e(z)$ and $\rho_e(z)$. A description of all variables is provided by Table 5.1. A complete basic description is provided by Li et al. (2018), who give an introduction to the rOPS moist-air retrieval algorithm for the state retrieval and a simplified, variance based, random uncertainty propagation.

If not stated otherwise, elements of the vector-type vertical profiles are addressed using subscript i , and optionally j (with $i, j = 1, \dots, N$), running from top downward towards the bottom of the profile, where N is the total number of vertical grid levels. For simplicity, vertical profiles (i.e., rank-1 vectors) are not denoted by bold symbols in C.1, as their rank can usually be identified from the argument of the variable. All steps of the rOPS moist-air retrieval described in this paper are performed on a 100 m grid, below $z^{\text{DM}} + \Delta z^{\text{DM}} = 17$ km and thus N is at most 170 elements, with the actual length for individual RO events depending on the tropospheric penetration depth of the event.

Above $z^{\text{DM}} + \Delta z^{\text{DM}}$, the moist quantities are essentially approximated by the dry-air variables, and the moist and dry profiles are merged by a half-sine transition weighting over $z^{\text{DM}} \pm \Delta z^{\text{DM}}$ (i.e., over 15 km to 17 km)

C.1.1 Direct Retrievals of Temperature and Specific Humidity

Retrieval of Temperature with Specific Humidity Prescribed

For the **state retrieval** of the directly retrieved temperature profile $T_q(z)$, the dry pressure profile $p_d(z)$, the dry temperature profile $T_d(z)$ and background specific humidity profile $q_b(z)$ serve as input. The first step is to compute the water vapor volume mixing ratio profile $V_{wb}(z)$ from $q_b(z)$,

$$V_{wb,i} = \frac{q_{b,i}}{a_w + b_w q_{b,i}}, \quad (\text{C.1})$$

where $a_w = 0.622$ is the gas constant ratio and $b_w = 1 - a_w$.

Next, to calculate $T_q(z)$, $T_q(z)$ and the retrieved pressure profile $p_q(z)$ are computed iteratively level by level top-downward from z_{topM} to z_{bot} . For this (Item 1.3 in Figure 5.2),

$$T_{q,i} = T_{d,i} \frac{p_{q,i}}{p_{d,i}} \left(1 + \frac{c_T}{T_{q,i}} V_{wb,i} \right) \quad (\text{C.2})$$

and (Item 1.2 in Figure 5.2)

$$p_{q,i} = p_{q,i-1} \left(\frac{p_{d,i}}{p_{d,i-1}} \right)^{\beta_{q,i-1/2}} \quad (\text{C.3})$$

are alternately calculated with the updated $p_{q,i}$ from Equation C.3 put into Equation C.2, and $T_{q,i}$ from Equation C.2 used to calculate the new pressure exponent $\beta_{q,i-1/2}$, until a convergence criteria is met. β is the ratio of the log-pressure change to the log-dry-pressure change between adjacent levels [*log-pressure change coefficient* hereafter]. $\beta_{q,i-1/2}$ is calculated between levels z_i and z_{i-1} ,

$$\beta_{q,i-1/2} = \frac{T_{d,i} + T_{d,i-1}}{T_{q,i} + T_{q,i-1}} \cdot \frac{1 + b_w \sqrt{V_{wb,i} V_{wb,i-1}}}{1 + 2b_w \sqrt{V_{wb,i} V_{wb,i-1}}}, \quad (\text{C.4})$$

Li et al. (2018) show how Equation C.2 can be derived through application of the ideal gas law and the Smith Weintraub refractivity formula.

To propagate the **estimated systematic uncertainty** to $T_q(z)$ and $p_q(z)$, linearized versions of Equations C.2 and C.3 need to be created using a Taylor expansion to first order (see Equation 5.4).

First, the estimated systematic uncertainty profile of the pressure $u_{p_q}^s(z)$ is derived from the systematic uncertainty of dry pressure $u_{p_d}^s$. To get a linear version of Equation C.3, Equation (8) in Li et al. (2018), i.e.,

$$d \ln p_{q,i} = \beta_{q,i} \cdot d \ln p_{d,i}, \quad (\text{C.5})$$

is reformulated using $d \ln x/dx = 1/x$, to get

$$\frac{dp_{q,i}}{p_{q,i}} = \beta_{q,i} \cdot \frac{dp_{d,i}}{p_{d,i}}, \quad (\text{C.6})$$

and hence

$$dp_{q,i} = \beta_{q,i} \cdot \frac{p_{q,i}}{p_{d,i}} dp_{d,i}. \quad (\text{C.7})$$

This linear relation is then used for the uncertainty propagation by adopting the first-order deviations to represent the estimated systematic uncertainties (this simplification may later be needed to be replaced by a more advanced formulation). We thus get the uncertainty profile $u_{pq}^s(z)$,

$$u_{pq,i}^s = \beta_{q,i} \cdot \frac{p_{q,i}}{p_{d,i}} u_{pd,i}^s. \quad (\text{C.8})$$

For the uncertainty propagation it is sufficiently accurate to calculate the log-pressure change coefficient β_q at the altitude levels of interest i , i.e.,

$$\beta_{q,i} = \frac{T_{d,i}(1 + b_w V_{wb,i})}{T_{q,i}(1 + 2b_w V_{wb,i})}. \quad (\text{C.9})$$

Secondly, for calculating estimated systematic uncertainty profile $u_{Tq}^s(z)$ of the retrieved temperature, Equation C.2 is linearized. To simplify the linearization, we use

$$V_{wb,i} \approx q_b/a_{w,i}, \quad (\text{C.10})$$

and

$$c_{q2T} := \frac{c_T}{a_w}, \quad (\text{C.11})$$

and then get

$$dT_{q,i} = \frac{\partial T_{q,i}}{\partial T_{d,i}} dT_{d,i} + \frac{\partial T_{q,i}}{\partial p_{d,i}} dp_{d,i} + \frac{\partial T_{q,i}}{\partial p_{q,i}} dp_{q,i} + \frac{\partial T_{q,i}}{\partial q_{b,i}} dq_{b,i}, \quad (\text{C.12})$$

for the Taylor expansion, when considering only first-order changes of input quantities p_d , T_d , p_q , q_b . The partial derivatives are

$$\frac{\partial T_{q,i}}{\partial T_{d,i}} = \frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}/p_{q,i} + c_{q2T} T_{d,i} q_{b,i}/T_{q,i}^2}, \quad (\text{C.13})$$

$$\frac{\partial T_{q,i}}{\partial p_{d,i}} = -\frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}^2/(T_{d,i} p_{q,i}) + c_{q2T} p_{d,i} q_{b,i}/T_{q,i}^2}, \quad (\text{C.14})$$

$$\frac{\partial T_{q,i}}{\partial p_{q,i}} = \frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}/T_{d,i} + c_{q2T} p_{q,i} q_{b,i}/T_{q,i}^2}, \quad (\text{C.15})$$

$$\frac{\partial T_{q,i}}{\partial q_{b,i}} = \frac{1}{p_{d,i} T_{q,i}/(T_{d,i} p_{q,i} c_{q2T}) + q_{b,i}/T_{q,i}}. \quad (\text{C.16})$$

Since Equation C.2 gives $T_q(z)$ only implicitly, the partial derivatives were calculated by implicit differentiation. Applying the derived systematic uncertainty propagation according to Equation 5.13, and adopting the first-order deviations to represent the systematic uncertainty profiles, we obtain

$$\begin{aligned}
 (u_{T_q,i}^s)^2 = & \left(\frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}/p_{q,i} + c_{q2T} T_{d,i} q_{b,i}/T_{q,i}^2} \cdot u_{T_{d,i}}^s \right)^2 + \\
 & + \left(\frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}^2/(T_{d,i} p_{q,i}) + c_{q2T} p_{d,i} q_{b,i}/T_{q,i}^2} \cdot u_{p_{d,i}}^s \right)^2 + \\
 & + \left(\frac{1 + c_{q2T} q_{b,i}/T_{q,i}}{p_{d,i}/T_{d,i} + c_{q2T} p_{q,i} q_{b,i}/T_{q,i}^2} \cdot u_{p_{q,i}}^s \right)^2 + \\
 & + \left(\frac{1}{p_{d,i} T_{q,i}/(T_{d,i} p_{q,i} c_{q2T}) + q_{b,i}/T_{q,i}} \cdot u_{q_{b,i}}^s \right)^2,
 \end{aligned} \tag{C.17}$$

for the square of the systematic uncertainty profile $u_{T_q}^s(z)$.

The linearized operators are also used to propagate the **estimated random uncertainty**. By using linear Equation C.7, the covariance matrix \mathbf{C}_{p_q} of the retrieved pressure profile $p_q(z)$ can be derived (Item 1.4 in Figure 5.2),

$$C_{p_q,ij} = \frac{\partial p_{q,i}}{\partial p_{d,i}} \frac{\partial p_{q,j}}{\partial p_{d,j}} C_{p_d,ij} = \beta_{q,i} \cdot \frac{p_{q,i}}{p_{d,i}} \beta_{q,j} \cdot \frac{p_{q,j}}{p_{d,j}} C_{p_d,ij}. \tag{C.18}$$

Hereby we made use of the relation expressed in Equation 5.8, i.e., a simplification of the covariance propagation, in this case with $P = 1$.

From \mathbf{C}_{p_q} , the rOPS derives the estimated random uncertainty profile $u_{p_q}^r$ (using Equation 5.9),

$$u_{p_q,i}^r = \sqrt{C_{p_q,ii}}, \tag{C.19}$$

and the correlation matrix \mathbf{R}_{p_q} (using Equation 5.10),

$$R_{p_q,ij} = \frac{C_{p_q,ij}}{u_{p_q,i}^r u_{p_q,j}^r}. \tag{C.20}$$

and also estimates the correlation length profile l_{p_q} (as described in Section 5.3.1), and estimated via

$$l_{p_q,i} = |z_i - z(R_{p_q,ij} = 1/e)|. \tag{C.21}$$

This decomposition of the covariance matrix \mathbf{C}_X into u_X^r and \mathbf{R}_X , and determination of l_X is undertaken after every major propagation step, but is for the sake of brevity not described repeatedly hereafter.

To determine the covariance matrix \mathbf{C}_{T_q} of the retrieved temperature T_q from the covariance matrices of $T_d(z)$, $p_d(z)$, $p_q(z)$, and $q_b(z)$, we use Equation C.12 and get (Item 1.5 in Figure 5.2)

$$C_{T_q,ij} = \frac{\partial T_{q,i}}{\partial T_{d,i}} \frac{\partial T_{q,j}}{\partial T_{d,j}} C_{T_d,ij} + \frac{\partial T_{q,i}}{\partial p_{d,i}} \frac{\partial T_{q,j}}{\partial p_{d,j}} C_{p_d,ij} + \frac{\partial T_{q,i}}{\partial p_{q,i}} \frac{\partial T_{q,j}}{\partial p_{q,j}} C_{p_q,ij} + \frac{\partial T_{q,i}}{\partial q_{b,i}} \frac{\partial T_{q,j}}{\partial q_{b,j}} C_{q_b,ij} \quad (\text{C.22})$$

for the elements of \mathbf{C}_{T_q} , where Equation 5.8 (with $P = 4$) is again valid, because cross-correlations are negligible and the i^{th} altitude level of $T_q(z)$ only depends on the i^{th} altitude level of the four input profiles. These conditions also apply for the following operators of the rOPS L2b processor. The partial derivatives in Equation C.22 are not substituted by Equations C.13 to C.16 because the resulting equation is lengthy and therefore not shown here.

Retrieval of Specific Humidity with Temperature Prescribed

In the **state retrieval**, parallel to the retrieved temperature with specific humidity prescribed $T_q(z)$, the retrieved humidity profile with temperature prescribed $q_T(z)$ and the corresponding retrieved pressure profile $p_T(z)$ are computed from $p_d(z)$, $T_d(z)$ and background temperature profile $T_b(z)$ through iteration. First, the water vapor volume mixing ratio $V_{wT}(z)$ and p_T are retrieved (Items 1.6 and 1.7 in Figure 5.2), with

$$V_{wT,i} = \frac{(p_{d,i} T_{b,i})/p_{T,i} - T_{d,i}}{(c_T T_{d,i})/T_{b,i}} > V_w^{\text{Min}}, \quad (\text{C.23})$$

and

$$p_{T,i} = p_{T,i} \left(\frac{p_{d,i}}{p_{d,i}} \right)^{\beta_{T,i-1/2}}, \quad (\text{C.24})$$

where the log-pressure change coefficient $\beta_{T,i-1/2}$ is again calculated at mid-level

$$\beta_{T,i-1/2} = \frac{T_{d,i} + T_{d,i-1}}{T_{b,i} + T_{b,i-1}} \cdot \frac{1 + b_w \sqrt{V_{wT,i} V_{wT,i-1}}}{1 + 2b_w \sqrt{V_{wT,i} V_{wT,i-1}}}, \quad (\text{C.25})$$

and where V_w^{Min} is a minimum threshold (corresponding to $1.5 \cdot 10^{-3}$ g/kg) that keeps the water vapor estimate positive as a lower bound to stratospheric concentrations. Then the specific humidity profile $q_T(z)$ is calculated from $V_{wT}(z)$,

$$q_{T,i} = \frac{a_w V_{wT,i}}{(1 + b_w V_{wT,i})}. \quad (\text{C.26})$$

For the propagation of the **estimated systematic uncertainty** we again linearize Equation C.24, analogously to Equation C.7 before (an approach that may need further advancement in future as mentioned there), and get

$$dp_{T,i} = \beta_{T,i} \cdot \frac{p_{T,i}}{p_{d,i}} dp_{d,i}, \quad (\text{C.27})$$

where the log-pressure change coefficient $\beta_{T,i}$ is

$$\beta_{T,i} = \frac{T_{d,i}(1 + b_w V_{wT,i})}{T_{b,i}(1 + 2b_w V_{wT,i})}. \quad (\text{C.28})$$

The systematic uncertainty of p_T is then

$$u_{pT,i}^s = \beta_{T,i} \cdot \frac{p_{T,i}}{p_{d,i}} u_{pd,i}^s. \quad (\text{C.29})$$

To get to a linearized version of Equation C.23, $V_{wb,i}$ is substituted by $q_{b,i}$ (using Equation C.10), before applying a Taylor expansion. The linearization to first order yields

$$dq_{T,i} = \frac{\partial q_{T,i}}{\partial T_{d,i}} dT_{d,i} + \frac{\partial q_{T,i}}{\partial p_{d,i}} dp_{d,i} + \frac{\partial q_{T,i}}{\partial p_{T,i}} dp_{T,i} + \frac{\partial q_{T,i}}{\partial T_{b,i}} dT_{b,i}, \quad (\text{C.30})$$

where the partial derivatives are

$$\frac{\partial q_{T,i}}{\partial T_{d,i}} = -c_{T2q} \frac{p_{d,i} T_{b,i}^2}{p_{T,i} T_{d,i}^2}, \quad (\text{C.31})$$

$$\frac{\partial q_{T,i}}{\partial p_{d,i}} = c_{T2q} \frac{T_{b,i}^2}{p_{T,i} T_{d,i}}, \quad (\text{C.32})$$

$$\frac{\partial q_{T,i}}{\partial p_{T,i}} = -c_{T2q} \frac{p_{d,i} T_{b,i}^2}{p_{T,i}^2 T_{d,i}}, \quad (\text{C.33})$$

$$\frac{\partial q_{T,i}}{\partial T_{b,i}} = 2c_{T2q} \frac{p_{d,i} T_{b,i}}{p_{T,i} T_{d,i}} - c_{T2q}. \quad (\text{C.34})$$

We thus obtain

$$\begin{aligned} (u_{qT,i}^s)^2 &= \left(c_{T2q} \frac{p_{d,i} T_{b,i}^2}{p_{T,i} T_{d,i}^2} \cdot u_{Td,i}^s \right)^2 + \\ &+ \left(c_{T2q} \frac{T_{b,i}^2}{p_{T,i} T_{d,i}} \cdot u_{pd,i}^s \right)^2 + \\ &+ \left(c_{T2q} \frac{p_{d,i} T_{b,i}^2}{p_{T,i}^2 T_{d,i}} \cdot u_{pT,i}^s \right)^2 + \\ &+ \left(2c_{T2q} \frac{p_{d,i} T_{b,i}}{p_{T,i} T_{d,i}} - c_{T2q} \right)^2 \cdot (u_{pT,i}^s)^2, \end{aligned} \quad (\text{C.35})$$

for the square of the estimated systematic uncertainty u_{qT}^s .

The **estimated random uncertainty** is determined analogously to C.1.1, i.e., using linearized retrieval Equations C.27 and C.30, while applying Equation 5.8, to simplify the covariance propagation, which then yields (Item 1.8 in Figure 5.2)

$$C_{pT,ij} = \beta_{T,i} \cdot \frac{p_{T,i}}{p_{d,i}} \beta_{T,j} \cdot \frac{p_{T,j}}{p_{d,j}} C_{pd,ij}, \quad (C.36)$$

for the elements of the covariance matrix of the retrieved pressure with temperature prescribed, C_{pT} . The elements of the covariance matrix C_{qT} are (Item 1.9 in Figure 5.2)

$$C_{qT,ij} = \frac{\partial q_{T,i}}{\partial T_{d,i}} \frac{\partial q_{T,j}}{\partial T_{d,j}} C_{Td,ij} + \frac{\partial q_{T,i}}{\partial p_{d,i}} \frac{\partial q_{T,j}}{\partial p_{d,j}} C_{pd,ij} + \quad (C.37)$$

$$+ \frac{\partial q_{T,i}}{\partial p_{T,i}} \frac{\partial q_{T,j}}{\partial p_{T,j}} C_{pT,ij} + \frac{\partial q_{T,i}}{\partial T_{b,i}} \frac{\partial q_{T,j}}{\partial T_{b,j}} C_{Tb,ij}, \quad (C.38)$$

where the partial derivatives need to be substituted by Equations C.31 to C.34. C_{pT} and C_{qT} are decomposed into u_{pT}^r , \mathbf{R}_{pT} and u_{qT}^r , \mathbf{R}_{qT} , respectively (as described for C_{pq} in Equations C.19 to C.20).

C.1.2 Optimal Estimation of Temperature and Specific Humidity

Compared to Li et al. (2018), the availability of the covariance matrices C_{Tq} and C_{qT} also allows for an improvement in the **state retrieval**. As shown by e.g. Healy (2001) or Rieder and Kirchengast (2001a) for the bending angle optimization, using full covariance weighting rather than merely relying on the diagonal elements can improve the weighting in the optimal estimation, particularly when the correlation lengths of retrieved and background state profiles are significantly different.

Here the optimal covariance weighting matrices (also called gain matrices by Rodgers 2000) are

$$\mathbf{A}^{Te} = \mathbf{C}_{Tb}(\mathbf{C}_{Tb} + \mathbf{C}_{Tq})^{-1}, \quad (C.39)$$

and

$$\mathbf{A}^{qe} = \mathbf{C}_{qb}(\mathbf{C}_{qb} + \mathbf{C}_{qT})^{-1}, \quad (C.40)$$

where \mathbf{A}^{Te} is the weighting matrix for the optimal estimation of temperature, calculated from the propagated covariance matrix C_{Tq} and background covariance matrix C_{Tb} , and \mathbf{A}^{qe} is the corresponding one for specific humidity.

Hence, using retrieved temperature profile $T_q(z)$ and background temperature profile T_b , the optimally estimated temperature profile T_e is (Item 2.1 in Figure 5.2)

$$T_e(z) = T_b(z) + \mathbf{A}^{Te} (T_q(z) - T_b(z)), \quad (C.41)$$

or

$$T_{e,i} = T_{b,i} + \sum_{j=1}^N A_{ij}^{Te} (T_{q,j} - T_{b,j}), \quad (\text{C.42})$$

in index notation. Similarly, using retrieved specific humidity profile $q_T(z)$ and background specific humidity profile $q_b(z)$, the optimally estimated specific humidity profile q_e is (Item 2.2 in Figure 5.2)

$$q_e(z) = q_b(z) + \mathbf{A}^{qe} (q_T(z) - q_b(z)), \quad (\text{C.43})$$

or equivalently, in index notation

$$q_{e,i} = q_{b,i} + \sum_{j=1}^N A_{ij}^{qe} (q_{T,i} - q_{b,i}). \quad (\text{C.44})$$

Using the **estimated systematic uncertainty** profile of the retrieved temperature $u_{Tq}^s(z)$ and the background temperature $u_{Tb}^s(z)$, the systematic uncertainty profile $u_{Te}^s(z)$ of the optimally estimated temperature profile $T_e(z)$ can be determined by again assuming perfect correlation along one profile, and no correlation between the systematic uncertainties of the retrieved and the background variable. Then, building on Equation C.42, we can cluster the two uncorrelated terms,

$$T_{e,i} = \sum_{j=1}^N A_{ij}^{Te} T_{q,j} + (T_{b,i} - \sum_{j=1}^N A_{ij}^{Te} T_{b,j}), \quad (\text{C.45})$$

and derive

$$(u_{Te,i}^s)^2 = \left(\sum_{j=1}^N A_{ij}^{Te} u_{Tq,j}^s \right)^2 + \left(u_{Tb,i}^s - \sum_{j=1}^N A_{ij}^{Te} u_{Tb,j}^s \right)^2, \quad (\text{C.46})$$

per element of the square of the estimated systematic uncertainty profile u_{Te}^s . The estimated systematic uncertainty profile $u_{qe}^s(z)$ of the optimally estimated specific humidity profile $q_e(z)$, can be calculated analogously from Equation C.44,

$$(u_{qe,i}^s)^2 = \left(\sum_{j=1}^N A_{ij}^{qe} u_{qT,j}^s \right)^2 + \left(u_{qb,i}^s - \sum_{j=1}^N A_{ij}^{qe} u_{qb,j}^s \right)^2, \quad (\text{C.47})$$

where $u_{qT}^s(z)$ is the systematic uncertainty profile of the retrieved temperature and $u_{qb}^s(z)$ the systematic uncertainty profile of the background temperature.

For the **estimated random uncertainty** through the optimal estimation we follow Rodgers (2000) and calculate the covariance matrix of T_e (Item 2.4 in Figure 5.2) as the weighted retrieved-profile covariance matrix according to

$$\mathbf{C}_{Te} = \mathbf{A}^{Te} \mathbf{C}_{Tq}, \quad (\text{C.48})$$

and, similarly, the covariance matrix of the specific humidity q_e (Item 2.6 in Figure 5.2), according to

$$\mathbf{C}_{q_e} = \mathbf{A}^{q_e} \mathbf{C}_{q_T}. \quad (\text{C.49})$$

Finally, as customary in our framework, we decompose \mathbf{C}_{T_e} and \mathbf{C}_{q_e} into $u_{T_e}^r$ and \mathbf{R}_{T_e} , and $u_{q_e}^r$ and \mathbf{R}_{q_e} , respectively (cf. Equations C.19 and C.20).

C.1.3 Derivation of Further Thermodynamic Variables

Based on the temperature profile $T_e(z)$ and specific humidity profile $q_e(z)$ from the optimal estimation, further thermodynamic variables, such as optimal water vapor volume mixing ratio $V_{we}(z)$, pressure $p_e(z)$, water vapor partial pressure $e_e(z)$ and density $\rho_e(z)$ can be derived. A thorough description of the **state retrieval** for these variables can be found in Li et al. (2018).

Briefly, using the specific humidity profile $q_e(z)$ from the optimal estimation, the optimal estimated water vapor volume mixing ratio profile $V_{we}(z)$ can be calculated (Item 3.2 in Figure 5.2),

$$V_{we,i} = \frac{q_{e,i}}{a_w + b_w q_{e,i}}. \quad (\text{C.50})$$

The optimal pressure $p_e(z)$ in turn can be computed using $p_d(z)$, the optimal temperature profile $T_e(z)$ and the obtained mixing ratio profile $V_{we}(z)$ (Item 3.1 in Figure 5.2)

$$p_{e,i} = p_{e,i-1} \left(\frac{p_{d,i}}{p_{d,i-1}} \right)^{\beta_{e,i-1/2}}, \quad (\text{C.51})$$

where the log-pressure change coefficient β_e is

$$\beta_{e,i-1/2} = \frac{T_{d,i} + T_{d,i-1}}{T_{e,i} + T_{e,i-1}} \cdot \frac{1 + b_w \sqrt{V_{we,i} V_{we,i-1}}}{1 + 2b_w \sqrt{V_{we,i} V_{we,i-1}}}. \quad (\text{C.52})$$

The water vapor partial pressure profile $e_e(z)$ is computed (Item 3.3 in Figure 5.2) using

$$e_{e,i} = V_{we,i} p_{e,i}, \quad (\text{C.53})$$

and the density profile $\rho_e(z)$ using the equation of state in moist air (Item 3.4 in Figure 5.2)

$$\rho_{e,i} = \frac{p_{e,i}}{R_d T_{e,i} (1 + c_w q_{e,i})}, \quad (\text{C.54})$$

where $c_w = 1/a_w - 1 = 0.608$ and $R_d = 287.06 \text{ J kg}^{-1} \text{ K}^{-1}$.

To propagate the **estimated systematic uncertainty** we again create linear versions of the operators in the state retrieval. To calculate the estimated systematic uncertainty of the water vapor volume mixing ratio $u_{V_{we}}^s(z)$, we linearize Equation C.50,

$$dV_{we,i} = \frac{\partial V_{we,i}}{\partial q_{e,i}} \cdot dq_{e,i}, \quad (\text{C.55})$$

and get

$$u_{V_{we,i}}^s = \frac{a_w}{a_w + b_w q_{e,i}} \cdot u_{q_{e,i}}^s, \quad (\text{C.56})$$

through substitution of the first-order changes by the systematic uncertainties. Linearizing Equation C.51 as we did with Equations C.7 and C.27 before (which may later need further advancement as mentioned), we get

$$dp_{e,i} = \frac{\partial p_{e,i}}{\partial p_{d,i}} \cdot dp_{d,i} \quad (\text{C.57})$$

and

$$u_{p_{e,i}}^s = \beta_{e,i} \cdot \frac{p_{e,i}}{p_{d,i}} \cdot u_{p_{d,i}}^s, \quad (\text{C.58})$$

for the estimated systematic uncertainty of pressure $u_{p_e}^s(z)$, where the log-pressure change coefficient β_e is

$$\beta_{e,i} = \frac{T_{d,i}(1 + b_w V_{we,i})}{T_{e,i}(1 + 2b_w V_{we,i})}. \quad (\text{C.59})$$

The systematic uncertainty profile of the partial pressure $u_{e_e}^s$ is found by linearizing Equation C.53,

$$de_{e,i} = \frac{\partial e_{e,i}}{\partial V_{we,i}} \cdot dV_{we,i} + \frac{\partial e_{e,i}}{\partial p_{e,i}} \cdot dp_{e,i}, \quad (\text{C.60})$$

and then applying Equation 5.13 (because the inter-variable systematic uncertainties between p_e and V_{we} are uncorrelated),

$$(u_{e_e,i}^s)^2 = (p_{e,i} \cdot u_{V_{we,i}}^s)^2 + (V_{we,i} \cdot u_{p_{e,i}}^s)^2. \quad (\text{C.61})$$

The systematic uncertainty of the density profile $u_{\rho_e}^s$ is calculated through

$$d\rho_{e,i} = \frac{\partial \rho_{e,i}}{\partial p_{e,i}} \cdot dp_{e,i} + \frac{\partial \rho_{e,i}}{\partial T_{e,i}} \cdot dT_{e,i} + \frac{\partial \rho_{e,i}}{\partial q_{e,i}} \cdot dq_{e,i}, \quad (\text{C.62})$$

and

$$\begin{aligned} (u_{\rho_e,i}^s)^2 = & \left(\frac{1}{R_d T_{e,i} (1 + c_w q_{e,i})} \cdot u_{p_{e,i}}^s \right)^2 + \\ & + \left(\frac{p_{e,i}}{R_d T_{e,i}^2 (1 + c_w q_{e,i})} \cdot u_{T_{e,i}}^s \right)^2 + \\ & + \left(\frac{c_w p_{e,i}}{R_d T_{e,i} (1 + c_w q_{e,i})^2} \cdot u_{q_{e,i}}^s \right)^2, \end{aligned} \quad (\text{C.63})$$

where Equation C.62 is the linearization of Equation C.54, for which the partial derivatives were calculated, and the systematic uncertainties then taken as first order deviations.

The **estimated random uncertainty** profiles of these further thermodynamic variables are again calculated through the simplified covariance propagation expressed by Equation 5.8.

Making use of the linear Equation C.55 created for the propagation of the systematic uncertainties, the random uncertainty of the water mixing ratio $V_{we}(z)$, as derived from the specific humidity $q_e(z)$ is determined through (Item 3.6 in Figure 5.2)

$$C_{V_{we},ij} = \frac{\partial V_{we,i}}{\partial q_{e,i}} \frac{\partial V_{we,j}}{\partial q_{e,j}} C_{q_e,ij} = \frac{a_w}{a_w + b_w q_{e,i}} \frac{a_w}{a_w + b_w q_{e,j}} C_{q_e,ij}. \quad (C.64)$$

To determine the estimated random uncertainty of the pressure profile, expressed by \mathbf{C}_{pe} , we apply the same formulas as for $p_T(z)$ and $p_q(z)$ and get (Item 3.5 in Figure 5.2)

$$C_{pe,ij} = \frac{\partial p_{e,i}}{\partial p_{d,i}} \frac{\partial p_{e,j}}{\partial p_{d,j}} C_{pd,ij} = \beta_{e,i} \cdot \frac{p_{e,i}}{p_{d,i}} \beta_{e,j} \cdot \frac{p_{e,j}}{p_{d,j}} C_{pd,ij}. \quad (C.65)$$

where the log-pressure change coefficient β_e is as defined in Equation C.59.

The the covariance matrix of $e_e(z)$, \mathbf{C}_{ee} , can be calculated from the pressure profile, the vator wapor mixing ratio profile and their uncertainties, based on Equation C.60 (Item 3.7 in Figure 5.2),

$$C_{ee,ij} = \frac{\partial e_{e,i}}{\partial V_{we,i}} \frac{\partial e_{e,j}}{\partial V_{we,j}} C_{V_{we},ij} + \frac{\partial e_{e,i}}{\partial p_{e,i}} \frac{\partial e_{e,j}}{\partial p_{e,j}} C_{pe,ij} = \quad (C.66)$$

$$= p_{e,i} p_{e,j} C_{V_{we},ij} + V_{we,i} V_{we,j} C_{pe,ij}. \quad (C.67)$$

Finally, the estimated random uncertainty of the density profile is determined by (Item 3.8 in Figure 5.2)

$$C_{\rho_e,ij} = \frac{\partial \rho_{e,i}}{\partial p_{e,i}} \frac{\partial \rho_{e,j}}{\partial p_{e,j}} C_{pe,ij} + \frac{\partial \rho_{e,i}}{\partial T_{e,i}} \frac{\partial \rho_{e,j}}{\partial T_{e,j}} C_{Te,ij} + \frac{\partial \rho_{e,i}}{\partial q_{e,i}} \frac{\partial \rho_{e,j}}{\partial q_{e,j}} C_{q_e,ij} \quad (C.68)$$

$$= \frac{1}{R_d T_{e,i} (1 + c_w q_{e,i})} \frac{1}{R_d T_{e,j} (1 + c_w q_{e,j})} C_{pe,ij} + \quad (C.69)$$

$$+ \frac{-p_{e,i}}{R_d T_{e,i}^2 (1 + c_w q_{e,i})} \frac{-p_{e,j}}{R_d T_{e,j}^2 (1 + c_w q_{e,j})} C_{Te,ij} + \quad (C.70)$$

$$+ \frac{-c_w p_{e,i}}{R_d T_{e,i} (1 + c_w q_{e,i})^2} \frac{-c_w p_{e,j}}{R_d T_{e,j} (1 + c_w q_{e,j})^2} C_{q_e,ij} \quad (C.71)$$

As with all other thermodynamic variables, the random uncertainty profile $u_{\rho_e}^r(z)$, the correlation matrix \mathbf{R}_{ρ_e} and correlation length l_{ρ_e} are determined from the covariance matrix.

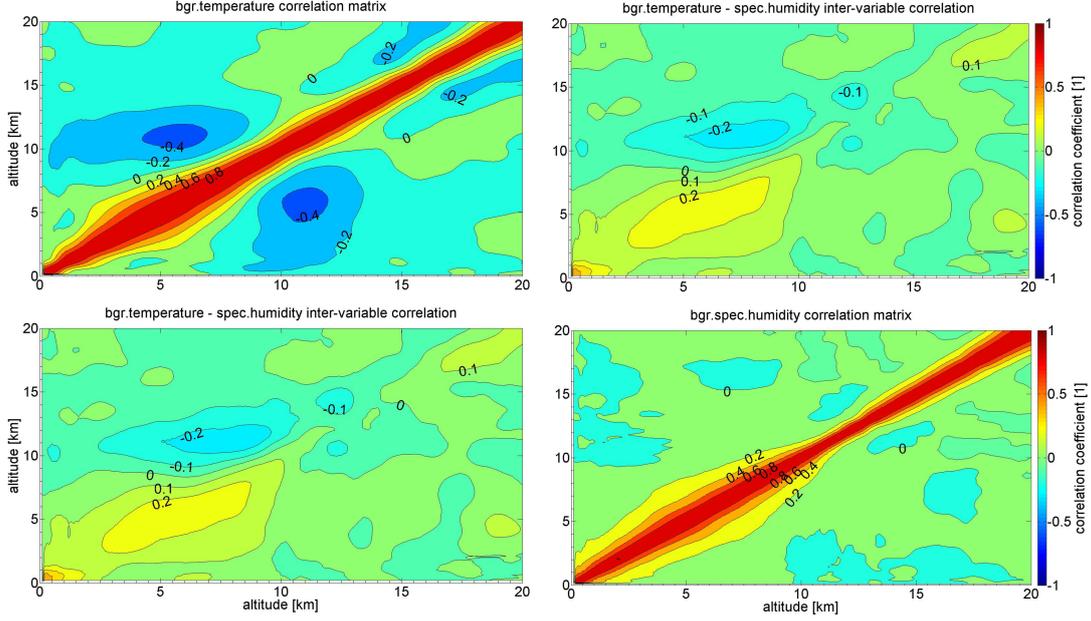


Figure C.1: Intra- and inter-variable correlation coefficients of temperature T_b and specific humidity q_b , for July 15th 2008, for vertical profiles from 0 km to 20 km altitude, in a zonal band from 60°S to 70°S, as constructed according to Li et al. (2013) from ECMWF analysis and forecast fields. Top left: correlation matrix of background temperature, \mathbf{R}_{T_b} . Bottom right: correlation matrix of background specific humidity, \mathbf{R}_{q_b} . Bottom left and top right: inter-variable correlation between background temperature and specific humidity, \mathbf{R}_{T_b, q_b} .

C.2 Proof of Simplification of Equation 5.2

In the rOPS retrieval we mainly have multivariate measurement models (operators) with correlated input quantities of type

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \quad (\text{C.72})$$

where \mathbf{y} , the rank-1 output quantity of length N , is a profile with elements

$$\mathbf{y} = (y_1, \dots, y_i, \dots, y_N)^T, \quad (\text{C.73})$$

with $i = 1, \dots, N$, and \mathbf{x} , the rank-1 input quantity of length K , consists of P profiles \mathbf{x}_μ (with $\mu = 1, \dots, P$), each of length N (and thus $K = P \cdot N$), with elements

$$\mathbf{x} = (x_1, \dots, x_k, \dots, x_K) = (\mathbf{x}_1^T, \dots, \mathbf{x}_\mu^T, \dots, \mathbf{x}_P^T)^T, \quad (\text{C.74})$$

where $k = 1, \dots, K$, and

$$\mathbf{x}_\mu^T = (x_{\mu 1}, \dots, x_{\mu i}, \dots, x_{\mu N}). \quad (\text{C.75})$$

For such a measurement model, the covariance matrix of the output, \mathbf{C}_y , is calculated from the covariance matrix of the input, \mathbf{C}_x , according to

$$\mathbf{C}_y = \mathbf{J} \mathbf{C}_x (\mathbf{J})^T, \quad (\text{C.76})$$

or element wise, i.e., $C_{y,ij} := [\mathbf{C}_y]_{ij}$, with $i, j = 1, \dots, N$,

$$C_{y,ij} = \sum_{k=1}^K \sum_{l=1}^K J_{ik} C_{x,kl} J_{jl}. \quad (\text{C.77})$$

Here $k, l = 1, \dots, K$, and \mathbf{J} is the Jacobi-Matrix, with elements $J_{ik} := [\mathbf{J}]_{ik}$. Both, \mathbf{J} with $1 \times P$ constituent blocks (matrices), and \mathbf{C}_x with $P \times P$ such blocks, are block matrices. Their elements can therefore also be formulated as $J_{\mu ij} := [\mathbf{J}_\mu]_{ij}$, where $\mathbf{J}_\mu := [\mathbf{J}]_\mu$, and $C_{x,\mu ij} := [\mathbf{C}_{x,\mu}]_{ij}$, where $\mathbf{C}_{x,\mu} := [\mathbf{C}_x]_{\mu\nu}$. The block indices are $\mu, \nu = 1, \dots, P$.

If the measurement model is non-linear, \mathbf{C}_y is actually only a first-order approximation of the output covariance matrix, but without limiting the validity of the following proof, we assume the approximation fulfills the rOPS uncertainty accuracy requirements.

We want to show that Equation C.76 can be simplified to

$$C_{y,ij} = \sum_{\mu=1}^P \frac{\partial y_i}{\partial x_{\mu i}} \frac{\partial y_j}{\partial x_{\mu j}} C_{x,\mu ij}, \quad (\text{C.78})$$

if

$$[\mathbf{C}_x]_{\mu\nu} = \mathbf{C}_{x,\mu\nu} \delta_{\mu\nu}, \quad (\text{C.79})$$

and

$$[\mathbf{J}_\mu]_{ij} = J_{\mu ij} \delta_{ij} = \frac{\partial y_i}{\partial x_{\mu j}} \delta_{ij}, \quad (\text{C.80})$$

where $\delta_{\mu\nu}$ and δ_{ij} are Kronecker-deltas. Equation C.79 means there is no inter-variable correlations, while Equation C.80 means that all operators work level-by-level.

Because \mathbf{J} and \mathbf{C}_x are block matrices, Equation C.76 can be formulated blockwise,

$$\mathbf{C}_y = \sum_{\mu=1}^P \sum_{\nu=1}^P [\mathbf{J}]_\mu [\mathbf{C}_x]_{\mu\nu} [(\mathbf{J})^T]_\nu. \quad (\text{C.81})$$

Using Equation C.79, this gives

$$\mathbf{C}_y = \sum_{\mu=1}^P \mathbf{J}_\mu \sum_{\nu=1}^P \mathbf{C}_{x,\mu\nu} \delta_{\mu\nu} (\mathbf{J}_\nu)^T = \sum_{\mu=1}^P \mathbf{J}_\mu \mathbf{C}_{x,\mu\mu} (\mathbf{J}_\mu)^T. \quad (\text{C.82})$$

Using index notation, and applying Equation C.80, we can deduce

$$\begin{aligned}
 C_{y,ij} &= \sum_{\mu=1}^P \sum_{m=1}^N \sum_{n=1}^N [\mathbf{J}_{\mu}]_{im} [\mathbf{C}_{x,\mu}]_{mn} [\mathbf{J}_{\mu}]_{jn} & (C.83) \\
 &= \sum_{\mu=1}^P \sum_{m=1}^N \frac{\partial y_i}{\partial x_{\mu m}} \delta_{im} \sum_{n=1}^N C_{x,\mu mn} \frac{\partial y_j}{\partial x_{\mu n}} \delta_{jn} \\
 &= \sum_{\mu=1}^M \sum_{m=1}^N J_{\mu im} \delta_{im} C_{x,\mu mj} \frac{\partial y_j}{\partial x_{\mu j}} \\
 &= \sum_{\mu=1}^P \frac{\partial y_i}{\partial x_{\mu i}} \frac{\partial y_j}{\partial x_{\mu j}} C_{x,\mu ij},
 \end{aligned}$$

which was to be shown.

List of Figures

1.1	The relation of GCOS to WCRP, IPCC,UNFCCC	6
1.2	ECVs in the domains atmosphere, ocean, land.	7
1.3	Scheme of the occultation geometry	12
1.4	Schematic view of the main rOPS processors: L1a	19
1.5	Schematic view of the main rOPS processors: L1b	20
1.6	Schematic view of the main rOPS processors: L2a	21
1.7	Schematic view of the main rOPS processors: L2b	22
1.8	Results from the BUC test, refractivity retrieval	25
1.9	BUC results from different Abel integral solutions	25
2.1	Illustration of GUM terminology	33
2.2	Illustration of the difference between error and uncertainty	34
3.1	Schematic workflow (L1b)	49
3.2	Detailed workflow (L1b)	51
3.3	Input profiles excess phase (L1b)	59
3.4	Results for filtered excess phase profiles (L1b)	63
3.5	Results for retrieved Doppler shift profiles (L1b)	64
3.6	Results for geometric optics bending angle profiles (L1b)	66
3.7	Results for filtered bending angle profiles (L1b)	68
3.8	Results for atmospheric bending angle profile (L1b)	69
3.9	Results from the validation of covariance propagation (L1b)	72
3.10	Results for real data ensembles, spaghetti (L1b)	76
3.11	Results for real data ensembles, means (L1b)	77
4.1	Schematic workflow (L2a)	85
4.2	Detailed workflow (L2a)	87
4.3	Input profiles bending angle (L2a)	93
4.4	Results for statistically optimized bending angle profile (L2a)	100
4.5	Results for retrieved refractivity profile (L2a)	101
4.6	Results for retrieved dry pressure profile (L2a)	102
4.7	Results for retrieved dry temperature profile (L2a)	103
4.8	Results from the validation of covariance propagation (L2a)	105
4.9	Results for different geographic cases (L2a)	107
4.10	Results for real data ensembles, spaghetti (L2a)	110

List of Figures

4.11	Results for real data ensembles, means (L2a)	111
5.1	Schematic workflow (L2b)	116
5.2	Detailed workflow (L2b)	118
5.3	Input profiles retrieved dry pressure (L2b)	125
5.4	Input profiles retrieved dry temperature (L2b)	126
5.5	Input profiles background temperature (L2b)	126
5.6	Input profiles background specific humidity (L2b)	127
5.7	Results for retrieved temperature (L2b)	128
5.8	Results for retrieved specific humidity (L2b)	128
5.9	Results for optimal estimation temperature (L2b)	129
5.10	Results for optimal estimation specific humidity (L2b)	130
5.11	Results for optimal estimation pressure (L2b)	131
5.12	Results for optimal estimation density (L2b)	132
A.1	Comparison of BWS low-pass filter and boxcar filters	139
C.1	Cross-correlation background temperature and specific humidity (L2b)	176

List of Tables

1.1	Overall objective and processing tracks and processors	18
1.2	The two main rOPS system layers and their subsystems	23
1.3	The four guiding principles of the rOPS processing design	24
1.4	Examples for innovations in rOPS component algorithms	26
2.1	Retrieval operators and corresponding propagation rule	46
3.1	Principal variables for uncertainty propagation (L1b)	53
3.2	Vertical grids and coordinate variables (L1b)	54
4.1	Principal input variables for uncertainty propagation (L2a)	88
4.2	Principal operators and output variables (L2a)	89
4.3	Vertical grids and coordinate variables (L2a)	90
4.4	Simulated test cases August 08 th , 2008 (L2a)	92
5.1	Principal variables for the uncertainty propagation (L2b)	119

Bibliography

- Anderson, G. P., S. A. Clough, F. X. Kneizys, J. H. Chetwynd and E. P. Shettle (1986). *AFGL atmospheric constituent profiles (0 -120 km)*. Environm. Res. Papers, 954 AFGL-TR-86-0110. Optical Physics Division Air Force Geophysics Laboratory Hanscom AFB, Massachusetts, USA.
- Angerer, B., F. Ladstädter, B. Scherllin-Pirscher, M. Schwärz, A. K. Steiner, U. Foelsche and G. Kirchengast (2017). ‘Quality aspects of the Wegener Center multi-satellite GPS radio occultation record OPSv5.6’. In: *Atmos. Meas. Tech.* 10, pp. 4845–4863. DOI: [10.5194/amt-10-4845-2017](https://doi.org/10.5194/amt-10-4845-2017).
- Anthes, R. A. (2011). ‘Exploring Earth’s atmosphere with radio occultation: contributions to weather, climate, and space weather’. In: *Atmos. Meas. Tech.* 4, pp. 1077–1103. DOI: [10.5194/amt-4-1077-2011](https://doi.org/10.5194/amt-4-1077-2011).
- Anthes, R. A. et al. (2008). ‘The COSMIC/FORMOSAT-3 mission: Early results’. In: *Bull. Amer. Meteor. Soc.* 89.3, pp. 313–333. DOI: [10.1175/BAMS-89-3-313](https://doi.org/10.1175/BAMS-89-3-313).
- Aparicio, J. M. and S. Laroche (2011). ‘An evaluation of the expression of the atmospheric refractivity for GPS signals’. In: *J. Geophys. Res.* 116, D11104. DOI: [10.1029/2010JD015214](https://doi.org/10.1029/2010JD015214).
- Arras, K. O. (1998). *An introduction to error propagation: Derivation, meaning and examples of $Cy = Fx Cx Fx$* . No. LSA-REPORT-1998-001. Swiss Federal Institute of Technology Lausanne, Switzerland: Autonomous Systems Lab, Institute of Robotic Systems.
- Bai, W.-H., C.-L. Liu, X.-G. Meng, Y.-Q. Sun, G. Kirchengast, Q.-F. Du, X.-Y. Wang, G.-L. Yang, M. Liao, Z.-D. Yang, D.-Y. Zhao, J.-M. Xia, Y.-R. Cai, L.-J. Liu and D.-W. Wang (2018). ‘Evaluation of atmospheric profiles derived from single- and zero-difference excess phase processing of BeiDou System radio occultation data of the FY-3C GNOS mission’. In: *Atmos. Meas. Tech.* 11, pp. 819–833. DOI: [10.5194/amt-11-819-2018](https://doi.org/10.5194/amt-11-819-2018).
- Bauer, P., E. Andersson and D. Richardson (2015). ‘The quiet revolution of numerical weather prediction’. In: *Nature* 525, pp. 47–55. DOI: [10.1038/nature14956](https://doi.org/10.1038/nature14956).
- Baum, W. A. and A. D. Code (1953). ‘A photometric observation of the occultation of σ ARIETIS by Jupiter’. In: *Astronomical Journal* 58, pp. 108–112. DOI: [10.1086/106829](https://doi.org/10.1086/106829).

Bibliography

- Beyerle, G., T. Schmidt, G. Michalak, S. Heise, J. Wickert and C. Reigber (2005). ‘GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique’. In: *Geophys. Res. Lett.* 32, L13806. DOI: [10.1029/2005GL023109](https://doi.org/10.1029/2005GL023109).
- Bodeker, G. E., S. Bojinski, D. Cimini, R. J. Dirksen, M. Haeffelin, J. W. Hannigan, D. F. Hurst, T. Leblanc, F. Madonna, M. Maturilli, A. C. Mikalsen, R. Philipona, T. Reale, D. J. Seidel, D. G. H. Tan, P. W. Thorne, H. Vömel and J. Wang (2016). ‘Reference Upper-Air Observations for Climate: From Concept to Reality’. In: *Bull. Amer. Meteor. Soc.* 97, pp. 123–135. DOI: [10.1175/BAMS-D-14-00072.1](https://doi.org/10.1175/BAMS-D-14-00072.1).
- Bojinski, S., M. Verstraete, T. C. Peterson, C. Richter, A. Simmons and M. Zemp (2014). ‘The concept of Essential Climate Variables in support of climate research, applications, and policy’. In: *Bull. Amer. Meteor. Soc.* 95 (9), pp. 1431–1443. DOI: [10.1175/BAMS-D-13-00047.1](https://doi.org/10.1175/BAMS-D-13-00047.1).
- Bonnedal, M., J. Christensen, A. Carlström and A. Berg (2010). ‘Metop-GRAS in-orbit instrument performance’. In: *GPS Solut.* 14 (1), pp. 109–120. DOI: [10.1007/s10291-009-0142-3](https://doi.org/10.1007/s10291-009-0142-3).
- Borowitz, M. (2014). *Challenges to developing a global satellite climate monitoring system*. Tech. rep. MD, USA: University of Maryland Libraries, p. 33. DOI: [10.13016/M2SG78](https://doi.org/10.13016/M2SG78). URL: <http://drum.lib.umd.edu/handle/1903/15991>.
- CEOS (2012). *The Response of the Committee on Earth Observation Satellites (CEOS) to the Global Climate Observing System Implementation Plan 2010 (GCOS IP-10)*. Technical report. Frascati, Italy: CEOS, p. 111. URL: http://ceos.org/document_management/Meetings/COP-21/COP-21_2015/CEOS-Response-GCOS-IP-10_24Sep2012.pdf.
- Danzer, J., B. Scherllin-Pirscher and U. Foelsche (2013). ‘Systematic residual ionospheric errors in radio occultation data and a potential way to minimize them’. In: *Atmos. Meas. Tech.* 6, pp. 2169–2179. DOI: [10.5194/amt-6-2169-2013](https://doi.org/10.5194/amt-6-2169-2013).
- Danzer, J., S. B. Healy and I. D. Culverwell (2015). ‘A simulation study with a new residual ionospheric error model for GPS radio occultation climatologies’. In: *Atmos. Meas. Tech.* 8, pp. 3395–3404. DOI: [10.5194/amt-8-3395-2015](https://doi.org/10.5194/amt-8-3395-2015).
- Dow, J. M., R. E. Neilan and C. Rizos (2009). ‘The International GNSS Service in a changing landscape of Global Navigation Satellite Systems’. In: *JG* 83.3-4, pp. 191–198. DOI: [10.1007/s00190-008-0300-3](https://doi.org/10.1007/s00190-008-0300-3).
- Elliot, J. L. (1996). ‘Probing Planetary Atmospheres with Stellar Occultations’. In: *Ann. Rev. Earth Planet. Sci.* 24, pp. 89–124. DOI: [10.1146/annurev.earth.24.1.89](https://doi.org/10.1146/annurev.earth.24.1.89).

- Engeln, A. von, G. Nedoluha, G. Kirchengast and S. Bühler (2003). ‘One-dimensional variational (1-D Var) retrieval of temperature, water vapor, and a reference pressure from radio occultation measurements: A sensitivity analysis’. In: *J. Geophys. Res.* 108.D11. DOI: [10.1029/2002JD002908](https://doi.org/10.1029/2002JD002908).
- Engeln, A. von, J. Teixeira, J. Wickert and S. A. Buehler (2005). ‘Using CHAMP radio occultation data to determine the top altitude of the Planetary Boundary Layer’. In: *Geophys. Res. Lett.* 32, L06815. DOI: [10.1029/2004GL022168](https://doi.org/10.1029/2004GL022168).
- ESA/EUMETSAT (1998). *The GRAS instrument on METOP*. ESA/EUMETSAT Rep. ESA VR/3021/PI, EUMETSAT EPS/MIS/IN/9. Noordwijk, The Netherlands: ESA, 38pp.
- Fjeldbo, G. and V. R. Eshleman (1968). ‘The atmosphere of Mars analyzed by integral inversion of the Mariner IV occultation data’. In: *Planet. Space Sci.* 16.8, pp. 1035–1059. DOI: [10.1016/0032-0633\(68\)90020-2](https://doi.org/10.1016/0032-0633(68)90020-2).
- Fjeldbo, G., A. J. Kliore and V. R. Eshleman (1971). ‘The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments’. In: *The Astronomical Journal* 76.2, pp. 123–140. DOI: [10.1086/111096](https://doi.org/10.1086/111096).
- Foelsche, U. and G. Kirchengast (2004). ‘Sensitivity of GNSS radio occultation data to horizontal variability in the troposphere’. In: *Phys. Chem. Earth* 29.2–3, pp. 225–240. DOI: [10.1016/j.pce.2004.01.007](https://doi.org/10.1016/j.pce.2004.01.007).
- Foelsche, U., M. Borsche, A. K. Steiner, A. Gobiet, B. Pirscher, G. Kirchengast, J. Wickert and T. Schmidt (2008). ‘Observing upper troposphere-lower stratosphere climate with radio occultation data from the CHAMP satellite’. In: *Climate Dyn.* 31, pp. 49–65. DOI: [10.1007/s00382-007-0337-7](https://doi.org/10.1007/s00382-007-0337-7).
- Foelsche, U., B. Pirscher, M. Borsche, G. Kirchengast and J. Wickert (2009). ‘Assessing the climate monitoring utility of radio occultation data: From CHAMP to FORMOSAT-3/COSMIC’. In: *Terr. Atmos. Ocean. Sci.* 20.1, pp. 155–170. DOI: [10.3319/TAO.2008.01.14.01\(F3C\)](https://doi.org/10.3319/TAO.2008.01.14.01(F3C)).
- Foelsche, U., S. Syndergaard, J. Fritzer and G. Kirchengast (2011). ‘Errors in GNSS radio occultation data: relevance of the measurement geometry and obliquity of profiles’. In: *Atmos. Meas. Tech.* 4, pp. 189–199. DOI: [10.5194/amt-4-189-2011](https://doi.org/10.5194/amt-4-189-2011).
- Fritzer, J., G. Kirchengast and M. Pock (2009). *EGOPS 5.5/SUM, End-to-End Generic Occultation Performance Simulation and Processing System, Version 5.5, Software User Manual*. WEGC-IGAM/Uni Graz Technical Report for ESA/ESTEC No. 1/2009, Doc-Id: WEGC-EGOPS-2009-TR01. Version 1. University of Graz, Austria: WEGC and IGAM.
- GCOS (2003). *The second report on the adequacy of the global observing systems for climate in support of the UNFCCC*. WMO-TD/No. 1143. GCOS-92. Geneva: WMO, p. 85.

- GCOS (2007). *GCOS Reference Upper-Air Network (GRUAN): Justification, requirements, siting and instrumentation options*. WMO-TD/No. 1379. GCOS-112. Geneva: WMO, p. 42. URL: <https://www.gruan.org/gruan/editor/documents/gcos/gcos-112.pdf>.
- (2010a). *Guideline for the Generation of Datasets and Products Meeting GCOS Requirements*. WMO-TD/No. 1530. GCOS-143, update to GCOS-128. Geneva: WMO, p. 12. URL: <http://www.wmo.int/pages/prog/gcos/Publications/gcos-143.pdf>.
 - (2010b). *Implementation plan for the global observing system for climate in support of the UNFCCC (2010 update)*. WMO-TD/No. 1523. GCOS-138, update to GCOS-92. Geneva: WMO, p. 186. URL: <http://www.wmo.int/pages/prog/gcos/Publications/gcos-138.pdf>.
 - (2013). *The GCOS Reference Upper-Air Network (GRUAN) Guide*. GCOS-171. Geneva: WMO, p. 122. URL: <http://www.wmo.int/pages/prog/gcos/Publications/gcos-171.pdf>.
 - (2015). *Status of the Global Observing System for Climate – Full Report*. GCOS-195. Geneva: WMO, p. 358. URL: http://www.wmo.int/pages/prog/gcos/Publications/GCOS-195_en.pdf.
- Gobiet, A., G. Kirchengast, G. L. Manney, M. Borsche, C. Retscher and G. Stiller (2007). ‘Retrieval of temperature profiles from CHAMP for climate monitoring: Intercomparison with Envisat MIPAS and GOMOS and different atmospheric analyses’. In: *Atmos. Chem. Phys.* 7, pp. 3519–3536.
- Gorbunov, M. E. (2002a). ‘Canonical transform method for processing radio occultation data in the lower troposphere’. In: *Radio Sci.* 37.5. DOI: [10.1029/2000RS002592](https://doi.org/10.1029/2000RS002592).
- (2002b). ‘Ionospheric correction and statistical optimization of radio occultation data’. In: *Radio Sci.* 37.5. DOI: [10.1029/2000RS002370](https://doi.org/10.1029/2000RS002370).
- Gorbunov, M. E. and G. Kirchengast (2015). ‘Uncertainty propagation through wave optics retrieval of bending angles from GPS radio occultation: Theory and simulation results’. In: *Radio Sci.* 50, RS6001. DOI: [10.1002/2015RS005730](https://doi.org/10.1002/2015RS005730).
- (2018). ‘Wave-optics uncertainty propagation and regression-based bias model in GNSS radio occultation bending angle retrievals’. In: *Atmos. Meas. Tech. Discuss.* 11, pp. 111–125. DOI: [10.5194/amt-11-111-2018](https://doi.org/10.5194/amt-11-111-2018).
- Gorbunov, M. E. and K. B. Lauritsen (2004). ‘Analysis of wave fields by Fourier integral operators and their application for radio occultations’. In: *Radio Sci.* 39, RS4010. DOI: [10.1029/2003RS002971](https://doi.org/10.1029/2003RS002971).
- Gorbunov, M. E., H.-H. Benzon, A. S. Jensen, M. S. Lohmann and A. S. Nielsen (2004). ‘Comparative analysis of radio occultation processing approaches based on Fourier integral operators’. In: *Radio Sci.* 39, RS6004. DOI: [10.1029/2003RS002916](https://doi.org/10.1029/2003RS002916).

- Griffiths, J. and J. R. Ray (2009). ‘On the precision and accuracy of IGS orbits’. In: *JG* 83, pp. 277–287. DOI: [10.1007/s00190-008-0237-6](https://doi.org/10.1007/s00190-008-0237-6).
- Hajj, G. A., E. R. Kursinski, L. J. Romans, W. I. Bertiger and S. S. Leroy (2002). ‘A technical description of atmospheric sounding by GPS occultation’. In: *J. Atmos. Solar-Terr. Phys.* 64.4, pp. 451–469. DOI: [10.1016/S1364-6826\(01\)00114-6](https://doi.org/10.1016/S1364-6826(01)00114-6).
- Hajj, G. A., C. O. Ao, B. A. Iijima, D. Kuang, E. R. Kursinski, A. J. Mannucci, T. K. Meehan, L. J. Romans, M. de la Torre Juarez and T. P. Yunck (2004). ‘CHAMP and SAC-C atmospheric occultation results and intercomparisons’. In: *J. Geophys. Res.* 109, D06109. DOI: [10.1029/2003JD003909](https://doi.org/10.1029/2003JD003909).
- Healy, S. B. (2001). ‘Smoothing radio occultation bending angles above 40 km’. In: *Ann. Geophys.* 19, pp. 459–468.
- Healy, S. B. and I. D. Culverwell (2015). ‘A modification to the standard ionospheric correction method used in GPS radio occultation’. In: *Atmos. Meas. Tech.* 8, pp. 3385–3393. DOI: [10.5194/amt-8-3385-2015](https://doi.org/10.5194/amt-8-3385-2015).
- Ho, S.-P. et al. (2012). ‘Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile inter-comparison of CHAMP climate records 2002 to 2008 from six data centers’. In: *J. Geophys. Res.* 117, D18111. DOI: [10.1029/2012JD017665](https://doi.org/10.1029/2012JD017665).
- Houghton, J., J. Townshend, K. Dawson, P. Mason, J. Zillman and A. Simmons (2012). ‘The GCOS at 20 years: the origin, achievement and future development of the Global Climate Observing System’. In: *Weather* 67.9, pp. 227–235. DOI: [10.1002/wea.1964](https://doi.org/10.1002/wea.1964).
- Innerkofler, J., C. Pock, G. Kirchengast, M. Schwärz, A. Jäggi, Y. Andres, C. Marquardt and J. Schwarz (2016). *SI-traceable radio occultation excess phase processing with integrated uncertainty estimation for climate applications*. Presentation at OPAC-IROWG International Workshop 8-14 September 2016, Seggau/Leibnitz, Austria, available online at <http://wegcwww.uni-graz.at/opacirowg2016> > Scient. Programme > Mon, Sep 12.
- Innerkofler, J., C. Pock, G. Kirchengast, M. Schwärz, C. Marquardt, Y. Andres and J. Schwarz (2017a). ‘GNSS radio occultation excess phase processing with integrated uncertainty estimation for climate applications’. In: *Atmos. Meas. Tech.* in preparation.

- Innerkofler, J., C. Pock, G. Kirchengast, M. Schwärz, A. Jäggi, Y. Andres, C. Marquardt and J. Schwarz (2017b). *GNSS radio occultation zero-differencing excess phase processing with integrated uncertainty estimation for climate applications*. Presentation at COSMIC-IROWG International Workshop 21-27 September 2017, Estes Park, Colorado, USA, available online at <https://cpaess.ucar.edu/sites/default/files/meetings/2017/cosmic/posters/innerkofler-josef-poster.pdf> (last access: 23 April 2018).
- IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, p. 1535.
- JCGM (2008a). *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement*. Rep. JCGM 100:2008. Joint Committee for Guides in Metrology, Office BIPM, Paris.
- (2008b). *Evaluation of Measurement Data – Supplement 1 to the “Guide to the Expression of Uncertainty in Measurement” – Propagation of distributions using a Monte Carlo method*. Rep. JCGM 101:2008. Joint Committee for Guides in Metrology, Office BIPM, Paris.
 - (2011). *Evaluation of Measurement Data – Supplement 2 to the “Guide to the Expression of Uncertainty in Measurement” – Extension to any number of output quantities*. Rep. JCGM 102:2011. Joint Committee for Guides in Metrology, Office BIPM, Paris.
 - (2012). *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM 3rd edition)*. Rep. JCGM 200:2012. Joint Committee for Guides in Metrology, Office BIPM, Paris.
- Karl, T. R., V. E. Derr, D. R. Easterling, C. K. Folland, D. J. Hofmann, S. Levitus, N. Nicholls, D. E. Parker and G. W. Withee (1995). ‘Critical issues for long-term climate monitoring’. In: *Clim. Change* 31 (2-4), pp. 185–221. DOI: [10.1007/BF01095146](https://doi.org/10.1007/BF01095146).
- Kennedy, M. C. and A. O’Hagan (2001). ‘Bayesian calibration of computer models’. In: *J. Roy. Stat. Soc.* 63 (3). Series B, pp. 425–464.
- Kirchengast, G. and S. Schweitzer (2011). ‘Climate benchmark profiling of greenhouse gases and thermodynamic structure and wind from space’. In: *Geophys. Res. Lett.* 38, L13701 (13). DOI: [10.1029/2011GL047617](https://doi.org/10.1029/2011GL047617).

- Kirchengast, G., M. Schwärz, J. Schwarz, B. Scherllin-Pirscher, C. Pock and J. Innerkofler (2015). *Reference Occultation Processing System (rOPS) for cal/val and climate: a new GNSS RO retrieval chain with integrated uncertainty propagation*. Presentation at the IROWG-4 international workshop, 16-22 April 2015, Bureau of Meteorology, Melbourne, Australia, available at: <http://irowg.org/wpcms/wp-content/uploads/2014/05/Kirchengast-IROWG-4.pdf> (last access: 20 November 2017).
- Kirchengast, G., M. Schwärz, J. Schwarz, B. Scherllin-Pirscher, C. Pock, J. Innerkofler, V. Proschek, A. K. Steiner, J. Danzer, F. Ladstädter and U. Foelsche (2016a). *Employing GNSS radio occultation for solving the global climate monitoring problem for the fundamental state of the atmosphere*. Presentation at the EGU 2016 General Assembly, 17–22 April 2016, Vienna, Austria, Geophys. Res. Abstracts, 18, EGU2016-12035-1, available at: <http://meetingorganizer.copernicus.org/EGU2016/EGU2016-12035-1.pdf> (last access: 20 November 2017).
- (2016b). *The reference occultation processing system approach to interpret GNSS radio occultation as SI-traceable planetary system refractometer*. Presentation at OPAC-IROWG International Workshop 8-14 September 2016, Seggau/Leibnitz, Austria, available online at http://wegcwww.uni-graz.at/opacirowg2016/data/public/files/opacirowg2016_Gottfried_Kirchengast_presentation_261.pdf (last access: 20 November 2017).
- Kirchengast, G., Y. Li, B. Scherllin-Pirscher, M. Schwärz, J. Schwarz and J. K. Nielsen (2017a). *A new retrieval algorithm for tropospheric temperature, humidity and pressure profiling based on GNSS radio occultation data*. Presentation at the EGU 2017 General Assembly, 23–28 April 2017, Vienna, Austria, Geophys. Res. Abstracts, 19, EGU2017-16328-2 available online at: <http://meetingorganizer.copernicus.org/EGU2017/EGU2017-16328-2.pdf> (last access: 23 April 2018).
- Kirchengast, G., M. Schwärz, J. Schwarz, J. Ramsauer, J. Fritzer, B. Scherllin-Pirscher, J. Innerkofler, V. Proschek, T. Rieckh and J. Danzer (2017b). *Reference OPS DAD—Reference Occultation Processing System (rOPS) Detailed Algorithm Description*. Tech. Rep. for ESA and FFG No. 1/2017, Doc-Id: WEGC-rOPS-2017-TR01, Issue 1.7. Wegener Center, Univ. of Graz.
- Kuo, Y.-H., T.-K. Wee, S. Sokolovskiy, C. Rocken, W. Schreiner, D. Hunt and R. A. Anthes (2004). ‘Inversion and error estimation of GPS radio occultation data’. In: *Journal of the Royal Statistical Society* 82.1B, pp. 507–531.
- Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield and K. R. Hardy (1997). ‘Observing Earth’s atmosphere with radio occultation measurements using the Global Positioning System’. In: *J. Geophys. Res.* 102.D19, pp. 23429–23465. DOI: [10.1029/97JD01569](https://doi.org/10.1029/97JD01569).

- Ladreiter, H. P. and G. Kirchengast (1996). ‘GPS/GLONASS sensing of the neutral atmosphere: Model independent correction of ionospheric influences’. In: *Radio Sci.* 31, pp. 877–891. DOI: [10.1029/96RS01094](https://doi.org/10.1029/96RS01094).
- Leroy, S. S., J. A. Dykema and J. G. Anderson (2006). ‘Climate benchmarking using GNSS occultation’. In: *Atmosphere and Climate: Studies by Occultation Methods*. Ed. by U. Foelsche, G. Kirchengast and A. K. Steiner. Berlin Heidelberg: Springer, pp. 287–301.
- Li, Y., G. Kirchengast, B. Scherllin-Pirscher, S. Wu, M. Schwaerz, J. Fritzer, S. Zhang, B. A. Carter and K. Zhang (2013). ‘A new dynamic approach for statistical optimization of GNSS radio occultation bending angles for optimal climate monitoring utility’. In: *J. Geophys. Res.: Atmos* 118, pp. 13022–13040. DOI: [10.1002/2013JD020763](https://doi.org/10.1002/2013JD020763).
- Li, Y., G. Kirchengast, B. Scherllin-Pirscher, R. Norman, J. F. Y. B. Yuan, M. Schwaerz and K. Zhang (2015). ‘Dynamic statistical optimization of GNSS radio occultation bending angles: advanced algorithm and performance analysis’. In: *Atmos. Meas. Tech.* 8, pp. 3447–3465. DOI: [10.5194/amt-8-3447-2015](https://doi.org/10.5194/amt-8-3447-2015).
- Li, Y., G. Kirchengast, B. Scherllin-Pirscher, M. Schwärz, J. K. Nielsen, T.-K. Wee, S.-P. Ho and Y.-B. Yuan (2018). ‘A new algorithm for the retrieval of atmospheric profiles from GNSS radio occultation data in moist air and cross-evaluation among processing centers’. In: *J. Geophys. Res.: Atmos.* in submission.
- Liao, M., P. Zhang, G.-L. Yang, Y.-M. Bi, Y. Liu, W.-H. Bai, X.-G. Meng, Q.-F. Du and Y.-Q. Sun (2016). ‘Preliminary validation of the refractivity from the new radio occultation sounder GNOS/FY-3C’. In: *Atmos. Meas. Tech.* 9.2, pp. 781–792. DOI: [10.5194/amt-9-781-2016](https://doi.org/10.5194/amt-9-781-2016).
- Liu, C.-L., G. Kirchengast, K. Zhang, R. Norman, Y. Li, S. C. Zhang, B. Carter, J. Fritzer, M. Schwaerz, S. L. Choy, S. Q. Wu and Z. X. Tan (2013). ‘Characterisation of residual ionospheric errors in bending angles using GNSS RO end-to-end simulations’. In: *Adv. Space Res.* 52.5, pp. 821–836. DOI: [10.1016/j.asr.2013.05.021](https://doi.org/10.1016/j.asr.2013.05.021).
- Liu, C.-L., G. Kirchengast, K. Zhang, R. Norman, Y. Li, S. C. Zhang, J. Fritzer, M. Schwaerz, S. Q. Wu and Z. X. Tan (2015). ‘Quantifying residual ionospheric errors in GNSS radio occultation bending angles based on ensembles of profiles from end-to-end simulations’. In: *Atmos. Meas. Tech.* 8 (7), pp. 2999–3019. DOI: [10.5194/amt-8-2999-2015](https://doi.org/10.5194/amt-8-2999-2015).
- Luntama, J.-P., G. Kirchengast, M. Borsche, U. Foelsche, A. Steiner, S. B. Healy, A. von Engeln, E. O’Clerigh and C. Marquardt (2008). ‘Prospects of the EPS GRAS mission for operational atmospheric applications’. In: *Bull. Amer. Meteor. Soc.* 89, pp. 1863–1875. DOI: [10.1175/2008BAMS2399.1](https://doi.org/10.1175/2008BAMS2399.1).

- Melbourne, W. G., E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, E. R. Kursinski, T. K. Meehan, L. E. Young and T. P. Yunck (1994). ‘The application of spaceborne GPS to atmospheric limb sounding and global change monitoring’. In: *JPL Publication* 94–18, p. 147.
- Montenbruck, O., M. Garcia-Fernandez, Y. Yoon, S. Schön and A. Jäggi (2009). ‘Antenna phase center calibration for precise positioning of LEO satellites’. In: *GPS Solut.* 13.1, pp. 23–34. DOI: [10.1007/s10291-008-0094-z](https://doi.org/10.1007/s10291-008-0094-z).
- NRC (2007). *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond (Decadal Survey)*. The National Academies Press, Washington, D.C. ISBN: 978-0-309-10387-9. DOI: [10.17226/11820](https://doi.org/10.17226/11820).
- Palmer, P. I., J. J. Barnett, J. R. Eyre and S. B. Healy (2000). ‘A nonlinear optimal estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure’. In: *J. Geophys. Res.* 105.D13, pp. 17513–17526. DOI: [10.1029/2000JD900151](https://doi.org/10.1029/2000JD900151).
- Pirscher, B. (2010). *Multi-satellite climatologies of fundamental atmospheric variables from radio occultation and their validation (Ph.D. thesis)*. Sci. Rep. 33-2010. Austria: Wegener Center Verlag Graz. ISBN: 978-3-9502940-3-3.
- Rieder, M. J. and G. Kirchengast (2001a). ‘Error analysis and characterization of atmospheric profiles retrieved from GNSS occultation data’. In: *J. Geophys. Res.* 106.D23, pp. 31755–31770.
- (2001b). ‘Error analysis for mesospheric temperature profiling by absorptive occultation sensors’. In: *Ann. Geophys.* 19, pp. 71–81.
- Rocken, C., R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.-H. Kuo, and X. Zuo (1997). ‘Analysis and validation of GPS/MET data in the neutral atmosphere’. In: *J. Geophys. Res.* 102.D25, pp. 29849–29866. DOI: [10.1029/97JD02400](https://doi.org/10.1029/97JD02400).
- Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Vol. 2. Series on Atmospheric Oceanic and Planetary Physics. Singapore: World Scientific Publishing Company.
- Rogelj, J., M. den Elzen, N. Höhne, T. Fransen, H. Fekete, H. Winkler, R. Schaeffer, F. Sha, K. Riahi and M. Meinshausen (2016). ‘Paris Agreement climate proposals need a boost to keep warming well below 2°C’. In: *Nature* 534, pp. 631–639. DOI: [10.1038/nature18307](https://doi.org/10.1038/nature18307).
- Rüeger, J. M. (2002). *Refractive index formulae for radio waves*. JS28 Integration of Techniques and Corrections to Achieve Accurate Engineering; FIG XXII International Congress. Washington DC, USA.

- Scherllin-Pirscher, B., A. K. Steiner, G. Kirchengast, Y.-H. Kuo and U. Foelsche (2011a). ‘Empirical analysis and modeling of errors of atmospheric profiles from GPS radio occultation’. In: *Atmos. Meas. Tech.* 4, pp. 1875–1890. DOI: [10.5194/amt-4-1875-2011](https://doi.org/10.5194/amt-4-1875-2011).
- Scherllin-Pirscher, B., G. Kirchengast, A. K. Steiner, Y.-H. Kuo and U. Foelsche (2011b). ‘Quantifying uncertainty in climatological fields from GPS radio occultation: an empirical-analytical error model’. In: *Atmos. Meas. Tech.* 4, pp. 2019–2034. DOI: [10.5194/amt-4-2019-2011](https://doi.org/10.5194/amt-4-2019-2011).
- Scherllin-Pirscher, B., A. K. Steiner, G. Kirchengast, M. Schwaerz and S. S. Leroy (2017). ‘The power of vertical geolocation of atmospheric profiles from GNSS radio occultation’. In: *J. Geophys. Res.: Atmos* 122, pp. 1595–1616. DOI: [10.1002/2016JD025902](https://doi.org/10.1002/2016JD025902).
- Schreiner, W. S., C. Rocken, S. Sokolovskiy and D. Hunt (2010). ‘Quality assessment of COSMIC/FORMOSAT-3 GPS radio occultation data derived from single- and double-difference atmospheric excess phase processing’. In: *GPS Solut.* 14.1, pp. 13–22. DOI: [10.1007/s10291-009-0132-5](https://doi.org/10.1007/s10291-009-0132-5).
- Schwarz, J. C., G. Kirchengast and M. Schwaerz (2017). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval: from bending angle to dry-air atmospheric profiles’. In: *Earth Space Sci.* 4.4, pp. 200–228. DOI: [10.1002/2016EA000234](https://doi.org/10.1002/2016EA000234).
- (2018a). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval : from excess phase to atmospheric bending angle profiles’. In: *Atmos. Meas. Tech.* 11, pp. 2601–2631. DOI: [10.5194/amt-11-2601-2018](https://doi.org/10.5194/amt-11-2601-2018).
- Schwarz, J. C., G. Kirchengast, M. Schwaerz, V. Proschek, Y. Li and J. Innerkofler (2018b). ‘Integrating uncertainty propagation in GNSS radio occultation retrieval: Advanced moist-air algorithm and performance analysis’. In: *Rem. Sens. Environ.* manuscript in preparation.
- Smith, E. and S. Weintraub (1953). ‘The constants in the equation for atmospheric refractive index at radio frequencies’. In: *Proc. IRE* 41, pp. 1035–1037.
- Smith, S. W. (1999). *The Scientist and Engineer’s Guide to Digital Signal Processing*. 2nd ed. San Diego, CA, USA: California Technical Publishing.
- Sokolovskiy, S. V., W. S. Schreiner, C. Rocken and D. Hunt (2009). ‘Optimal noise filtering for the ionospheric correction of GPS radio occultation signals’. In: *J. Atmos. Oceanic Technol.* 26, pp. 1398–1403. DOI: [10.1175/2009JTECHA1192.1](https://doi.org/10.1175/2009JTECHA1192.1).
- Steiner, A. K. and G. Kirchengast (2005). ‘Error analysis of GNSS radio occultation data based on ensembles of profiles from end-to-end simulations’. In: *J. Geophys. Res.* 110, D15307. DOI: [10.1029/2004JD005251](https://doi.org/10.1029/2004JD005251).

- Steiner, A. K., G. Kirchengast and H. P. Ladreiter (1999). ‘Inversion, error analysis, and validation of GPS/MET occultation data’. In: *Ann. Geophys.* 17.1, pp. 122–138. DOI: [10.1007/s00585-999-0122-5](https://doi.org/10.1007/s00585-999-0122-5).
- Steiner, A. K., B. C. Lackner, F. Ladstädter, B. Scherllin-Pirscher, U. Foelsche and G. Kirchengast (2011). ‘GPS radio occultation for climate monitoring and change detection’. In: *Radio Sci.* 46, RS0D24. DOI: [10.1029/2010RS004614](https://doi.org/10.1029/2010RS004614).
- Steiner, A. K. et al. (2013). ‘Quantification of structural uncertainty in climate data records from GPS radio occultation’. In: *Atmos. Chem. Phys.* 13, pp. 1469–1484. DOI: [10.5194/acp-13-1469-2013](https://doi.org/10.5194/acp-13-1469-2013).
- Syndergaard, S. (1998). ‘Modeling the impact of the Earth’s oblateness on the retrieval of temperature and pressure profiles from limb sounding’. In: *J. Atmos. Solar-Terr. Phys.* 60.2, pp. 171–180. DOI: [10.1016/S1364-6826\(97\)00056-4](https://doi.org/10.1016/S1364-6826(97)00056-4).
- (1999). *Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique (PhD thesis)*. DMI Sci Rep 99-6. Danish Meteorological Institute, Copenhagen, Denmark, p. 131.
- (2000). ‘On the ionosphere calibration in GPS radio occultation measurements’. In: *Radio Sci.* 35.3, pp. 865–883.
- Syndergaard, S. and G. Kirchengast (2016). ‘An Abel transform for deriving line-of-sight wind profiles from LEO-LEO infrared laser occultation measurements’. In: *J. Geophys. Res.* 121. DOI: [10.1002/2015JD023535](https://doi.org/10.1002/2015JD023535).
- Thayer, G. D. (1974). ‘An improved equation for the radio refractive index of air’. In: *Radio Sci.* 9.10, pp. 803–807. DOI: [10.1029/RS009i010p00803](https://doi.org/10.1029/RS009i010p00803).
- Untch, A., M. Miller, M. Hortal, R. Buizza and P. Janssen (2006). ‘Towards a global meso-scale model: The high-resolution system T799L91 and T399L62 EPS’. In: *ECMWF Newsletter* 108, pp. 6–13.
- Vorob’ev, V. V. and T. G. Krasil’nikova (1994). ‘Estimation of the accuracy of the atmospheric refractive index recovery from Doppler shift measurements at frequencies used in the NAVSTAR system’. In: *Izv. Atmos. Ocean. Phys.* 29, pp. 602–609.
- Ware, R., M. Exner, D. Feng, M. Gorbunov, K. Hardy, B. Herman, Y.-H. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Solheim, X. Zou, R. Anthes, S. Businger and K. Trenberth (1996). ‘GPS sounding of the atmosphere from low Earth orbit: Preliminary results’. In: *Bull. Amer. Meteor. Soc.* 77.1, pp. 19–40. DOI: [10.1175/1520-0477\(1996\)077<0019:GSOTAF>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0019:GSOTAF>2.0.CO;2).
- Wickert, J., C. Reigber, G. Beyerle, R. König, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. Meehan, W. Melbourne and K. Hocke (2001). ‘Atmosphere sounding by GPS radio occultation: First results from CHAMP’. In: *Geophys. Res. Lett.* 28.17, pp. 3263–3266.

Bibliography

- Wickert, J., T. Schmidt, G. Beyerle, G. Michalak, R. König, S. Heise and C. Reigber (2006). 'GPS radio occultation with CHAMP and GRACE: Recent results'. In: *Atmosphere and Climate: Studies by Occultation Methods*. Ed. by U. Foelsche, G. Kirchengast and A. K. Steiner. Berlin, Heidelberg: Springer-Verlag, pp. 3–16.
- Yunck, T. P., G. F. Lindal and C. H. Liu (1988). 'The role of GPS in precise earth observation'. In: *PLANS '88 - IEEE Position Location and Navigation Symposium*, pp. 251–258. DOI: [10.1109/PLANS.1988.195491](https://doi.org/10.1109/PLANS.1988.195491).

Abstract:

GNSS radio occultation (RO) observations have the potential to provide climate data records of benchmark quality, due to the unique properties of the RO technique. RO records are highly accurate, long-term stable, globally available and provide Essential Climate Variables (ECVs) for the thermodynamic state of the atmosphere. These ECVs, pressure, temperature and tropospheric water vapor profiles, can be derived from the raw RO observations through an atmospheric profiles retrieval chain. To realize the climate benchmark potential, these RO retrievals need to be very accurate and the remaining uncertainties quantified and traced throughout the retrieval chain from raw observations to the ECVs. The new Reference Occultation Processing System (rOPS) at the Wegener Center aims to deliver such an accurate chain with integrated uncertainty propagation.

This thesis contributed to the development and implementation of algorithms in the rOPS to propagate uncertainty, first, from RO excess phase profiles to atmospheric bending angle profiles, then further through refractivity to dry-air profiles, and finally, using also background profiles, to moist-air thermodynamic profiles (the ECVs). In this context, propagation of estimated systematic and random uncertainties, vertical error correlations and resolution estimates, and also of observation-to-background weighting ratio profiles was implemented into the rOPS retrieval chain. Results from the covariance propagation were validated using Monte Carlo ensemble methods. The algorithm performance was demonstrated by test-day ensembles of simulated and real RO event data from the satellite missions CHAMP, COSMIC, and MetOp, which show that the new uncertainty estimation chain performs robustly and delivers reliable results. The thesis thus contributes to the benchmark capability of the rOPS for the benefit of climate change monitoring and other applications.

Zum Inhalt:

Aufgrund der Eigenschaften der Methode der GNSS Radio-Okkultation (RO) haben RO Beobachtungen das Potenzial Klimadaten in Referenz-Qualität zur Verfügung zu stellen. RO Messungen zeichnen sich durch hohe Genauigkeit, Langzeitstabilität und globale Abdeckung aus. Wesentliche Klimavariablen (ECVs) bezüglich des thermodynamischen Zustandes der freien Atmosphäre können davon abgeleitet werden. Diese ECVs, nämlich vertikale Druck-, Temperatur- und Feuchteprofile, werden mithilfe eines Prozessierungssystems aus den RO Rohdaten errechnet. Um Referenz-Klimabeobachtungen bereitstellen zu können, muss die Genauigkeit der Prozessierung sehr hoch sein. Verbleibende Unsicherheiten müssen quantifiziert und von den Rohdaten durch die gesamte Prozessierungskette bis zu den ECVs fortgepflanzt werden. Das neue Referenz-Okkultations-Prozessierungssystem (rOPS) am Wegener Center hat zum Ziel, eine solche hoch genaue Prozessierung mit integrierter Unsicherheitsfortpflanzung zu realisieren.

Diese Dissertation hat zur Entwicklung und rOPS Implementierung von Algorithmen beigetragen, um Unsicherheitsprofile zuerst von Exzess-Phasen zu Brechungswinkeln, dann über Refraktivitäten zu sogenannten trockenen Parametern, und letztlich mithilfe externer Zusatzinformationen zu den thermodynamischen ECVs fortzupflanzen. Es wurden die Algorithmen zur Fortpflanzung von Profilen zufälliger und systematischer Unsicherheiten, vertikaler Korrelationen, vertikaler Auflösung, und dem Gewichtungsverhältnis zwischen Messung und Zusatzinformation entwickelt und implementiert.

Die korrekte Berechnung der Resultate der Fortpflanzung wurde mittels Monte-Carlo Validierung unabhängig geprüft. Die Robustheit des Algorithmus wurde durch Prozessierung von realen Messdaten für verschiedene RO Satellitenmissionen demonstriert. Diese Dissertation trägt damit zur Etablierung der Referenzfähigkeit des rOPS bei, zum Nutzen des Klimawandel-Monitoring und anderer Anwendungen.