

Novel Methods for Atmospheric Carbon Dioxide Retrieval from the JAXA GOSAT and NASA OCO-2 Satellites

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

by Peter Somkuti Department of Physics and Astronomy Earth Observation Science Group

July 2018

Main Supervisor:Prof. Hartmut BöschSecond Supervisor:Dr. Alessandro Battaglia

Abstract

Space-based measurements of atmospheric carbon dioxide (CO_2) provide global coverage with repetition times on the order of days. These measurements are used in combination with flux inversion models to track and identify sources and sinks of carbon. The ultimate goal is a better understanding of natural and anthropogenic contributions to the global carbon cycle, from which mitigation strategies and policies can be derived to deal with the effects of climate change.

The algorithms responsible for inferring the atmospheric concentrations of CO_2 from the high-resolution spectroscopic measurements are the so-called retrieval algorithms. This thesis focuses on two main aspects that are important for a successful retrieval strategy, and both have applications beyond CO_2 retreivals.

The first part of this thesis is centred around solar-induced chlorophyll fluorescence (SIF), a naturally occurring radiance signal produced by vegetation as a by-product of photosynthesis. Due to its spectral signature, it is observed by satellite measurements in the O_2 A-band at ~0.76 µm. Based on an established retrieval concept, the SIF retrieval was implemented and its impact on CO_2 retrievals has been evaluated. The SIF retrievals themselves are of great interest to carbon cycle science, and have been used for two case studies: relating SIF to primary production, and tracking the biosphere response to the 2012 North American drought.

In the second part, the focus of the thesis is on fast radiative transfer (RT) methods, which are acceleration techniques to speed up the computationally very expensive line-by-line RT calculations. A novel method based on principal component analysis has been implemented and further advanced. This allowed for the PCA-based method to be used in CO_2 retrievals for measurements from the OCO-2 instrument. Finally, for the first time, a comparison of three popular fast RT schemes has been performed in a consistent way using the same retrieval algorithm.

© Peter Somkuti, July 2018

This thesis is copyright material and no quotation from it may be published without proper acknowledgement.

Declaration

I hereby declare that no part of this thesis has been previously submitted to this or any other University as part of the requirement for a higher degree. The work described herein was conducted by the undersigned except for contributions from colleagues as acknowledged in the text.

Peter Somkuti

Acknowledgements

There are many people who I must thank for helping me along the way. The biggest portion of my gratitude belongs to my supervisor Hartmut Bösch. He not only gave me this chance in the first place, but helped me become a better scientist throughout the last years. He was a constant source of constructive criticism, and also one of motivation - I believe it is not an overstatement to say he was more convinced and motivated about the results of my work, than I ever was. I realise that I was in a highly privileged position with Hartmut being almost always available for so many discussions to drive forward the research. His support throughout this doctorate is one of the main reasons this Ph.D. experience turned out as successful as it did. I also have to acknowledge Robert Parker, whom I have had many fruitful chats with. It is indispensable to have someone like Robert as close to one's office as I had, and I have benefited from his experience many times. I hope to have given some back during those chats we had about the mysteries of the Python programming language! I extend my gratitude to all members of the Earth Observation Science team, past and present, and especially the often unsung heroes residing in the Michael Atiyah building: Colette, Jane and Bev - thank you! I thank especially Will Hewson for helping me become a better programmer.

My research was sponsored by the ESA GHG-CCI project, and I want to thank everyone involved. I do not consider it a given to be able participate in such an international endeavour. On the other side of the pond, I thank Vijay Natraj for cooperating on what essentially became a core part of my research.

A big shoutout goes to my Ph.D. brother Alex and my Ph.D. cousin Mike. You two have made me feel very welcome in Leicester when I arrived four years ago, and your friendship has made this place a little more like home. While I spent a fair amount of the last four years inside Room 112, there was a life outside the office as well sometimes. I have been very lucky to have had incredible people to share my living space with, so thank you Emily Jane for putting up with me, and especially thank you, Becky - your friendship and support in these last months of write-up madness will never be forgotten. Thank you, other members of my adventuring party: Liam (and the Black Spider), Matt, James, Hannah and Stuart. Thank you Tom, Jonathan and Neil, as well as Andrew (when you could be bothered) for the musical intermezzos. And thanks to all the other physicists stuck in that derelict building.

I probably would have been lost in Leicester were it not for my anchors in Vienna, who made me feel I never truly left. Especially Hakan, Adina, Davide and of course the magnificent Ozzy.

Finally, I thank my family for all the support they have given me throughout these many years, I could not have done it without them.

Thank you all, and now off to new adventures!

Tools

All of the presented results have been possible due to a large number of mostly computational tools. This document was typeset using the LATEX document preparation system via the TeXShop front-end, as well as many additional packages available from the Comprehensive TeX Archive Network (CTAN). I have made extensive use of the Python programming language for all data analysis tasks, especially the scipy stack, but also packages like statsmodels and scikit-learn. A conservative estimate puts the lower limit of the total number of code lines that I have written, including the additions to the University of Leicester Full-Physics algorithm (FORTRAN), at 60 000. All plots in this thesis have been made using matplotlib in combination with the basemap toolkit for map projections. My day-to-day work was done on the University of Leicester ALICE(2) high-performance computing facilities, and many retrievals were done on the NERC/NCEO computing platform CEMS/JASMIN. I am grateful to those working on and maintaining both HPC systems.

Contents

Li	st of	Figures	ix
Li	st of	Tables	xxii
Pı	eface	e and Thesis Structure	1
		I Background	5
1	Mo	tivation	6
	1.1	The Global Carbon Cycle	6
	1.2	Space-Based CO ₂ Measurements and Flux Inversions	10
2	Rer	note Sensing of Carbon Dioxide	13
	2.1	Molecular Spectroscopy of O ₂ and CO ₂	13
	2.2	The O ₂ A-band, and the Weak and Strong CO ₂ Bands	15
	2.3	Three-Band Approach to Concentration Measurement	18
3	GO	SAT and OCO-2	22
	3.1	Mission and Spacecraft Description	22
	3.2	The TANSO-FTS and OCO-2 Instruments	23
	3.3	Sampling Strategy	27
4	The	e University of Leicester Full-Physics Algorithm	31
	4.1	Overview and Workflow Description	31
	4.2	Pressure Grid and Model Atmosphere	34
	4.3	Extinction due to Gas Absorption and Rayleigh Scattering	36
	4.4	Aerosols and the Composite Phase Function	37
	4.5	Instrument Model	42
	4.6	Derivatives with Respect to Atmospheric and Surface Parameters	44
5	Inv	erse Method	47
	5.1	Inverse Problem Formulation	47
	5.2	Bayesian Optimal Estimation	49

5.4 Retrieval Characterisation and Errors 5.5 General Retrieval Setup II Remote Sensing of Chlorophyll Fluorescence 6 Solar-Induced Chlorophyll Fluorescence 6.1 Chloroplasts, Chlorophyll, and Photosystems 6.2 Light-Dependent and Light-Independent Reactions 6.3 Non-Photochemical Quenching 6.4 Chlorophyll Fluorescence 6.5 SIF and Primary Production 7 7 SIF Retrieval from GOSAT 7.1 Fraunhofer Line Discrimination Principle 7.2 Physically-based Retrieval 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO2 Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns 8.3 Time Series 8.4		5.3	Iterative Approach to the Non-Linear Problem	5
5.5 General Retrieval Setup II Remote Sensing of Chlorophyll Fluorescence 6 Solar-Induced Chlorophyll Fluorescence 6.1 Chloroplasts, Chlorophyll, and Photosystems 6.2 Light-Dependent and Light-Independent Reactions 6.3 Non-Photochemical Quenching 6.4 Chlorophyll Fluorescence 6.5 SIF and Primary Production 7		5.4	Retrieval Characterisation and Errors	52
II Remote Sensing of Chlorophyll Fluorescence 6 Solar-Induced Chlorophyll Fluorescence 6.1 Chloroplasts, Chlorophyll, and Photosystems. 6.2 Light-Dependent and Light-Independent Reactions. 6.3 Non-Photochemical Quenching. 6.4 Chlorophyll Fluorescence. 6.5 SIF and Primary Production 7 7 Fetrieval from GOSAT 7.1 Fraunhofer Line Discrimination Principle 7.2 Physically-based Retrieval 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results. 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO ₂ Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns. 8.3 Time Series 8.4 Effect on Flux Inversions		5.5	General Retrieval Setup	53
6 Solar-Induced Chlorophyll Fluorescence 6.1 Chloroplasts, Chlorophyll, and Photosystems			II Remote Sensing of Chlorophyll Fluorescence	58
6.1 Chloroplasts, Chlorophyll, and Photosystems 6.2 Light-Dependent and Light-Independent Reactions 6.3 Non-Photochemical Quenching 6.4 Chlorophyll Fluorescence 6.5 SIF and Primary Production 7 SIF Retrieval from GOSAT 7.1 Fraunhofer Line Discrimination Principle 7.2 Physically-based Retrieval 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO2 Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns. 8.3 Time Series 8.4 Effect on Flux Inversions 9.1 SIF-GPP Relationship 9.2 2012 North American Drought	6	Sola	ar-Induced Chlorophyll Fluorescence	6
 6.2 Light-Dependent and Light-Independent Reactions 6.3 Non-Photochemical Quenching		6.1	Chloroplasts, Chlorophyll, and Photosystems	6
6.3 Non-Photochemical Quenching. 6.4 Chlorophyll Fluorescence. 6.5 SIF and Primary Production . 7 SIF Retrieval from GOSAT 7.1 Fraunhofer Line Discrimination Principle . 7.2 Physically-based Retrieval . 7.3 Retrieval Set-Up . 7.4 Collocation of Auxiliary Data . 7.5 Pre-selection and Initial Filtering of the Retrieval Results . 7.6 Scale-Offset Correlation . 7.7 Bias Correction . 7.8 Assessment of Retrievals and Comparison to GOME-2 . 7.9 Uncertainties of the SIF Retrieval . 8 Influence of SIF on XCO ₂ Retrievals . 8.1 Analysing Retrieval Results . 8.2 Spatial and Seasonal Patterns . 8.3 Time Series . 8.4 Effect on Flux Inversions . 9 SIF Case Studies . 9.1 SIF-GPP Relationship . 9.2 2012 North American Drought .		6.2	Light-Dependent and Light-Independent Reactions	6
6.4Chlorophyll Fluorescence.6.5SIF and Primary Production7SIF Retrieval from GOSAT7.1Fraunhofer Line Discrimination Principle7.2Physically-based Retrieval7.3Retrieval Set-Up7.4Collocation of Auxiliary Data7.5Pre-selection and Initial Filtering of the Retrieval Results.7.6Scale-Offset Correlation7.7Bias Correction7.8Assessment of Retrievals and Comparison to GOME-27.9Uncertainties of the SIF Retrieval8Influence of SIF on XCO2 Retrievals8.1Analysing Retrieval Results8.2Spatial and Seasonal Patterns.8.3Time Series8.4Effect on Flux Inversions9SIF Case Studies9.1SIF-GPP Relationship9.22012 North American Drought		6.3	Non-Photochemical Quenching	6
 6.5 SIF and Primary Production		6.4	Chlorophyll Fluorescence	6
7 SIF Retrieval from GOSAT 7.1 Fraunhofer Line Discrimination Principle 7.2 Physically-based Retrieval 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO ₂ Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns 8.3 Time Series 8.4 Effect on Flux Inversions 8 SIF Case Studies 9.1 SIF-GPP Relationship 9.2 2012 North American Drought		6.5	SIF and Primary Production	6
7.1 Fraunhofer Line Discrimination Principle 7.2 Physically-based Retrieval 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO ₂ Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns 8.3 Time Series 8.4 Effect on Flux Inversions 9 SIF Case Studies 9.1 SIF-GPP Relationship 9.2 2012 North American Drought	7	SIF	Retrieval from GOSAT	6
 7.2 Physically-based Retrieval		7.1	Fraunhofer Line Discrimination Principle	6
 7.3 Retrieval Set-Up 7.4 Collocation of Auxiliary Data 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation 7.7 Bias Correction 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval 8 Influence of SIF on XCO₂ Retrievals 8.1 Analysing Retrieval Results 8.2 Spatial and Seasonal Patterns 8.3 Time Series 8.4 Effect on Flux Inversions 9 SIF Case Studies 9.1 SIF-GPP Relationship 9.2 2012 North American Drought 		7.2	Physically-based Retrieval	7
 7.4 Collocation of Auxiliary Data		7.3	Retrieval Set-Up	7
 7.5 Pre-selection and Initial Filtering of the Retrieval Results 7.6 Scale-Offset Correlation		7.4	Collocation of Auxiliary Data	7
 7.6 Scale-Offset Correlation		7.5	Pre-selection and Initial Filtering of the Retrieval Results	7
 7.7 Bias Correction		7.6	Scale-Offset Correlation	7
 7.8 Assessment of Retrievals and Comparison to GOME-2 7.9 Uncertainties of the SIF Retrieval		7.7	Bias Correction	8
 7.9 Uncertainties of the SIF Retrieval		7.8	Assessment of Retrievals and Comparison to GOME-2	8
 8 Influence of SIF on XCO₂ Retrievals 8.1 Analysing Retrieval Results		7.9	Uncertainties of the SIF Retrieval	9
 8.1 Analysing Retrieval Results	8	Infl	uence of SIF on XCO ₂ Retrievals	10
 8.2 Spatial and Seasonal Patterns		8.1	Analysing Retrieval Results	10
 8.3 Time Series		8.2	Spatial and Seasonal Patterns	10
 8.4 Effect on Flux Inversions		8.3	Time Series	10
9SIF Case Studies19.1SIF-GPP Relationship19.22012 North American Drought1		8.4	Effect on Flux Inversions	11
9.1 SIF-GPP Relationship9.2 2012 North American Drought	9	SIF	Case Studies	11
9.2 2012 North American Drought		9.1	SIF-GPP Relationship	11-
		9.2	2012 North American Drought	12
Summary	Su	ımma	ary	13
III Fast Radiative Transfer for XCO ₂ Retrievals 1			III Fast Radiative Transfer for XCO ₂ Retrievals	14(

10 Atmospheric Radiative Transfer	141
10.1 The Radiative Transfer Equation	141

10.2 Solving the Scalar RT Equation	144
10.3 Implementations and Performance Considerations	147
11 Fast Radiative Transfer Methods	149
11.1 General Overview and Motivation	149
11.2 Low-Streams Interpolation	151
11.3 Linear-k	156
11.4 PCA-based Fast RT Method	157
11.5 Detailed Analysis of the PCA-based Method	167
11.6 Performance Aspects of the Fast RT Methods	176
12 CO ₂ Retrievals from OCO-2	181
12.1 Retrieval Setup and Implementation of Fast RT Methods	181
12.2 OCO-2 Retrievals Using the PCA-based Method	184
12.3 Intercomparison of Fast RT Methods	188
12.4 Conclusions	205
Summary	206
Final Remarks and Outlook	208
Appendices	212
A The Discrete Ordinates Method	213
Bibliography	Ι

List of Figures

1.1	The famous Keeling curve, recorded at the Manua Loa Observatory	
	from 1955 onward in weekly intervals (Keeling et al. 2001). The	
	steady rise due to anthropogenic emissions is superposed on the	
	strong seasonal cycle. This cycle arises mainly due to forest leaf-	
	out in the northern hemisphere causing a drawdown in spring.	
	Data taken from "Primary Mauna Loa CO ₂ Record Scripps CO ₂	
	Program" (2018)	7
1.2	Simplified summary of the major fluxes and pools of the global	
	carbon cycle. For a more complete illustration see e.g. IPCC (2014,	
	Figure 6-1). Inspired by the same figure. Values for carbon reservoirs	
	are in Pg C, fluxes (arrows) are in Pg C yr ⁻¹	8
1.3	Partitioned carbon dioxide sources and sinks, taken from Le Quéré	
	et al. (2018).	9
1.4	ObsPack (NOAA Earth System Research Laboratory 2016) surface	
	(red circles) and tower (orange squares) measurement locations;	
	aircraft measurements not shown	10
2.1	Carbon dioxide vibration types. In Figure 2.1c, the cross and	
	dot symbols represent arrows pointing into and out of the plane	
	respectively. The two bend modes are energetically degenerate	14
2.2	Absorption cross sections σ for molecular oxygen in the A-band,	
	as seen by an instrument with GOSAT-like resolution (see Table 3.1).	16
2.3	The effect of pressure broadening demonstrated on a line in the	
	O_2 A-band. As opposed to Figure 2.2, no convolution has taken	
	place and the cross sections are shown on a native grid of 0.01 cm^{-1} .	
	$T = 290 \mathrm{K}$ for both.	17
2.4	Weak CO_2 band along with interfering species CH_4 and H_2O . The	
	cross sections for methane and water vapour were scaled for better	
	visibility.	18

2.5	Strong CO ₂ band at $T = 290$ K and $p = 1120$ hPa. Compare the ordinate tick labels with Figure 2.4, showing the significant	
2.6	difference in line strengths Example of simulated spectra corresponding to the three bands, calculated using the GOSAT instrument model (see Section 3.2).	18
	The radiances are normalised with respect to the largest value in the band.	21
3.1	Simplified TANSO-FTS schematic showing the main components of the instrument, inspired by the layout from JAXA/NIES ("TANSO-	
~ ~	FTS - GOSAT - JAXA" 2018).	25
3.2	Simplified schematic of one out of three (with same optical set-	
	up, different components) spectrometers of the OCO-2 instrument.	0.0
0.0	Inspired by Figure 2-3 from the OCO-2 Data User Guide.	26
3.3	GOSAT sampling pattern over the Amazon for three days in Septem-	
	ber, different years. Inspired by a figure from Webb (2017, Fig-	
	ure 3.14). Starting from a 5-point CT scan (left), GOSAT switched	
	to a 3-point CT pattern (middle) and a special dithering strategy	00
0.4	for cloudy regions (right).	28
3.4	GOSAT sampling for three consecutive days in September 2014	
	showing the global coverage. These locations are not screened for	
	cloud cover; taken from the UoL-FP pre-processing output, only	00
9.5	selected for the main LIB quality flag	29
3.5	OCO-2 global measurement pattern, divided into hadir (left) and	
	glint (right) measurement geometries. These locations have passed	
	a cloud-screening filter as well as other quality criteria; the colour	00
0.6	(see legend) represents the day of measurement in September 2016.	29
3.6	Typical GOSAT and OCO-2 ground footprints. The GOSAT scene	
	(southern Illinois, USA) and the OCO-2 footprints (central Aus-	
	tralia) are shown in plots at almost equal scale. Background im-	
	ages acquired through the Esri ArcGIS REST Services using the	0.0
	World_Imagery layer	30
4.1	Simplified schematic to illustrate the data aggregation and compu-	
	tation tasks required for a full-physics CO_2 retrieval. Data-related	
	items are shown as red rectangles, computational items as blue	
	rectangles with rounded corners and dashed edges. Tasks involving	
	the UoL-FP algorithm are shown as green rectangles with rounded	
	corners. Inspired by the schematic from Detmers et al	33

4.2	Layered model atmosphere which is the basis of the UoL-FP algo-	
	rithm. Every layer is considered to be optically homogeneous and	
	plane-parallel to the surface. Sun and satellite icons are taken from	
	www.openclipart.org	35
4.3	Rayleigh scattering phase function from Equation 4.5	38
4.4	Phase functions of the base types - radial components of the plot	
	are logarithmically scaled. The hydrophilic black carbon type is	
	not displayed, as it is almost overlapping the hydrophobic type	39
4.5	Example of the GOSAT and OCO-2 instrument line shape func-	
	tions. The GOSAT ILS (left) is taken for the reference wavenumber	
	$13200\mathrm{cm}^{-1}$ (O ₂ A-band), whereas the one for OCO-2 (right) is	
	taken from the same band, for footprint 1, and the first detector	
	sample	43
F 1		
5.1	GOSAT and OCO-2 column averaging kernels showing the near-	
	surface sensitivity for both instruments. The retrievals are taken	
	from May 2010, and the three lines correspond to solar zenith angle	59
5.0	bins (in degrees) to which the retrievals were grouped.	53
5.2	CO_2 prior covariance matrix S_a (left) and the corresponding corre-	
	lation matrix \mathbf{C} (right). The axes labels indicate the level (0 is the	
	top of the atmosphere). For the relationship between level index	<i></i>
5.9	and pressure, see Section 4.2.	55
5.3	Example of a prior covariance matrix for the small type aerosol	
	mixture (left). The values for stratospheric levels are very small,	
	and will now allow the retrieval to adjust the aerosol in those levels.	
	Correlations (right) show that the aerosol extinction at a given level	F7
	is coupled to the levels below and above.	57
6.1	Illustration of the inner structure of a chloroplast. The disk-shaped	
	thylakoids are arranged in stacks (grana lamellae), interconnected	
	via the stroma lamellae.	62
6.2	Chlorophylls a $(C_{55}H_{72}MgN_4O_5)$ and b $(C_{55}H_{70}MgN_4O_6)$. The	
	two chloropylls are identical apart from the functional group R,	
	which for chlorophyll a is another methyl (CH_3) , and for chlorophyll	
	b is an aldehyde (CHO).	62
6.3	This illustration depicts a heavily simplified version of the Z scheme.	63

6.4	This figure illustrates a typical broadband SIF emission spectrum. The shaded area between 755 nm and 772 nm marks the location of the O_2 A-band as measured by GOSAT/TANSO-FTS and other instruments. Note that the SIF magnitude at 755 nm is about 1.8 times higher compared to 772 nm. The emission spectrum was calculated using FluorMODgui (Zarco-Tejada et al. 2006)	66
7.1	Illustration of the FLD principle, inspired by Alonso et al. (2008). The solar line, centred at λ_{in} , is filled in via the presence of the	
	additive fluorescence radiance F_{s}	69
7.2	A modelled O_2 A-band spectrum with shaded regions marking the two micro-windows at 755 nm and 772 nm that are utilised in the retrieval scheme. The two red arrows mark the positions of two weak O_2 doublettes in the 772 nm micro-window	71
7.3	Close-up of the two micro-windows used in the physically-based retrieval scheme	71
7.4	Example of the collocation/aggregation of land cover data. This (almost circular) GOSAT footprint, located in Manitoba (Canada), contains about 65% needleleaf evergreen tree cover, about 17% shrubland and about 17% water body LC classes (total of ~1800 pixels). Hatched pixels are not collocated, as the areal overlap is < 100%.	74
7.5	Collocation example for the various target data sets at different res- olutions to illustrate how the spatial extent of the GOSAT footprint was taken into account. Coordinate labels are suppressed here since the location of the footprint is the same as in Figure 7.4. Hatched pixels are not aggregated. For (b), the projection was switched to a cylindrical equal area map projection. [†] Sources: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community	75
7.6	TransCom basis function map ("The Transcom Experiment" 2018), which divides the landmasses (excluding Greenland and Antarctica) into 11 separate regions. Ocean regions have been ignored in this visualisation.	76
7.7	An example of a successful fit and corresponding residual. Radi- ances and residuals are in units of $1 \cdot 10^{-7} \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$	77
7.8	Two box plots showing the distributions of the χ^2 for both micro- windows and the three polarisations separately (only for the year 2010). The S-polarised retrievals are closer to 1, and also exhibit lower spread than the P- or PS-polarised ones	78

7.9	Mean correlation matrices for the $755\mathrm{nm}$ (left) and $772\mathrm{nm}$ (right)	
	micro-windows (S-polarised spectra depicted, but P and PS correla-	
	tions are very similar)	79
7.10	A systematic retrieval bias for two scenarios where only the true	
	albedo differed (SZA: 5°, true fluorescence: $3 \cdot 10^{-8} \mathrm{W}\mathrm{cm}^{-2}\mathrm{sr}^{-1}\mathrm{cm}^{-1}$)	
	with ρ = 0.25 and ρ = 0.35. The markers are filled according to the	
	prior albedo as a percentage of the true albedo	81
7.11	The difference between largest and smallest value of the deviations	
	\hat{z}_0-z_0 (in units of $1\cdot 10^{-11}{ m Wcm^{-2}sr^{-1}cm^{-1}})$ when the prior albedo	
	is perturbed between 90 $\%$ and 110 $\%$ of the true albedo	82
7.12	Gridded (2° × 2°) retrieved z_0 for the year 2012 (P-polarised, 755 nm	
	window)	83
7.13	Radiance-dependent bias for P-polarised spectra, split up into the	
	two micro-windows and the two gain modes. The solid, black line	
	(derived using the LOWESS technique) is used to highlight the	
	shape of the bias	85
7.14	For easier visualisation, the point clouds as seen in Figure 7.13 are	
	replaced by LOWESS-smoothed curves. For both gain modes, the	
	same observation is made: S-polarised correction curves are shifted	
	up (compared to P-polarised ones), as are the curves for the 772 nm	
	window	86
7.15	Same data as Figure 7.12, but using bias corrected values. Re-	
	gions that are considered permanently free of vegetation (Sahara,	
	Antarctica, Greenland, etc.) now exhibit roughly 0 fluorescence	87
7.16	The correction value $z_0^{(ext{corr})}$ as a function of $ar{I}^{(1)}$ and time. The	
	left-hand (right-hand) side plot shows the values for high-gain (low-	
	gain) measurements. As an example, a trapezoidal area for a winter	
	season is highlighted where measurements over Antarctica dominate	
	the derivation of the correction	87
7.17	For the noted bin in $ar{I}^{(1)}$ -space, the values of $z_0^{(\mathrm{corr})}$ (S-polarised,	
	$755\mathrm{nm},$ high-gain) are aggregated and the mean and standard	
	deviation are shown in the above figure	88
7.18	Out of roughly 4 million retrievals, about 3.5 million pass the quality	
	filtering, with the S-polarised measurements surpassing the P- and	
	PS-polarised ones.	89

7.19	Stability of the bias-corrected SIF over Antarctica. Solid lines	
	represent means, while coloured and shaded areas are one standard	
	deviation away from the mean for that time period. The dashed,	
	grey lines are the mean state vector uncertainties for the entire time	00
-	series. Absolute SIF are in units of mW m ² sr ⁻¹ nm ⁻¹	90
7.20	Full time-series averaged on 2° by 2° degree bins, P-polarised spectra.	
	The intensity of the green shading for tropical Africa and the eastern	
	US clearly show the stronger SIF signal for the 755 nm micro-window. (755) (779)	91
7.21	Ratio $F_{\rm s}^{(133)}/F_{\rm s}^{(112)}$ of SIF for the two retrieval windows. Grid cells	
	with values lower than $0.2 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ were masked and ap-	
	pear white in order to highlight areas with a strong SIF signal	92
7.22	The maps show the difference between $1.8 \cdot F_{\rm s}^{(7/2)}$ and $F_{\rm s}^{(755)}$ for	
	S- and P-polarised spectra. The tropical forests in South Amer-	
	ica, Africa, and Indonesia exhibit the largest differences. In these	
	regions, $F_{ m s}^{(755)}$ is considerably larger than $1.8 \cdot F_{ m s}^{(772)}$	93
7.23	Overall comparison of Spline-corrected GOSAT SIF against GOME-	
	2 SIF (Köhler et al. 2015). The 1:1 line is shown in solid, black; the	
	best-fit line is dashed, red	93
7.24	GOSAT/GOME-2 comparisons for various Transcom regions after	
	2° by 2° binning. Only collocated GOME-2 measurements were used	
	where the time difference to the GOSAT measurement was less than	
	5 hours. The standard error of the estimated slopes and intercepts	
	is given as well. Note that the axes are switched when comparing to	
	Figure 7.23. This figure illustrates the region-dependent differences	
	between GOME-2 and GOSAT SIF	95
7.25	$F_{\rm s}$ (blue, circles) and associated uncertainties (orange, squares).	
	Error bars of F_s are close to the mean single-sounding uncertainties.	
	This confirms that the width of the posteriori distribution accurately	
	represents the uncertainty of the retrievals	97
	represents the uncertainty of the retrevals.	07
8.1	The violin plot visualises the change of number of iterations (SIF -	
	noSIF) when compared the set of retrievals with a zero SIF prior	
	(noSIF). When more than four iterations were required for the noSIF	
	case, the SIF case shows less needed iterations in general	103
8.2	XCO ₂ differences versus SIF prior on a sounding-by-sounding basis.	
	The while, dashed line indicates the gradient of $1 \mathrm{ppm} \%^{-1}$ (Franken-	
	berg et al. 2012), and the solid, black line is the actual best-fit to	
	the data	104
8.3	Seasonal maps of $\Delta XCO_2 = XCO_2^{(SIF)} - XCO_2^{(noSIF)}$ showing the	
	regional pattern which correspond to the vegetated areas.	105

8.4	ΔXCO_2 as a function of the SIF prior, like Figure 8.2, but with grid	
	cell averages, based on seasonal aggregation	106
8.5	Seasonal aggregation as in Figure 8.4, but showing the effect on	
	surface pressure instead	107
8.6	XCO_2 for Tropical Asia. The middle panel shows the difference	
	between using a SIF prior, and not using one, ΔXCO_2 , along with	
	the magnitude of the SIF prior. For this specific region, F_s and	
	ΔXCO_2 are anti-correlated ($R = -0.51$). The bottom panel shows	
	the relative change in retrieved AOD for the three mixtures separately.	108
8.7	This figure maps each region into a space spanned by the mean	
	of the ΔXCO_2 time series, and its seasonal amplitude. Error bars	
	represent the standard deviation of the seasonal amplitudes	109
8.8	Differences of $\Delta XCO_2 = \Delta XCO_2^{(corr)} - \Delta XCO_2^{(nocorr)}$. This figure	
	illustrates the net effect of one particular bias correction (see Equa-	
	tion 8.1) on the concentration. The differences tend to be lower	
	than 0.1 ppm, however the spatial pattern still persists	111
0.1		
9.1	Seasonal averages for SIF (755 nm, P-polarised) and EVI. The	
	rightmost column shows the number of measurements in each grid	119
0.0		113
9.2	Same SIF data as in Figure 9.1, but averaged into 2 latitude bins. The later limit is the SIF in the for $W_{1} = \frac{2}{3} = \frac{1}{3} = \frac{1}{3}$ the init	
	The left ordinate is for SIF in units of mW m ² sr ² nm ² , the right	
	ordinate is for EVI and NDVI. EVI was scaled by a factor of 1.6 to	110
0.0	make it align with the NDVI curve.	113
9.3	CASA uptake and MODIS NDVI compared to SIF for a few TransCom	
	regions. Iemperate and boreal regions show that all time series	
	are very much in sync in terms of the seasonal cycle. For tropical	
	regions, like South Asia and the Amazon, the NDVI time series	
	shows the same seasonal cycle as SIF, but the CASA uptake exhibits	
0.4	a significant phase shift.	115
9.4	Illustration of the global distribution of scale factors α (left), and	
~ ~	as latitudinal aggregates $\alpha(\phi)$ (right)	116
9.5	SIF-GPP relationship for various TransCom regions, aggregated by	
	season. Each circle represents a monthly mean, and the size of the	
	circle is proportional to the standard deviation for that month. The	
	value eta is inferred through ordinary least-squares	118

9.6	SIF from GOME-2 and GOSAT, compared to collocated FluxNet GPP measurements as described in the text. Circles represent monthly average, and their areas scale with the GOSAT SIF standard deviation of the measurements in that month. Dashed, red lines are	
	the best-fit linear models, whereas the dashed, blue lines represent	190
9.7	Similar as Figure 9.6, but with <i>scaled</i> GOSAT SIF. The black line is	120
011	the linear function using the mean slope from Figure 9.5 (excluding	
	DIF)	121
9.8	Each point corresponds in the left panel to a linear fit for a set	
	of collocation using the collocation radius indicated by the colour.	
	The right panel shows histograms of the fitted slopes β	122
9.9	Illustrating the sampling bias, the fitted values for β' are shown like	
	in Figure 9.8 - the true value of $\beta' = 1$ is reached only for larger	
	collocation radii $\geq 500 \text{km}$	123
9.10	Global mean annual uptake between 2010 and 2016 and the cor-	
	responding fractional cover of the globe in terms of the grid cells.	
	Full coverage (100%) is obtained if all grid cells with a land fraction	
	>10% contain valid SIF measurements	125
9.11	Annual global fluxes and global land coverage using a grid cell	
	size of 3.5° and four different aggregation intervals. The significant	
	coverage drop in the year 2014 for aggregation intervals smaller	
	than a year is caused by a gap in the GOSAT data stream between	
	December 2014 and January 2015. The dashed lines in the left	
	panel show variability in the result due to $\sigma_{\beta} = 0.229$	126
9.12	Precipitation and soil moisture anomalies for 2012 against a 2010 to	
	2016 baseline (gridded to $2^{\circ} \times 2^{\circ}$ cells). TRMM precipitation data	
	does not extend further than 50° N. The bounding box shows the	
	region of interest onto which the data is sub-set for the following	
	time-series analysis.	128
9.13	GOSAT sampling density for 2009-2016. The western half of the	
	bounding box is sampled through the regular pattern with ${\sim}150{ m km}$	
	distance between point triplets, as well a track of more continuous	
	measurements. The eastern region is less densely sampled	128

9.14	TRMM precipitation time-series. The measurements are re-sampled	
	to 15 day intervals, and for every interval the mean (thick, blue line)	
	and the standard deviation (shaded area) are computed. For times	
	between May and August of each year, the values are highlighted.	
	The horizontal, red-dashed line marks the mean (standard deviation	
	again shown shaded) of all values between May and August for all	
	years, thus representing the climatology for the region	129
9.15	This figure is designed in the same way as Figure 9.14. Summer	
	2012 exhibits anomalously low soil moisture for all three time series,	
	but is best seen for the active measurements (middle panel)	129
9.16	The standardised precipitation-evapotranspiration index shows a	
	significant departure from the average during the 2012 summer	130
9.17	Ensemble medians (solid line) and ensemble maxima and minima	
	(shaded) show the relative consistency between polarisations and	
	calibration methods. The unphysical spikes near data gaps are a	
	result of data sparsity	131
9.18	Summer SIF differences $\Delta F_{\rm s}$ compared to the 2012 level. All four	
	data sets (S-polarised) show similar trends with small deviations.	
	For all sets, 2009 and 2012 summers are low compared to post-2012	
	levels. For relative SIF at $755\mathrm{nm}$, the summer average for 2009 was	
	even lower than for 2012	131
9.19	Monthly mean gross uptake from CASA (left) along with the SIF-	
	derived uptake values (right). Both CASA and SIF show the same	
	warm spring in 2012	133
9.20	The location of five flux tower sites, which have a comparatively long	
	record of measurements. Three sites (US-PFa, US-Los and US-WCr,	
	slightly outside the region of interest outlined in Figure 9.12) are	
	clustered very closely together, the remaining three sites, while still	
	within the region of interest, are several hundred kilometres to the	
	east and south	134
9.21	FluxNet GPP from the six selected locations compared to scaled	
	SIF for the entire region of interest. Both panels show monthly means.	135
9.22	Monthly aggregated time series of the four stations US-WCr, US-	
	UMB, US-Oho, and US-MMS (blue) and the region-averaged scaled	
	SIF. The dashed, red line is the mean of the annual maxima of the	
	full FluxNet time series from 1999 onward (not shown)	135
9.23	Monthly mean aggregates of the four FluxNet sites from Figure 9.22,	
	but overlaid to show the early spring onset in 2012	136
9.24	Relative fluxes compared to the 2010 baseline (see Table 9.2)	136

9.25	Net fluxes as given by the CASA prior (left) and the inversion performed by the University of Edinburgh, where prior fluxes for 2012 were set to be those of 2011 - thus showing the early onset is not merely a result of the prior.	137
11.1	Intensity differences $\Delta I = I_{low} - I_{high}$ as functions of wavelength and total gas optical depth. The right panels are a histogram rather than point plots, and the point density is indicated by the hue (dark green = low, light yellow = high). Residuals for the right column are in the same units as on the left	152
11.2	The same intensity differences as Figure 11.1c, but colours repre- senting $\sqrt{\xi}$. For larger values of $\sqrt{\xi}$, the spectral points lie mostly on the upper branch	153
11.3	Example of reconstructed radiance in the strong CO_2 band, nor- malised to the maximal value of I_{Full} to better highlight the residual slope	154
11.4	LSI-reconstructed radiances with the applied slope correction (strong CO_2 band, same as in Figure 11.3). The residuals appear spectrally flat	155
11.5	Representation of $\overline{\mathbf{F}}$ (excluding the surface albedo) for the first bin in the O ₂ A-band. Each colour represents a different percentile based on the set of all spectral points within the bin	160
11.6	Same as Figure 11.5, but for the last bin $(5 < \tau \le \infty)$	163
11.7	The first five scaled EOFs W for the first and last bins. For bin 22, the components corresponding to the Rayleigh optical depth profiles, aerosol scattering coefficient and fraction have been multiplied by 20 for visualisation purposes. ρ was omitted	163
11.8	PCs in the vicinity of three weak oxygen lines in the O_2 A-band, from the first bin	164
11.9	PCs at the position of a strong oxygen line, taken from bin number 21	164
11.10	Reconstruction of τ around a strong oxygen line using $N_{\text{EOF}} = 1 \dots 5$ PCs. For $N_{\text{EOF}} = 1$, the error (as a percentage of the original τ) was scaled by 0.1, and for $N_{\text{EOF}} = 2$ by 0.5 so that all five curves could be fit into the same graph.	165
	courd be ne meto the same graph.	100

11.11	Locations for which forward model simulations are performed, and	
	corresponding total aerosol optical depths. For both seasons, the	
	coverage is almost global, where the shift towards high southern	
	latitudes is due to the changing SZA for the winter season. Note	
	the two different latitudinal bands covered by the glint observations	
	over the oceans in the two seasons. Apart from boreal forests and	
	Greenland, most landmasses are covered in both seasons	169
11.12	Histograms of the spectral residuals. The abscissas represent the	
	interquartile ranges of the absolute values of the relative errors	
	$ I_{\text{rel},i} $ (see Equation 11.43). The inset text boxes in each sub-plot	
	show the median and the interquartile ratios for each distribution	
	(number of EOFs).	170
11.13	Global map of residual magnitudes when using 1 EOF. Bins for	
	ocean simulations have been chosen to be large enough to allow for	
	continuous patches despite the spacing of orbits (see Figure 11.11),	
	which makes a visual assessment easier. The factors shown in the	
	top-right corner of each map indicate the factors, with which the	
	respective ensembles have been multiplied (L for land, G for glint)	
	such that the values can be plotted using the same colormap. Re-	
	gions over land with larger residuals are easily identified, being dusty	
	deserts (Sahara, Arabian peninsula), regions with large amounts	
	of sea salt (high southern latitudes) and aerosol outflow regions	
	(mid-Atlantic, western North-Pacific) as well as high latitudes with	
	higher solar zenith angles	171
11.14	Estimated XCO_2 errors displayed as both maps and histograms. For	
	the case of using only 1 EOF, the errors on the map were scaled by	
	a factor of 0.1 such that the same colorbar could be used. Regions	
	with higher surface reflectance and thus higher SNR clearly result	
	in higher XCO_2 errors (deserts) over land. Over ocean, errors	
	are increased in the mid-Atlantic due to larger AODs (compare	
	Figure 11.11) and the larger viewing angles at the edges of the glint	
	observation bands.	172
11.15	XCO ₂ errors at various cirrus cloud optical depths for a small subset	
	of the winter season scenes. There is a dramatic reduction of the	
	error magnitude going from 1 EOF to 2 EOFs, especially for cases	
	with cirrus optical depth larger than 0.1. The tick labels for the 1	
	EOF case are seen on the left hand side of the three plots, whereas	
	the tick labels for the cases of 2 and 3 EOFs are seen on the right	
	hand side. Whiskers in this plot have been suppressed.	174
	have been approximated and have been approximated ap	- / 1

11.16	XCO_2 errors (using 3 EOFs) of scenes grouped into aerosol optical depth bins (the dashed, grey lines indicate the bin boundaries) for the four instruments listed in Table 11.3 (left to right, GOSAT: grey,	
	A: blue, B: red, C: green) for different surfaces.	177
11.17	Speedup \mathcal{S} as a function of η and the number of bins N_{bin} . For the	
	PCA-based method, the number of EOFs was fixed at 3 for every bin.	180
12.1	Location of the 13 TCCON sites that were used for the validation	184
12.2	Footprint biases for each measurement mode separately	185
12.3	Comparison of land scenes (nadir & glint) against TCCON mea-	
	surements, which were collocated by Brendan Fisher (JPL). The	
	left panel shows the point cloud coloured by the mode of the foot-	
	prints within the bin. An apparent footprint bias would be seen	
	as differently coloured and distinct regions in the correlation plot.	
	Footprint number eight is slightly overrepresented in the sample,	
	hence why they are also represented as the dominant colour in this	
	plot	186
12.4	Same as Figure 12.4, but for ocean scenes	187
12.5	Distribution of footprints in the subset after minimal filtering for	
	convergence and fit quality	187
12.6	This figure demonstrates a systematic surface pressure bias for	
	land retrievals, which explains almost 23% of the variance of the	
	UoL-FP/TCCON differences. The bias is roughly -0.35 ppm hPa ⁻¹ .	
	Inspired by C.W. O'Dell's diagnostic plots	188
12.7	Residuals of the O_2 A-band for a challenging scenario with high	
	aerosol loadings and large solar zenith angle. The top three panels	
	show the relative point-by-point residuals $\Delta I_{\rm rel}$ (as percentage), and	
	Stokes fractions Δq and Δu . They are based on the full SS+MS	
	Stokes vectors and are thus instrument-independent. The last panel	
	shows the total intensity after \mathbf{S} is multiplied with the Müller matrix	
	components (m_{11}, m_{12}, m_{13}) and subsequently convolved with a	
	Gaussian instrument line shape function.	190
12.8	Same as Figure 12.7, but for the weak CO_2 band. Surprisingly,	
	the linear- k method exhibits the smallest residuals, however this	
	can be explained through the dynamic binning strategy that the	
	implementation of the other two methods lack	191
12.9	Same as Figure 12.7, but for the strong CO_2 band	192

12.10	Distribution of selected OCO-2 soundings from May 2016, gridded	
	to a rough 2° by 2° grid. The set covers most landmasses and	
	oceans apart from Antarctica and the southern oceans, with some	
	noticeable gaps near the Indian ocean, Patagonia, tropical Africa	
	and patches in Eastern Asia and Central Northern America.	19
12.11	Retrieval runtimes per iteration in minutes, showing that linear- k is,	
	on average, 18% shorter than LSI, which is about 25% shorter than	
	the PCA-based method using one EOF. Using five EOFs increases	
	the runtime by roughly 65% compared to using only one EOF	19
12.12	Same as Figure 12.11, but for land soundings. The overall picture	
	is the same, however the distributions feature more prominent tails.	19
12.13	Distributions of numbers of iterations (left) and numbers of diver-	
	gent steps (right) for ocean soundings. The retrievals were set up	
	to abort after more than 3 divergent steps	19
12.14	Same as Figure 12.13 but for land soundings	19
12.15	Box-plots illustrating the χ^2 values for the three bands and the four	
	different runs (ocean retrievals)	19
12.16	Same as Figure 12.15 but for land retrievals	19
12.17	Fit quality (χ^2) versus signal-to-noise ratio for the PCA(5) run	20
12.18	Latitudinally binned XCO_2 displayed as differences to the median of	
	all four values of that bin (4° spacing). The shaded areas represent	
	the standard error of the latitudinal bin	20
12.19	Matrix of pairwise differences between the different runs, separated	
	into land (upper right) and ocean (lower left) retrievals. Each map	
	in the inner rectangle (blue colour-scale) represents the difference	
	between (row) and (column), and the inset text box shows the range	
	of XCO ₂ difference for each map	20
12.20	Differences between the respective retrieval runs to the median (like	
	in Figure 12.19) as functions of the prior aerosol optical depths (for	
	small and large type aerosol mixtures) and the geometric airmass	
	factor. The shaded areas represent the standard error of the ΔXCO_2	
	values in the respective AOD or AMF bin	20
12.21	Similar to Figure 12.20 (and same colour coding), but with retrieved	
	AODs rather than the retrieved XCO ₂ . For the last row (cirrus	
	AODs), the retrievals were binned according to their latitude, as	
	the prior is the same for every scene.	20
	• ·	

List of Tables

3.1	Spectral ranges and approximate resolution for TANSO-FTS and	
	OCO-2. Values for TANSO-FTS are taken from "TANSO-FTS -	
	GOSAT - JAXA" (2018) (resolution averaged for P and S), values	
	for OCO-2 are from Frankenberg et al. (2015)	27
7.1	Initial filter thresholds used to select retrievals for the further bias	
	correction procedure. Mean radiances are in units of $ m Wcm^{-2}sr^{-1}cm^{-1}.$	78
7.2	Time series means and root-mean-squares (RMS) of SIF for the	
	three calibration modes as seen in Figure 7.19. Values are in units	
	of 10^{-8} W cm ⁻² sr ⁻¹ cm ⁻¹ . (S-polarised measurements only)	90
7.3	Same as Table 7.2, but for relative fluorescence (values in %). No-	
	tably, the Spline-based correction procedure performs worse than	
	the other two methods in terms of the mean, the RMS is however	
	lower	91
7.4	R^2 values between retrieved GOSAT SIF and GOME-2 SIF, com-	
	pared on a cell-by-cell basis. Retrievals in the 755 nm micro-window	
	exhibit a higher correlation, as do measurements using P-polarised	
	spectra	94
7.5	Slopes of fitted linear functions between GOME-2 SIF and GOSAT	
	SIF (see Figure 7.24, intercepts are not shown). The values in	
	brackets indicate the R^2 value for the subset, and bold-faced slopes	
	are those for which the R^2 is larger than 0.5	96
7.6	For the subsets, in which N random retrievals (absolute SIF) over	
	the Antarctic are drawn, the D -values / p -values of the Kolmogorov-	
	Smirnov test are listed. The D -values are smaller than the p -values	
	for all N , all windows, and all calibration types, meaning that the	
	test fails to reject the hypothesis of the SIF values being drawn from	
	a normal distribution	98

7.7	Estimation of regional uncertainties. In the first group, σ is the regional standard deviation including all measurements. For the <i>k</i> -means and land cover (LC) methods, <i>N</i> is the number of clusters with more than 50 soundings. For means and uncertainties, values are in units of mW m ⁻² sr ⁻¹ nm ⁻¹ . The last row is the mean (± standard deviation) of all σ main to be the standard deviation.	100
7.8	Same as Table 7.7, but for July 2014. For this month, there are no soundings over Antarctica.	100
8.1	A table summarising the fit quality indicators for both sets of re- trievals. Using a physical SIF prior appears to lead to overall slightly worse fits	104
8.2	Peak-to-peak seasonal amplitudes for both XCO ₂ and Δ XCO ₂ per region. The third column is the ratio of the mean Δ XCO ₂ /XCO ₂ , expressed as a percentage, the last column is the correlation coefficient between $F_{\rm s}$ and Δ XCO ₂	109
9.1 9.2	Coefficients α_i for the polynomial representation of α (see Figure 9.4). Mean annual SIF-derived and CASA-sampled fluxes as absolute and relative differences to 2010 levels. Absolute fluxes are in units of g C m ⁻² yr ⁻¹ . The third column pair (CASA [†]) are the annual flux differences to 2010 using non-collocated model data, where the entire region of interest was averaged, not taking into account the GOSAT sampling.	117
11.1	Percentage of scenes with XCO_2 errors larger than ± 0.1 ppm	173
11.2	Land cover class indices according to the ESA-CCI Land Cover map (epoch 2008-2012) associated with the four surface types	175
11.3	Instrument model characteristics, the FWHMs of the Gaussian instrument line shape functions are stated for all three bands (O_2 A-band, weak CO_2 and strong CO_2). Sampling ratio is the number of spectral points per FWHM.	175
11.4	Instrument model coefficients. Radiances (L, L_{ref}) and coefficient b are in ph s ⁻¹ m ⁻² sr ⁻¹ μ m ⁻¹ , a in $(ph s^{-1} m^{-2} sr^{-1} \mu m^{-1})^2$, and C_{back} , C_{ph} and SNR _{ref} are unitless	176
12.1	Bias correction coefficients.	188

12.2	Statistics (bias Δ and scatter σ) of the UoL-FP/TCCON matching	
	and comparison, after the bias correction procedure. Station-wise	
	biases are calculated as the mean and standard deviation of the	
	per-station means. Information on the sites themselves can be found	
	at the official TCCON website ("TCCON - Tccon-Wiki" 2018)	189
12.3	The first three lines in the table header define the scenario. Aerosol	
	optical depths for small and large-types, cirrus cloud optical depth,	
	and solar zenith angle. For each of the eight scenarios, the interquar-	
	tile range of the absolute value of the residuals are listed. $\Delta I_{ m rel}$ and	
	$\Delta I_{ m rel}$ (conv.) are in units of 0.01 %, Δq and Δu are in units of $1\cdot 10^{-5}$.	
	XCO ₂ errors are in ppm	193
12.4	Same as Table 12.3, but for the weak CO ₂ band	194
12.5	Same as Table 12.3, but for the strong CO ₂ band	194
12.6	Summary of the XCO_2 errors per band from Tables 12.3 to 12.5,	
	including the total error as the sum of all contributions	195
12.7	Runtimes per iteration (in minutes) as well as total runtime in	
	hours for roughly 31000 (ocean) and 85000 (land) retrievals. A	
	full line-by-line computation takes about 45 min per iteration	197

Preface and Thesis Structure

This thesis is the result of almost four years' worth of efforts and work as part of the Earth Observation Science group at the University of Leicester. All of the presented research revolves around the central topic of improving on the space-based carbon dioxide retrievals. I was lucky to be tightly integrated in the greenhouse gases-team and the ESA Climate Change Initiative project, which provided funding for my research. It can be divided into two major parts: solar-induced chlorophyll fluorescence and fast radiative transfer.

The research surrounding space-based measurements of fluorescence gained momentum after Frankenberg, Fisher, et al. (2011) showed that physically reasonable values could be retrieved on a global scale if an appropriate bias correction is applied. As my very first task in June 2014, I implemented their retrieval scheme using the University of Leicester Full-Physics (UoL-FP) algorithm and processed the then-available GOSAT time series. During the preparation of the climate data research package (CRDP3) XCO_2 data, I have contributed to the logistics of the retrieval set-up and the lastminute move to a different computing infrastructure, which required an impromptu re-write of major aspects of the processing pipeline. Along with the introduction of the MACC-II/CAMS-based aerosol scheme (Leif Vogel), fluorescence was added to the retrieval as a contribution to the radiances in the O_2 A-band. This again required modifications to the algorithm code to introduce fluorescence as a state vector-element. I have derived a new calibration method that builds on the established scheme, and also released a fluorescence data package for use by scientists in the UK and international Earth Observation community. As part of the ECMWF project CAMS41, I have set up a near-real time processing stream that acquires GOSAT data from the ESA third-party archive, performs the necessary cloud clearing, fluorescence retrieval and bias correction - every day. The finished daily data package is then provided to ECMWF in a CF-compliant format. While the spatial sparsity of GOSAT measurements made it difficult for the project partners to utilise the data to its fullest extent, I contributed with my data for a poster at the AGU 2017 Fall Meeting (Bacour et al. 2017).

The second big research topic was fast radiative transfer. Originally born out of

the need to increase the computational performance of the UoL-FP algorithm, the task began with reviewing the then implemented radiative transfer models. I have completely removed all traces of an outdated and rarely used model, as well as updated both LIDORT and 2OS radiative transfer models to their newest versions. Additionally, I have added the TWOSTR radiative transfer to the algorithm. Switching out and adding a radiative transfer model was a major effort that took me the best part of the 2015 summer, including all validation exercises to ensure that the forward model output is consistent with older versions of the algorithm.

Using the new models, I started investigating state-of-the-art fast radiative transfer models by implementing them. To facilitate a quicker development time for such a complex problem, I have chosen to modify the UoL-FP code to allow for external calls to arbitrary radiative transfer drivers. I achieved this by writing a module which saves all required atmospheric and surface properties for a given band into an HDF file, which can then be used by any external code. Once the external code finished its calculations, the results are fed back into the algorithm and the computation continues. Using this scheme I have written a fresh implementation of the LSI algorithm, of which an older and outdated version was part of the UoL-FP code. Having LSI as a reference, I began work on the PCA-based algorithm and thanks to my supervisor, had the chance to work for two weeks closely with its inventor Vijay Natraj at the California Institute of Technology. This visit allowed for some major progress. Around six months later, the PCA-based method was fully implemented and now accounted for the spectral variation of aerosol properties, which proved to be a major problem beforehand. I started an extensive study on the impact of this scheme on GOSAT retrievals, which culminated in a publication in the Journal of Geophysical Research (Somkuti, Boesch, Natraj, et al. 2017a). Very early on, I was interested in how different fast RT methods would compare when they are used consistently in one algorithm. By January 2018, I had finally implemented the three major methods natively into the UoL-FP code without the use of the external interface.

The thesis is structured in three parts. The first one (Page 5) discusses the overall motivation for space-based carbon dioxide measurements and introduces the key concepts needed to perform them. It includes a detailed description of the algorithm as well as the utilised satellites and instruments. The second part (Page 58) deals with fluorescence, both as an interesting phenomenon that can be related to vegetation status, but also as an influence in XCO_2 retrievals. The third and last major research part (Page 140) contains a detailed description of atmospheric radiative transfer and its application to trace gas retrievals. I introduce three established acceleration methods, and study their effects on the final result of XCO_2 retrievals.

My work regarding the PCA-based fast RT scheme was published in the Journal of Geophysical Research (see below), and contents therein are found throughout Part III (Section 11.4).

This thesis contains 124 figures and 24 tables, the total word count of the main body of the document obtained through texcount is 53138.

Peer-reviewed

P. Kopparla, V. Natraj, D. Limpasuvan, R. Spurr, D. Crisp, R.-L. Shia, P. **Somkuti**, and Y. L. Yung. 2017. "PCA-Based Radiative Transfer: Improvements to Aerosol Scheme, Vertical Layering and Spectral Binning." *Journal of Quantitative Spectroscopy and Radiative Transfer* 198 (September 1): 104–111

P. Somkuti, H. Boesch, V. Natraj, and P. Kopparla. 2017a. "Application of a PCA-Based Fast Radiative Transfer Model to XCO₂ Retrievals in the Shortwave Infrared." *Journal of Geophysical Research: Atmospheres* 122, no. 19 (October 12): 10, 477–10, 496

M. Buchwitz, O. Schneising, M. Reuter, J. Heymann, S. Krautwurst, H. Bovensmann, J. P. Burrows, H. Boesch, R. J. Parker, P. **Somkuti**, R. G. Detmers, O. P. Hasekamp, I. Aben, A. Butz, C. Frankenberg, and A. J. Turner. 2017. "Satellite-Derived Methane Hotspot Emission Estimates Using a Fast Data-Driven Method." *Atmos. Chem. Phys.* 17, no. 9 (May 9): 5751–5774

M. Buchwitz, M. Reuter, O. Schneising, S. Noël, B. Gier, H. Bovensmann, J. P. Burrows, H. Boesch, J. Anand, R. J. Parker, P. **Somkuti**, R. G. Detmers, O. P. Hasekamp, I. Aben, A. Butz, A. Kuze, H. Suto, Y. Yoshida, D. Crisp, and C. O'Dell. 2018. "Computation and Analysis of Atmospheric Carbon Dioxide Annual Mean Growth Rates from Satellite Observations during 2003-2016." Under review as of 22/03/2018, *Atmos. Chem. Phys. Discuss.* 2018 (March 14): 1–22

T. Trent, H. Boesch, and P. **Somkuti**. 2018. "Observing Water Vapour in the Planetary Boundary Layer From the Short-Wave Infrared." To be submitted, *Remote Sensing*

Talks

P. **Somkuti**. 2015. "Remote Sensing of Chlorophyll Fluorescence." CLCR, University of Leicester, UK, February 4

P. **Somkuti**, H. Boesch, V. Natraj, and P. Kopparla. 2016. "Fast Radiative Transfer Methods and Their Application on Greenhouse Gas Retrievals." Remote sensing in the O_2 A-band. KNMI, De Bilt, Netherlands, July 6

P. Somkuti. 2017. "CAMS41 Science Meeting." Paris, France, March 21

Posters

P. **Somkuti**, H. Boesch, and R. Parker. 2015. "Remote Sensing of Chlorophyll Fluorescence with GOSAT," 17:9582. EGU General Assembly 2015. Vienna, Austria, April 1

P. **Somkuti** and H. Boesch. 2016. "OCO-2 Retrievals Using the UoL-FP Algorithm." OCO-2 Science Team Meeting Fall 2016. Boulder, USA, October 25

P. **Somkuti**. 2016. "Remote Sensing of Chlorophyll Fluorescence." NCEO Science Conference 2016 - Poster prize. University of Warwick, UK, June 29

P. **Somkuti**, H. Boesch, V. Natraj, and P. Kopparla. 2017b. "XCO2 Retrieval Errors from a PCA-Based Approach to Fast Radiative Transfer," 19:6702. EGU General Assembly 2017. Vienna, Austria, April 1

P. **Somkuti**, H. Boesch, J. Anand, R. Parker, V. Natraj, and P. Kopparla. 2017. "Implementing a PCA-Based Fast Radiative Transfer Method for XCO₂ Retrievals." International Workshop on Greenhouse Gas Measurements from Space 2017. Helsinki, Finland, June 6

T. Trent, H. Boesch, P. **Somkuti**, and J. Remedios. 2017. "Observing Near Surface Layer Water Vapour From The SWIR." In *NCEO & CEOI Earth Observation Conference*. Bath, UK, June

C. Bacour, F. Maignan, N. MacBean, P. Köhler, M. Vountas, N. Khosravi, L. Guanter, J. Joiner, C. Frankenberg, P. **Somkuti**, and P. Peylin. 2017. "Comparison of Several Satellite-Derived Sun-Induced Fluorescence Products." AGU Fall Meeting 2017. New Orleans, USA, December 11

Part I



BACKGROUND

Chapter One

Motivation

1.1 The Global Carbon Cycle

ARTH's atmosphere, excluding water vapour (H₂O, ~1%), is composed of mostly nitrogen (N₂, 78.09%) and oxygen (O₂, 20.95%), as well as a small amount of argon (Ar, 0.93%). The remaining fraction, about 0.04%, is collectively known as *trace gases*. With more than 90%, the major component of the trace gases is carbon dioxide (CO₂).

Despite its deceivingly low abundance in the atmosphere, carbon dioxide plays a major role in Earth's climate and biosphere. A well-known consequence of atmospheric carbon dioxide is the *greenhouse effect*. While being somewhat transparent to the incoming short-wave solar radiation, CO_2 in its gas phase absorbs strongly in the infrared wavelength range at around 15 µm. Thus, the thermal long-wave radiation leaving Earth will be partly absorbed by the atmosphere due to CO_2 and other greenhouse gases such as water vapour (H₂O) and methane (CH₄). CO_2 is the most important of the anthropogenic greenhouse gases due to both the amount in the atmosphere as well as the longevity of the gas. As of 2018, the atmospheric concentration of CO_2 is steadily over 400 parts per million (ppm) - daily updated measurements for the famous Manua Loa Observatory can be found at "The Keeling Curve" (2018) (Figure 1.1). The growth rate of atmospheric CO_2 has been over 1.5 ppm yr⁻¹ for the last 15 years (NOAA 2018).

Apart from contributing to the "opaqueness" of Earth's atmosphere to infrared radiation and facilitating warming on a global scale, atmospheric carbon dioxide is a key player in interactions with the biosphere. Carbon, as the basis of all organic life on this planet, is found in various reservoirs. Not only in gaseous form in the atmosphere as CO_2 and CH_4 , but also in bound form in vegetation, soil, the oceans and in permafrost. Carbon is regularly exchanged between reservoirs, via so-called *fluxes*. A forest fire, for example, represents a carbon flux from the vegetation carbon pool into the atmospheric one. Photosynthesis, the process by which plants sequester atmospheric CO_2 into sugars, is another example of a flux with the opposite sign (atmosphere \rightarrow terrestrial biosphere). The sum of all reservoirs and fluxes on Earth



Figure 1.1: The famous Keeling curve, recorded at the Manua Loa Observatory from 1955 onward in weekly intervals (Keeling et al. 2001). The steady rise due to anthropogenic emissions is superposed on the strong seasonal cycle. This cycle arises mainly due to forest leaf-out in the northern hemisphere causing a drawdown in spring. Data taken from "Primary Mauna Loa CO₂ Record | Scripps CO₂ Program" (2018).

is termed the global carbon cycle.

Highlighted and detailed in Wigley et al. (2005), the importance of the carbon cycle has gained attention when in the mid-1990's the Intergovernmental Panel on Climate Change (IPCC) recognised the intricate, yet not well-understood link between emissions and atmospheric concentrations. A schematic overview of the global carbon cycle is shown in Figure 1.2.

According to IPCC (2014), the anthropogenic emissions since 1750 leading to an increase of the atmospheric CO_2 concentration have totalled $555 \pm 85 Pg C$. It is attributable mostly to fossil fuel combustion and cement production $(375 \pm 30 Pg C)$ as well as land use change $(180 \pm 80 Pg C)$. Inspecting Figure 1.2, it becomes clear that the anthropogenic fluxes (fossil fuel burning, cement production and land use change) are dwarfed by both vegetation-driven fluxes and ocean gas-exchange. Photosynthesis-driven gross primary production (GPP) causes the largest biospheric carbon flux, however almost the same flux is seen going into the atmosphere (for details on photosynthesis see Chapter 6, Page 61). The net flux is still negative (from the point of view of the atmosphere), meaning that vegetation, on a global scale, is a *net sink* of carbon. Similar net fluxes are observed at the ocean-atmosphere interface.

Balancing the estimated emissions due to human activities with the increase in atmospheric concentrations of CO_2 leads to a surprising result. Considering the total amount of human emissions from 1750 to 2011 (555 ± 85 Pg C), only about half (240 ± 10 Pg C) remained in the atmosphere! There is comparatively good agreement regarding the carbon fluxes from the atmosphere to the ocean, which account for 155 ± 30 Pg C yr⁻¹. Land use change is less well constrained, and is potentially a net positive flux to the atmosphere (1.1 ± 0.8 Pg C yr⁻¹). This imbalance has been known for some time (Kerr 1977), and considerable effort was made to pin down this missing



Figure 1.2: Simplified summary of the major fluxes and pools of the global carbon cycle. For a more complete illustration see e.g. IPCC (2014, Figure 6-1). Inspired by the same figure. Values for carbon reservoirs are in Pg C, fluxes (arrows) are in $Pg C yr^{-1}$.

sink. One of the most recent syntheses on the global carbon budget was published by Le Quéré et al. (2018). For the year 2016, they state the anthropogenic emissions to be 9.9 ± 0.5 Pg C yr⁻¹ due to fossil fuels and industry, and another 1.3 ± 0.7 Pg C yr⁻¹ from land use change. A change in methodology regarding the missing terrestrial carbon sink was introduced in Le Quéré et al. (2018). In previous versions of the Global Carbon Budget, e.g. Le Quéré et al. (2016), the residual terrestrial sink was calculated through a balance equation including estimated emissions, atmospheric growth rate and the modelled ocean uptake. Le Quéré et al. (2016) state this residual land sink as 1.9 ± 0.9 Pg C yr⁻¹. The new method involves dynamic global vegetation models (DGVMs, previously only used to constrain emissions due to land use change) and treats land carbon uptake explicitly, rather than just a residual. The remaining budget imbalance is therefore much smaller (0.2 Pg C yr⁻¹ for 2016) when ocean sinks (2.6 ± 0.5 Pg C yr⁻¹ for 2016) and land sinks (2.7 ± 1.0 Pg C yr⁻¹ for 2016) are explicitly accounted for.



Figure 1.3 shows how the increasing anthropogenic emissions due to fossil fuels and industry were partially compensated for by land and oceans sinks. The terrestrial carbon sink increased as a result of CO_2 fertilisation, nitrogen deposition and longer growing seasons due to the effects of climate change. The land sink, however, also exhibits fluctuations up to $2 Pg C yr^{-1}$, especially due to El Niño events which reduce the uptake. These events can have lasting effects on the global carbon cycle, as shown e.g. in the study by Poulter et al. (2014). They investigated the impact of the 2010-2011 La Niña on the semi-arid ecosystems in Australia. The La Niña event triggered several years of increased precipitation which led to a "greening" of Australia which then was the major contribution of the 2011 land sink anomaly.

Local fluxes can be measured using various techniques, like the prevalent eddy covariance method which applies concentration measurements alongside with accurate wind direction and wind speed measurements. Carbon flux is derived through the assumed covariance between vertical wind speed and the concentration of carbon dioxide. The so-called *flux towers* are found throughout the globe, and more than 140 of those are part of the FluxNet global network (Baldocchi et al. 2001). Comprehensively discussed in Andrew D. et al. (2006), FluxNet can be exploited to calibrate global production efficiency models¹ (PEMs), which then allow for a globally more dense view on carbon uptake (compared to the rather sparse FluxNet coverage).

The PEM approach, however, essentially relies on a "simple" conversion of the available radiation using other factors determining the uptake rate, such as leaf coverage through the leaf area index (LAI), canopy greenness as described through indices like the normalised differential vegetation index (NDVI), and so forth. The

¹Models which predict carbon uptake by vegetation and are driven by satellite measurements of absorbed photosynthetically available radiation (APAR).

actual atmospheric concentration of CO_2 is not taken into account.

A different approach to inferring carbon fluxes are *flux inversions*. While methodically very different from the eddy covariance technique, the principle behind it is quite similar. Given measured concentrations and knowledge of meteorological conditions, carbon fluxes can be estimated through an inverse procedure, similar to the inverse method explained in Chapter 5 (Page 47). From a prior assumption on fluxes, concentration measurements are used to constrain fluxes with the help of an atmospheric chemistry transport model. Ciais et al. (2000) describe how the inversion technique can be used in combination with 77 GLOBALVIEW-CO2 monitoring sites providing atmospheric concentrations from different sources (surface, airborne, towers). They found the largest land sinks to be in North America and Siberia.

1.2 Space-Based CO₂ Measurements and Flux Inversions

Surface stations and aircraft campaigns provide concentration measurements that make inversion studies possible in the first place. The sparsity of the measurement locations, however, constitute a significant problem - especially when attempting to quantify fluxes on a global scale. A map of the measurement locations (excluding aircraft campaigns) of the most recent ObsPack (NOAA Earth System Research Laboratory 2016) is shown in Figure 1.4. The map highlights the general issue that most measurement infrastructure is found in Europe and North America. South America with its biomass-rich Amazon region is not covered apart from a few surface stations near the coast. Africa and most of Asia are not covered by the ObsPack surface and tower stations.



Figure 1.4: ObsPack (NOAA Earth System Research Laboratory 2016) surface (red circles) and tower (orange squares) measurement locations; aircraft measurements not shown.

What Figure 1.4 means for flux inversion investigations, is that fluxes in some of the

most productive regions on Earth, the tropical forests in Africa and South America, are not constrained by measurements. Ideally, one would want spatially highly dense measurements, across the globe, with a high repetition time. The answer to such a requirement can only come by way of satellite remote sensing. As an example, the moderate-resolution imaging spectroradiometer instruments (MODIS) on board the Aqua and Terra satellites manage to image most parts of the Earth on a daily basis.

The first space-based instrument to infer atmospheric CO_2 concentrations through the measurement of back-scattered solar light in the short-wave infrared (SWIR) region of the electromagnetic spectrum (see Chapter 2, Page 13 for details on the method) was the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) deployed on the ENVISAT satellite (launched 1st March 2002). SCIAMACHY facilitated the first direct observation of the atmospheric CO_2 growth from space (Buchwitz et al. 2007) using SWIR bands. Studies using the thermal infrared region have been performed before by Chedin et al. (2002) using a microwave sounder on board the TIROS-N satellite. However, as Buchwitz et al. (2007) argue in their corrigendum to the article, TIR radiances are strongly dependent on temperature, and only weakly dependent on CO_2 concentration, whereas in the SWIR, it is the other way round.

GOSAT (see Chapter 3, Page 22) was launched in 2009, three years before the ENVISAT satellite was finally lost due to spacecraft failure. In a study published soon after GOSAT's launch, Chevallier et al. (2009) investigated its potential to inform flux inversion systems. They found that GOSAT measurements will help reduce the uncertainty in carbon fluxes by $\sim 30\%$ for most land regions, with a few regions even showing a reduction of $\sim 60\%$. For the planned Orbiting Carbon Observatory mission (Crisp et al. 2004), Chevallier et al. (2007) have investigated the sensitivity of inverted fluxes to systematic regional biases. Their far-reaching conclusion is that measurement biases as low as a few tenths of a ppm can already lead to significant differences in the inferred carbon fluxes, in their case more than 0.8 Pg C yr^{-1} for the Eurasian boreal region.

Almost ten years before the launch of GOSAT, Rayner et al. (2001) investigated the precision requirement of space-based measurements in order to be equivalently performing as surface stations (< 2.5 ppm). As part of the European Space Agency's Climate Change Initiative project for greenhouse gases (GHG-CCI), these requirements were re-evaluated and specified from a "user" point of view; the user here being the inverse modelling community who makes use of space-based concentration measurements. Formulated in Chevallier et al. (2016), the goal requirements on systematic errors (biases) are stated as < 0.2 ppm. Single-sounding uncertanties² are

²Note that Chevallier et al. (2016) use the term *random error* rather than "uncertainty".
required to be smaller than 0.3 ppm. Considering an atmospheric concentration of 400 ppm, the goal values in relative terms amount to 0.5% and 0.75% respectively! These goals, unlike (Rayner et al. 2001), represent the requirements that space-borne CO₂ retrievals need to fulfil in order for flux inversions to be able to quantify carbon sources and sinks on a regional scale. Regarding the relationship between flux inversions and the forest land sinks of the global carbon budget, the study by Pan et al. (2011) highlights the importance of regional carbon budgets, especially for the prediction of the future atmospheric growth rates.

Given these requirements, the development of a robust trace gas retrieval technique is critical. The research in this thesis focuses on two particular topics relevant to these broad goals: solar-induced chlorophyll fluorescence (Part II, Page 58) and fast radiative transfer (Part III, Page 140).

Chapter Two

Remote Sensing of Carbon Dioxide

PACE-BASED remote sensing of atmospheric carbon dioxide is based on the spectroscopic measurement of light in the vicinity of molecular rotational-vibrational modes. This chapter introduces the basic fundamental principles of molecular spectroscopy and how they are exploited in remote sensing applications.

2.1 Molecular Spectroscopy of O_2 and CO_2

The background on molecular spectroscopy follows Haken et al. (2003).

The basic physical principle behind molecular spectroscopy is the quantum mechanical nature of energy levels in molecules. A photon with frequency ν can be absorbed by a molecule if and only if that molecule at energy state E can be excited into a new energy state E' > E, where $\Delta E = E' - E = h\nu$ with h being the well-known Planck constant. This principle gives rise to the typical discrete molecular spectra. The available energy levels depend mainly on the geometry of the molecule and by the degrees of freedom of the possible excitations. A semi-classical approach is helpful in picturing the various excitation modes.

The most intuitive of excitations are rotations: a molecule of any given configuration in a gas phase can be excited to rotate. The number of possible rotations is given by the symmetry of the molecular configuration. Both molecular oxygen (O=O) and carbon dioxide (O=C=O) are linear molecules, meaning that all (double) bond angles within the respective molecule are 180°. This particular symmetry has consequences for the available excitation energies. Considering rotations with rotation axes along the symmetry axis, the associated moment of inertia is negligible compared to others. With negligible moments of inertia, the associated angular momenta can be assumed to be near-zero as well. Ignoring this negligible rotation, linear molecules have two distinct rotational axes with the same moment of inertia.

Molecules are able to vibrate by changing the relative position of one or more atoms in the molecule. There are several possibilities as to how vibrations can manifest themselves, and they increase with the number of atoms. For molecular oxygen, there is only one possible vibration type, the symmetric stretch. As a general rule, linear molecules with N atoms, can undergo 3N - 5 possible vibration types - thus resulting in only one for molecular oxygen (N = 2).

 CO_2 on the other hand possesses $3 \cdot 3 - 5 = 4$ different ways to vibrate, which are illustrated in Figure 2.1.



Figure 2.1: Carbon dioxide vibration types. In Figure 2.1c, the cross and dot symbols represent arrows pointing into and out of the plane respectively. The two bend modes are energetically degenerate.

In gases, rotations and vibrations of molecules appear together, with rotations being energetically much more favourable. From the quantised rotational-vibrational modes, the most likely to occur are therefore rotational transitions from and into in the same vibrational mode. A given vibrational mode is always accompanied by rotational modes.

The total energy of a molecule can be written as a sum of its constituent parts, and so can energy differences due to transitions:

$$\Delta E = \Delta E_{\text{rotation}} + \Delta E_{\text{vibration}} + \Delta E_{\text{electronic}} + \Delta E_{\text{other}}.$$
 (2.1)

Changes in rotational states are related to the change in the total spin angular moment quantum number, ΔJ , describing the change of J during a transition $J' \rightarrow J''$. The rotational energy can be semi-classically approximated as $\Delta E_{\text{rotation}} \propto J'J''$ where the proportionality constant determines the energy spacing between two adjacent levels with the transition $J'' = J' \pm 1$. Vibrational energy states in first approximation are introduced as a harmonic oscillator-type potential, $E_{\text{vibration}}(\nu) \propto (\frac{1}{2} + \nu)$.

Molecular spectroscopy relates the change of excitation state with the frequency of photons via $\Delta E = hv$. In order for a photon to be absorbed by a molecule to then trigger rotations or vibrations, the transition must be favoured by the *selection rules* of electromagnetic transitions. The notion of *allowed* and *forbidden* transitions (transitions which obey selection rules, and those which do not) is slightly misleading, as both can occur - forbidden transitions, however, occur at much lower rates. Selection rules can be derived via first-principles from Fermi's Golden Rule, which discribes the first-order probability of a transition from an initial state $|i\rangle$ to a final state $|f\rangle$ under a small perturbation described by W, the Hamiltonian of the interaction. The corresponding matrix element $\langle i|W|f\rangle$ then provides a proportionality constant for the calculation of the probability with which the transition $|i\rangle \rightarrow |f\rangle$ can occur. Considering electric and magnetic dipole transitions, one finds that $\Delta J = \pm 1$ or $\Delta J = 0$ (however $J = 0 \rightarrow 0$ is not allowed) so that $\langle J'|W|J'' \rangle \neq 0$. For vibrational modes, the selection rules for dipole transitions state that the vibrational quantum number ν must change by at least one $\Delta \nu = \pm 1$.

Analogously, electronic transitions have to be considered. These occur when bound electrons of a molecule move to an unoccupied position. The electronic energy of a molecule is usually written as the sum of all contributions from the electromagnetic interactions between all constituents (*Coulomb Hamiltonian*). An electron moving to a higher energetic state thus changes the total electronic energy of the molecule.

Other components to the total molecular energy, transitional and nuclear, are currently negligible for the purposes of remote sensing.

The symmetric stretching mode of carbon dioxide (Figure 2.1b) is *infrared inactive*, since the permanent dipole moment of the molecule does not change as a result of the oxygen atoms moving symmetrically relative to the carbon atom. The bending modes, however, change the dipole moment of the molecule, thus results light absorption in the thermal infra-red. The most important CO₂ absorption band in that regime, the vibrational mode v_2 , is located at 15 µm (667 cm⁻¹).

2.2 The O₂ A-band, and the Weak and Strong CO₂ Bands

The utilised bands in the GOSAT and OCO-2 missions are the O₂ A-band at 0.76 µm, the weak CO₂ band at 1.6 µm and the strong CO₂ band at 2.06 µm. The O₂ A-band is the result of an electronic transition (at $\nu = 0 \rightarrow 0$ and $\nu = 1 \rightarrow 1$), $X^3 \Sigma_g^- \rightarrow b^1 \Sigma_g^+$ in standard molecular term notation. As summarised by e.g. Wark et al. (1965) and Brown et al. (2000), this is a forbidden magnetic dipole transition. The fact that space-based instruments can still measure this band is solely due to the large abundance of molecular oxygen in the atmosphere (~20.95%) that overcomes the low occurrence of this transition. Figure 2.2 shows an example of absorption cross sections ("GES DISC - ABSCO" 2018) in the O₂ A-band at 290 K temperature and 1120 hPa pressure. It shows the characteristic double-wing structure of rotationalvibrational spectra, centred at around 13125 cm⁻¹. This gap is the location of the



Figure 2.2: Absorption cross sections σ for molecular oxygen in the A-band, as seen by an instrument with GOSAT-like resolution (see Table 3.1).

non-existent *Q-branch*, the term used for transitions in which the total spin angular moment does not change: $J = 0 \rightarrow 0$. For transitions $\Delta J = 0$, the change must come from a change in the rotational axis, of which there are two in the case of O_2 . The O_2 A-band, however, is a transition from one excited Σ state to the triple-degenerate ground state with the same Σ , where Σ denotes the projection of the total electron angular momentum of 0 as projected onto the molecular bond axis. Thus, for this spectral band, there is no Q-branch. Next to the missing Q-branch location, the band is separated into a P- and an R-branch. Both are related to changes in the rotational states via J'' = J' - 1 and J'' = J' + 1 respectively. The different modes in the P-branch are spaced further apart than those in the R-branch as a result of rotationalvibrational coupling. The strength of each line is influenced by the temperature and additionally follow a Boltzmann distribution. Actually three distinct bands can be observed near 13125 cm⁻¹, which are due to oxygen isotopes, for molecules ¹⁶O¹⁷O and ¹⁶O¹⁸O. Due to the low abundance of these molecules, they are not considered. The O₂ A-band features line pairs in both bands, more easily observed in the P-branch due to the larger spacing of pairs. The pairs themselves are essentially a result of electron spin, which adds to the total spin angular momentum J; see for example Geddes (2015, Chapter 2.2) and sources therein.

Molecular absorption lines as shown in Figure 2.2 do not only have a finite width due to instrument limitations, but due to a number of physical effects. Each line has a natural line width, determined by the time-energy uncertainty $\Delta E \Delta t \sim \hbar$, where Δt is the lifetime of the excited state. In addition to this, line broadening occurs due to a combination of temperature and pressure. Temperature-dependent line broadening arises due to the Doppler shift associated with the molecules moving at a certain average velocity and results in a Gaussian line shape. Molecules moving away from the detector would result in the shift of the line to higher wave numbers, but molecules naturally also move towards the detector, causing a blue shift of the line to lower wave numbers. Since the velocity vector directions of molecules average out in the thermodynamic limit, finite temperature leads to a broadening of spectral lines. Pressure broadening is more intricate, but a major contribution is collisions of molecules which results in a Lorentzian shape. Such collisions may cause a molecule to prematurely transition to a different energy state, thus reducing the aforementioned mean time of the excited state. Through the uncertainty principle, ΔE thus increases. An example showing the effect of pressure on the line shape is depicted in Figure 2.3.



Figure 2.3: The effect of pressure broadening demonstrated on a line in the O₂ A-band. As opposed to Figure 2.2, no convolution has taken place and the cross sections are shown on a native grid of 0.01 cm^{-1} . T = 290 K for both.

The so-called weak CO₂ band is located at 1.6 µm (or 6228 cm⁻¹) and originates from a combination of the fundamental vibrational modes v_1 (1388 cm⁻¹, symmetric stretch) and v_3 (2349 cm⁻¹, asymmetric stretch), and overtones of v_2 (667 cm⁻¹, bend). Overtones are transitions from the fundamental vibrational mode to a higher mode $0 \rightarrow n > 1$. For the weak band, the resulting central frequency is given by $v_1 + 4v_2 + v_3$; such a band is called *combination band* as different vibrational modes are excited simultaneously. The cross section for the band at 6228 cm⁻¹ is shown in Figure 2.4, along with spectral lines from CH₄ and H₂O. These gases appear in the weak band along with CO₂ and have to be accounted for in the set-up of the retrieval. As with the O₂ A-band, no Q-branch is present.

Finally, the strong CO₂ band at 2.06 μ m (4 $\nu_2 + \nu_3$) is shown in Figure 2.5. The cross sections in this window are larger than in the 1.6 μ m band (hence the name strong CO₂ band), and methane lines are not present.

For the OCO mission (details introduced later in Section 3.1), it was found that the then-available cross sections were not sufficient to meet the mission's science goals. Thus, the development of new tabulated cross sections was commissioned, and the resulting data is now freely available at the OCO-2 Data Center ("GES DISC - ABSCO" 2018). They are based on the works of Benner et al. (2016), Devi



Figure 2.4: Weak CO_2 band along with interfering species CH_4 and H_2O . The cross sections for methane and water vapour were scaled for better visibility.



Figure 2.5: Strong CO₂ band at T = 290 K and p = 1120 hPa. Compare the ordinate tick labels with Figure 2.4, showing the significant difference in line strengths.

et al. (2016), and Drouin et al. (2017) and many others referenced within. These socalled *ABSCO* tables contain the molecular cross sections in three- or four-dimensional data structures, as a function of wavenumber, temperature, pressure and from version 4.2 onwards also relative humidity. In this thesis, GOSAT retrievals utilised version 4.1.1, whereas the OCO-2 retrievals were performed with the version 5.0 tables. These tables however differ from the publicly available ones, as they were calculated for 71 pressure levels instead of 64 - courtesy of Vivian Payne (JPL).

2.3 Three-Band Approach to Concentration Measurement

GOSAT and OCO-2 (see Chapter 3) both utilise a three-band approach to measure the atmospheric concentration of carbon dioxide. The CO_2 bands described in the previous section are located in the short-wave infrared (SWIR) part of the electromagnetic spectrum, and the O_2 A-band is part of the visible to near-infra red range (VIS-NIR). Compared to the various thermal bands of CO_2 , the SWIR bands have an advantage with regards to the vertical measurement sensitivity. These bands also feature small contamination from other gas species, with only water vapour and methane lines additionally appearing in non-negligible quantities. The measured signal in the SWIR band is sensitive almost all the way down to the surface, whereas CO_2 measurements in the thermal range are most sensitive in the mid-troposphere. This is owed to the fact that absorption in the thermal range is highly dependent on the temperature in a given atmospheric layer, which is not the case in the SWIR range. Thus, through the contribution of all atmospheric layers through which the light travels, sensitivity peaks in the mid-troposphere for thermal bands, and in the lower troposphere for SWIR bands, where the number of CO_2 molecules is highest.

The measurement concept utilised by GOSAT and OCO-2 relies on the inherent relationship between molecular cross sections and the intensity of a light beam travelling through a gas in which said molecular species occurs. Explained in more detail in Section 10.1 (Page 141), the relationship is described through the Beer-Lambert law:

$$I(\lambda) = I_0(\lambda) \exp\left[-\int_0^L \sigma(\lambda) \cdot \mathbf{n}(s) \, \mathrm{d}s\right]. \tag{2.2}$$

Given an initial intensity I_0 , a monochromatic light beam of wavelength λ travelling through a (mono-molecular) gas for a distance L, experiences an decrease in intensity. The attenuation is determined by the cross section $\sigma(\lambda)$ and the number density of molecules n(s) as a function of the geometric path length s. For satellite-based observations, the intensity incident at the top of the atmosphere can be calculated through solar models, or in some cases be measured by the satellite itself. The viewing and solar geometry is generally known, and thus the geometric path through the atmosphere can be determined. If now the cross sections are sufficiently well known, the number density n integrated along the total line of view can be calculated

$$n \sim -\frac{1}{L \cdot \sigma(\lambda)} \ln \frac{I(\lambda)}{I_0(\lambda)}.$$
 (2.3)

In a few words: through measuring the attenuation of the intensity, GOSAT and OCO-2 are essentially counting molecules! The approach in Equation 2.3 is commonly known as differential optical absorption spectroscopy (or DOAS).

Unfortunately, the geometric path as determined by both solar and satellite positions and viewing angles, is not the same as the optical light path. Earth's atmosphere does not just contain gases, but also particles suspended in the air, collectively known as aerosols and clouds. Both are not only absorbing, but also cause light to be scattered. Even the air molecules cause scattering due to the effect of Rayleigh scattering. Several scatter events can cause an increase in the light path and thereby an additional attenuation compared to a non-scattering model calculation. Scattering can also shorten the optical path, if for example high-altitude aerosols can cause light to be scattered towards the satellite. This light will then not traverse the entire atmosphere down to the surface. Not accounting for multiple scattering thus leads (in general) to an overestimation of the gas concentration, while not account for aerosols at all will generally lead to an underestimation.

Rather than inferring the absolute number of CO_2 molecules, it is more appropriate to normalise to the number of air molecules. The reasoning behind this choice is the variation in surface pressure, which obviously changes between measurement locations due to topographic variations, as well as times which change due to weather. Going one step further, the amount of water vapour in the atmosphere can induce similarly high variations. It is thus preferable to define a quantity XCO₂, which is the fraction of CO₂ molecules per molecule of dry air:

$$XCO_2 = \frac{n_{CO_2}}{n_{dry-air}}.$$
 (2.4)

Space-based measurements performed by OCO-2 or GOSAT intrinsically measure the total column of air, hence why Equation 2.4 is referred to as the *column-averaged dry-air mole fraction* of carbon dioxide (along the geometric path s)

$$XCO_2 = \frac{\int n_{CO_2}(s) \, \mathrm{d}s}{\int n_{\mathrm{dry-air}}(s) \, \mathrm{d}s}.$$
(2.5)

An alternative, but equivalent term is *volume mixing ratio*. It is usually given in parts per million $(10^{-6}, \text{ ppm})$. O'Dell et al. (2012) have pointed out that the method of averaging is not unique (e.g. weighting with regards to pressure, or total number of air molecules), thus XCO₂ values might differ depending on the definition. In this thesis, the column weighting is performed using pressure, as described in O'Dell et al. (2012) (see Section 5.4, Page 52).

In the so-called *three-band approach*, the number density of air is informed by the O_2 A-band. The strong oxygen lines in this band carry information about surface pressure (or the oxygen concentration) to then calculate XCO₂ via Equation 2.5. Both weak and strong CO₂ bands naturally carry information on the amount of carbon dioxide in the light path. The reasons for using more than one band for CO₂ are again related to scattering by aerosols. Since the bands are separated by roughly 0.4 µm, the retrieval algorithm can utilise the spectral information to constrain the scattering properties of aerosols, which are potentially wavelength-dependent. The O₂ A-band, being another ~0.85 µm away from the weak CO₂ band, also provides information about aerosols - as the path length increase due to scattering will be different for the various spectral bands.

The term *full-physics* is not rigorously defined. Some refer to full-physics retrievals as algorithms that not only account for scattering by aerosols and clouds, but also attempt to retrieve information about them. In a more general sense, full-physics algorithms can be defined as those algorithms, which attempt to describe every aspect of the forward model through a physical model. Contrary to full-physics algorithms, the DOAS approach treats light scattering as a broadband phenomenon which is represented as an additive polynomial function.

Examples of the spectroscopic measurements are shown in Figure 2.6, which are taken straight from simulations that correspond to a GOSAT measurement location in Alaska, USA.



Figure 2.6: Example of simulated spectra corresponding to the three bands, calculated using the GOSAT instrument model (see Section 3.2). The radiances are normalised with respect to the largest value in the band.

Chapter Three

GOSAT and OCO-2

OTH GOSAT and OCO-2 satellites are dedicated missions to measure concentrations of atmospheric carbon dioxide (and methane for GOSAT). The measurement concept, i.e. the acquisition of high-resolution spectra of the rotational-vibrational bands of CO₂ in the short-wave infrared, is equally applied in both mission concepts. The satellites were originally planned to be launched around the same time, which is why they feature complimentary sampling approaches. GOSAT was designed to quantify continental-scale sources and sinks, whereas OCO-2 provides denser sampling and can even potentially provide information on point-like sources. In this chapter, the missions and instruments are introduced and direct comparisons, where applicable, are made to highlight the distinct characteristics. The two instrument concepts lead to different measurement acquisition rates, which then result in different global sampling strategies, which are discussed in the last section of this chapter.

3.1 Mission and Spacecraft Description

GOSAT (Kuze et al. 2009; Yokota et al. 2009), the first satellite mission dedicated to the measurement of atmospheric carbon dioxide, was developed by the Japanese Aerospace eXploration Association (JAXA), the National Institute of Environmental Studies (NIES), and the Ministry of the Environment of Japan (MOE). It was launched on 3rd January 2009. The satellite was placed into a polar, sun-synchronous low-earth orbit at roughly 666 km altitude. It orbits Earth in about 98.2 min, and the revisit time for a measurement location is every three days. GOSAT's equatorial crossing is at 13:00 local time.

OCO-2 (Crisp et al. 2017), a mission by the National Aeronautics and Space Administration (NASA), is a carbon-copy of the flight instrument OCO which met its premature end as a result of a launch failure (24th February 2009) when the payload fairing failed to separate from the spacecraft. It was finally launched on 2nd July 2014 into a polar, sun-synchronous low-earth orbit at 705 km altitude and 98.8 min orbital period. Nominal revisit time for a location is 16 days, and the equatorial crossing of OCO-2 is at 13:30 local time. OCO-2 is part of the so-called *A-train* (afternoon train), a satellite constellation of six earth observation satellites.

OCO-2 has its ascending node on the illuminated side of Earth, i.e. they travel from south to north for daytime measurements, whereas GOSAT orbits in the opposite direction. GOSAT does have an additional spectral band in the thermal infrared, which makes nighttime measurements of atmospheric CO_2 and CH_4 possible. Mission objectives for both GOSAT and OCO-2 are similar, and science goals can be summarised as follows: the satellites facilitate the measurement of atmospheric carbon dioxide on a global scale, which subsequently allows tracking the sources and sinks of carbon on continental and sub-continental scales.

3.2 The TANSO-FTS and OCO-2 Instruments

The instrument on board GOSAT responsible of measuring the high-resolution spectra is the Thermal And Near-infrared Sensor for carbon Observation Fourier-Transform Spectrometer (TANSO-FTS). This spectrometer type is based on the Michelson interferometry principle: incoming light is divided at a beam-splitter into two arms, at the end of which each beam is reflected back towards the splitter. The combined beam is then directed towards a detector where the intensity is measured. By extending the optical path length of one of beam arms, usually by moving one of the mirrors, and thus inducing a path length difference, interference occurs due to the two light beams having a relative phase difference. Recording the intensity at the detector as a function of mirror position yields a so-called *interferogram*. Translating the interferogram into wavenumber space through an inverse Fourier transformation results in the commonly used spectra.

TANSO-FTS is a double-pendulum type spectrometer, a modification to the cubecorner interferometer. The concept was first published by Jaacks et al. (1989), and operates with the mirrors being statically mounted onto a double-pendulum structure. By rotating the double-pendulum around its flex pivot for a distance x, the optical path length difference changes by 4x (Griffiths et al. 2007, Chapter 5.2.1). A simplified schematic of the TANSO-FTS instrument is shown in Figure 3.1. Light is captured through the two-axis main mirror which can point into both cross-track ($\pm 35^{\circ}$) and along-track ($\pm 20^{\circ}$) directions. On 26th January 2015, the primary mirror was switched off in favour of the secondary one as a result of mechanical degradation. Each scene is also imaged with a CMOS camera that shows the pointing location for every exposure. The light then hits the beam splitter of the double-pendulum interferometer, which induces the interference pattern through an optical path difference of ± 2.5 cm. The combined light then directed towards a collecting and a collimating mirror. For each of the three bands, a separate dichroic filter and band pass is installed, after which the light is split into two polarisation components P and S. This setup allows simultaneous measurement of both P and S components. The first SWIR band is measured through a silicon detector, whereas the second and third SWIR bands operate via InGaAs detectors. TANSO-FTS has a fourth band for measurements in the thermal infra-red, featuring a mercury cadmium telluride detector, however this band is not relevant for the work in this thesis.

The readout mechanism for TANSO-FTS operates through an analogue-digitalconverter (ADC) to turn the electric signal of the interferograms into digital, processable values. As part of the pre-amplifier unit, two different gain circuits are available. For most scenes, TANSO-FTS operates using the *high-gain* mode, only for bright scenes (mostly deserts), the *medium-gain* mode is activated. A slight non-linearity in the analogue-to-digital conversion has been known to cause unphysical zero-level offsets in the spectra, which is the main reason why the fluorescence retrieval requires a bias correction procedure (Section 7.7, Page 83).

The signal-to-noise ratio of a measured interferogram (and the converted spectrum) for a given scene is mainly determined by the integration time, which is set to four seconds. TANSO-FTS is capable of higher acquisition intervals, however there is a lower limit imposed due to the turnaround-time of the pointing system (0.6 s).

The OCO-2 instrument, carrying the same name as the spacecraft, is a grating-type spectrometer. In this instrument concept, light is focused on an optical element with a series of regular grooves. The geometry of the grooves causes a relative path difference between reflected light rays and cause interference at a detector. OCO-2 is equipped with a common telescope for all SWIR bands. Again, three filters direct the main beam to the respective section in the instrument, however the main difference is that each band is equipped with a different grating - compared to TANSO-FTS, where the interference is generated for the common beam. Through the camera lens, the diffracted beam is focused onto a focal plane array (FPA) with a total of 1024 by 1024 pixels. As with TANSO-FTS, OCO-2's band 1 detector is silicon based, whereas bands 2 and 3 are mercury cadmium telluride based (HgCdTe). A simplified version of the optical schematic is shown in Figure 3.2, which represents one of the three spectrometers. The incident light is fed through dichroic splitters and pre-disperser filters that take the role of bandpass filters, and is then directed into the relay optics of the given spectrometer. A polariser retains only the linearly polarised component of the incident light that is perpendicular to the long axis of the entry slit (3 mm by $0.25\,\mu$ m), since the diffraction grating is most effective for this particular polarisation direction. The light diffracted by the grating is then focused through the camera lens onto the cooled FPA. In contrast to TANSO-FTS, the OCO-2 instrument does



Figure 3.1: Simplified TANSO-FTS schematic showing the main components of the instrument, inspired by the layout from JAXA/NIES ("TANSO-FTS - GOSAT - JAXA" 2018).

not feature two distinct gain-modes, meaning that all spectra are recorded using the same pre-amplifier circuit and setting.

TANSO-FTS integrates one scene over four seconds, whereas OCO-2 performs detector readouts with a frequency of 3 Hz. The diffraction grating maps the incoming light onto the two-dimensional detector, where now one dimension corresponds to wavelength, and the other dimension represents a spatial coordinate. Due to imperfect alignment of optical elements, this mapping is not always successful, and spectral and spatial components are mixed. So-called "clocking-corrections" have to be performed for the calibrated spectra in order to mitigate these effects, which would otherwise result in discontinuities in the spectra. The spatial dimension is binned to eight so-called footprints, each of them exhibiting slightly different spectral characteristics. One can consider the eight footprints almost to be eight different instruments - dispersion relation, residual waveforms and other spectrally dependent quantities have to be considered for each footprint independently. This is attributed



Figure 3.2: Simplified schematic of one out of three (with same optical set-up, different components) spectrometers of the OCO-2 instrument. Inspired by Figure 2-3 from the OCO-2 Data User Guide.

to the fact that the radiometric calibration procedure is performed for each footprint individually.

Table 3.1 lists the spectral ranges and resolution for TANSO-FTS and OCO-2. The numbers provided by literature are not intuitively comparable as they are stated in wave numbers for TANSO-FTS (as the more natural units used in Fourier spectroscopy), while values for OCO-2 are in wavelengths. Thus, converted values are given as well. TANSO-FTS covers larger spectral ranges, especially in band 2, as it was designed to measure CH₄ as well. In terms of resolution, TANSO-FTS has higher resolving power ($\lambda/\Delta\lambda$, $\nu/\Delta\nu$) in band 1 (~36 000) compared to OCO-2 (~18 000), and are comparable in band 2 and 3 (~21 000). Even though TANSO-FTS provides higher-resolution spectra, OCO-2 has significantly higher signal-to-noise ratio for comparable scenes.

The instantaneous field of view (IFOV) that TANSO-FTS sees is 15.8 mrad, which equates to a circular ground footprint at nadir with roughly 10.5 km diameter. For OCO-2, the IFOV is 0.1°, and the ground footprint is determined by the integration time for each exposure, and results in a maximal footprint size of 1.29 km by 2.25 km. This footprint size, however, is dependent on the viewing geometry and can be substantially smaller.

Both instruments can operate in similar measurement geometries. TANSO-FTS achieves this exclusively via its pointing mirror(s). Apart from the down-looking nadir-type measurements, ocean scenes are obtained additionally by pointing the mirror towards the wide spot of specular reflection. In the same manner, TANSO-FTS can select predefined targets, track coastlines, and so forth. The OCO-2 instrument has a single static bore sight, which means that pointing the instrument to a specific

	Band 1	Band 2	Band 3	
	O_2 A-band	Weak CO ₂	Strong CO ₂	
Spectral range				
TANSO-FTS	12900 - 13200	5800 - 6400	4800 - 5200	$[cm^{-1}]$
	0.758 - 0.775	1.563 - 1.724	1.923 - 2.083	[µm]
OCO-2	0.758 - 0.772	1.594 - 1.619	2.042 - 2.082	[µm]
	12953 - 13193	6177 - 6273	4803 - 4897	$[cm^{-1}]$
Resolution (FWHM)				
TANSO-FTS	0.3615	0.2575	0.2625	$[cm^{-1}]$
OCO-2	0.042	0.076	0.097	[nm]
Resolving power				
TANSO-FTS	36 000	23500	19000	
OCO-2	18000	21500	21000	

Table 3.1: Spectral ranges and approximate resolution for TANSO-FTS and OCO-2. Values for TANSO-FTS are taken from "TANSO-FTS - GOSAT - JAXA" (2018) (resolution averaged for P and S), values for OCO-2 are from Frankenberg et al. (2015).

location involves rotating the entire spacecraft using its reaction wheels. OCO-2 routinely measures in nadir and glint modes. However, it is also capable of a special target mode in which the instrument rotates during an orbit to keep the bore sight pointed at the target location. This results in several thousands of measurements in the vicinity of a chosen target, however no other nadir or glint-type measurements can be taken on that orbit.

The higher readout frequency of OCO-2 increases the number of measurements in a given time frame, when compared to GOSAT. While TANSO-FTS requires four seconds per exposure, OCO-2 captures eight footprints three times per second. Thus, OCO-2 captures about 96-times more spectra than TANSO-FTS.

3.3 Sampling Strategy

Due to the fundamental differences of the instruments themselves, the overall measurement sampling pattern is different between the missions. GOSAT employs a pointing mechanism, compared to OCO-2's imaging-like approach. Throughout their mission lifetimes, the sampling pattern has undergone various updates. At mission start, GOSAT was measuring along a regular 5-point cross-track (CT) pattern for land scenes, in which 5 locations in a "row" were measured, as shown in Figure 3.3. In August 2010, due to instrument stability issues at the extremes, the pattern was changed to a 3-point mode, in which each row was reduced to three locations, and every location was measured three times. In a further attempt to increase coverage over particularly cloudy areas, a *dithering* strategy was employed from 2014 onward. This resulted in a more densely spaced pattern over problematic regions, like the Amazon. The middle and rightmost panels in Figure 3.3 also shows GOSAT's target capability, in which specific points of interest can be targeted. For the middle panel, target measurement locations were chosen to be along the Brazilian coastline near São Paolo, and in the western Amazon basin for the right panel.



Figure 3.3: GOSAT sampling pattern over the Amazon for three days in September, different years. Inspired by a figure from Webb (2017, Figure 3.14). Starting from a 5-point CT scan (left), GOSAT switched to a 3-point CT pattern (middle) and a special dithering strategy for cloudy regions (right).

The locations in 3-point CT scanning mode are around 150 km apart, and allows mapping continental and sub-continental areas. This measurement density provides a near global coverage for land masses, if the season and solar position permit. In addition to nadir-type land measurements, GOSAT also measures over the ocean in nadir viewing geometry, however these measurements are generally not usable for XCO₂ retrievals as the surface reflectance of water is too low. For ocean measurements, the glint-type geometry is preferred. Here, the instrument is pointed at the specular reflection such that viewing and solar zenith angles are almost the same. The resulting spectra exhibit high signal-to-noise ratios. Due to the geometry constraint of $\theta_0 \approx \theta$, glint-type measurements are only possible within a limited latitude band, which moves during the different seasons of the year. An illustrative example is shown in Figure 3.4, which displays the global measurement pattern for three days during September 2014. From a global perspective, GOSAT covers the main land masses every three days.

Owing to OCO-2's different instrument concept, its sampling strategy also differs from GOSAT. Originally, the global sampling pattern consisted of 16-day intervals, with nadir and glint observation models being exclusively utilised in each interval. From 12th November 2015 onward, the strategy was further optimised to maximise



Figure 3.4: GOSAT sampling for three consecutive days in September 2014 showing the global coverage. These locations are not screened for cloud cover; taken from the UoL-FP pre-processing output, only selected for the main L1B quality flag.

coverage and minimise the gap between nadir orbits. Nadir and glint orbits are now alternating, which is advantageous for SIF retrievals - a gap of 16 days would potentially lead to missed events. Additionally, if an orbit would cover mainly ocean, then the instrument is positioned into glint observation mode. Figure 3.5 thus shows a more dense coverage of the oceans. As opposed to GOSAT, OCO-2 also takes measurements over land in glint mode.



Figure 3.5: OCO-2 global measurement pattern, divided into nadir (left) and glint (right) measurement geometries. These locations have passed a cloud-screening filter as well as other quality criteria; the colour (see legend) represents the day of measurement in September 2016.

Apart from the global sampling pattern, the ground footprint size has additional consequences for the coverage. Figure 3.6 shows the increased measurement density for OCO-2. One OCO-2 swath comprising eight footprints is approximately as wide

as a single GOSAT footprint. Since OCO-2 measures continuously, the spatial vicinity of the various footprints can be used to analyse the retrieved XCO_2 with respect to biases. Assuming that the CO₂ concentration does not vary within a few kilometres (in the absence of strong, local sources), the retrieved XCO_2 should be near-constant for all measurements. This way, retrieval biases related to the surface (e.g. albedo, pressure) can be explored. Despite the OCO-2 swath being comparatively narrow, it does allow for some applications involving an imaging-like approach.



Figure 3.6: Typical GOSAT and OCO-2 ground footprints. The GOSAT scene (southern Illinois, USA) and the OCO-2 footprints (central Australia) are shown in plots at almost equal scale. Background images acquired through the Esri ArcGIS REST Services using the World_Imagery layer.

Nassar et al. (2017) have demonstrated that flux estimates from individual power plants can be inferred from OCO-2 overpasses as long as the overpass was close to the point-like source. They report their flux estimates to be between 1 to 17% of reported daily emissions. While they were using a simple Gaussian plume model, their work shows the potential of future space-based carbon monitoring systems with full imaging capability.

Chapter Four

The University of Leicester Full-Physics Algorithm

4.1 Overview and Workflow Description

RIGINALLY descended from the Orbiting Carbon Observatory (OCO) retrieval algorithm, the University of Leicester Full-Physics (UoL-FP) algorithm was independently further developed by various members of the Earth Observation Science research group at Leicester - mainly Hartmut Bösch, Robert Parker, Austin Cogan, Will Hewson, and recently, myself. As part of the work towards this thesis, I have put substantial effort into adding features to the algorithm, as well as improving on existing ones.

The radiative transfer models in the algorithm, LIDORT (Spurr et al. 2001) and 2OS (Natraj et al. 2007), were updated to the latest respective versions at the time. I have conducted a number of detailed comparisons to ensure that the simulated radiances did not change as a result of this update. To further increase the performance of the algorithm, I have added the TWOSTR (Spurr et al. 2011) radiative transfer model to the algorithm, a fast two-stream multiple-scattering solver for use in fast radiative transfer schemes, which are the main focus of Part III. In addition, I have implemented a new mode in the algorithm which bypasses the sophisticated radiative transfer models altogether. Instead, a pure non-scattering Beer-Lambert-type propagation of the radiances is calculated, including the necessary weighting functions. This particular mode is significantly faster and was used to perform the fluorescence retrievals in Part II. The fluorescence radiance itself was included as a state vector-element in the algorithm such that it can be adjusted for during an XCO₂ retrieval. The implementation of the different fast radiative transfer methods are discussed in the beginning of Part III.

For the processing of OCO-2 measurements, Will Hewson adapted the algorithm to ingest the operational OCO-2 L1B data files. Going forward, I set up and tested the algorithm using both a self-developed pre-processing routine to sample all required meteorological fields, as well a routine that makes use of the publicly available OCO-

2-collocated MET data ("GES DISC" 2018). OCO-2 retrievals also require a residual fitting technique, which I added to the algorithm as additional state vector-elements. These empirical residual waveforms can be used in both OCO-2 and GOSAT retrievals to fit systematic forward model shortcomings, usually related to the spectroscopy or the solar model. Details are explained in Section 12.2.

The term algorithm refers here to the combination of the so-called forward model and the inverse method (Chapter 5). The forward model produces modelled radiances that can be compared to the measured ones, whereas the inverse method is a scheme to infer information from the measurement. The algorithm requires a number of inputs whose preparation is usually not considered to be part of the algorithm. This so-called *pre-processing* involves the gathering of required meteorological data to construct a model atmosphere, calculating an estimate of instrument noise, as well as other scene-dependent quantities (e.g. surface elevation or surface reflectivity). The methods are introduced in Chapter 4 and Chapter 5, and the details on a particular retrieval setup, representative of many scenarios involving XCO_2 retrievals, are found in Section 5.5.

Figure 4.1 shows a simplified workflow diagram of the necessary tasks to run a fullphysics CO_2 retrieval for GOSAT measurements using the UoL-FP algorithm. The very first step involves acquiring the GOSAT L1B data from the GOSAT Data Archive Service (GDAS) ("GOSAT Data Archive Service" 2018). "L1B" signifies that these data packages include geolocated measurements and calibrated radiances. Meteorological data sets include vertically resolved specific humidity and air temperature, as well as geopotential height and surface pressure. Elevation data is derived from the Shuttle Radar Topography Mission (Farr et al. 2007) digital elevation map (30 m resolution). For the vertically resolved prior information on trace gas concentrations (CO_2 , CH_4), the LMDZ global climate model (constrained by in-situ measurements) outputs (Hourdin et al. 2006) acquired through the MACC-II/CAMS services are used. Similarly, aerosol mass mixing ratios (MMRs) for tropospheric aerosols from MACC-II/CAMS are used to inform the aerosol scheme.

From here on, the pre-processor takes in a list of soundings, based on the contents of the GOSAT L1B data, and prepares the algorithm inputs for every scene. This involves sampling (in time and space) all meteorological data, aerosol MMRs and trace gas priors, as well as the surface elevation (altitude) and surface pressure at the specific GOSAT measurement locations. These inputs are all required to compute the properties of the model atmosphere (see Section 4.2). Any required signal degradation corrections to the calibrated radiances are performed by the pre-processor - however



Figure 4.1: Simplified schematic to illustrate the data aggregation and computation tasks required for a full-physics CO_2 retrieval. Data-related items are shown as red rectangles, computational items as blue rectangles with rounded corners and dashed edges. Tasks involving the UoL-FP algorithm are shown as green rectangles with rounded corners. Inspired by the schematic from Detmers et al.

since version V201.201 of the GOSAT L1B data, the radiances in the data files are already corrected. Instrument dispersion, the relationship between detector pixel number and wavelength or wavenumber, is performed by fitting the position of a known, strong solar line. A per-pixel noise-equivalent radiance is assigned to every spectral point, based on the radiances themselves. Finally, the Lambertian albedo of the surface is estimated to provide the algorithm with a sensible prior value.

With the per-sounding preparation in place, the algorithm now has all required inputs to perform retrievals. The cloud-clearing procedure is based on retrieving the apparent surface pressure using exclusively the O_2 A-band. It is assumed that the surface pressure as given by ECMWF is an accurate prediction of the true value. Should the measurement have been obstructed by a thick cloud, the apparent surface pressure should be smaller than the true value. This is due to the high reflectivity of clouds in the SWIR, resulting in a shorter light path, shallower oxygen absorption lines and thus a reduced apparent surface pressure. In the current set-up, a measurement is considered "clear" if the difference between the retrieved (apparent)

and predicted value of the surface pressure is below an empirically derived threshold of $|\Delta p_{\rm surf}| < 20$ hPa. The sub-set of soundings considered to be of low or no cloud contamination is stored in the so-called *clearlist*.

The clearlist is the basis upon which a simplified XCO_2 retrieval is subsequently run. This strong/weak CO_2 retrieval uses an aerosol-free model atmosphere and retrieves a scaling factor to the a-priori CO_2 profile. It is run for the strong and weak CO_2 band independently, and the scaling factor between the two bands are used to compute a ratio. Ideally, the ratio is exactly one, and any deviation from one indicates heavy influence due to aerosol scattering. Since for these retrievals, the data quality of the full-physics retrievals tends to be low, it saves time to not attempt these retrievals in the first place.

Fluorescence retrievals (see Chapter 7, Page 69) are based on a relaxed clearlist¹ to provide fluorescence radiances in the O_2 A-band for the full-physics run.

Running the UoL-FP algorithm generally also requires a set of tabulated absorption coefficient tables (ABSCO, see Section 2.1) in order to calculate the wavelength-dependent transmissivity of the atmosphere.

4.2 Pressure Grid and Model Atmosphere

To avoid confusion about the terminology of "scene", here it is meant to describe all relevant geophysical and instrument-related entities for a specific GOSAT measurement. In the GOSAT documentation, the term scene is also used to specify a group of measurements that are spatially close together and generally form 1/60 of an orbit.

During the pre-processing, the basic elements of a model atmosphere are constructed for the specific location of the GOSAT measurement. Vertically resolved atmospheric quantities are laid out on a pressure level grid: air temperature and volume mixing ratios of CO₂, CH₄ and H₂O (O₂ is assumed constant). The pressure grid is constructed starting from the top of the atmosphere: five stratospheric levels at 10 Pa, 100 Pa, 10 hPa, 50 hPa and 80 hPa. The sixth level is situated halfway between 80 hPa and the tropopause, whereas the seventh level lies at the tropopause itself. The tropopause is estimated from the temperature profile following Reichler et al. (2003). From the tropopause down to the surface pressure, the remaining levels are equally spaced. As a last step, the bottom level pressure is increased by 40 hPa. A sketch of the model atmosphere is shown in Figure 4.2. The model consists of N levels, where the first index 0 refers to the level at the top of the atmosphere. The bottom-most level is at a pressure below the surface pressure, the surface is indicated by the

¹Threshold for $|\Delta p_{surf}|$ at 100 hPa instead of 20 hPa due to SIF being less affected by cloud contamination, see Köhler et al. (2015).



Figure 4.2: Layered model atmosphere which is the basis of the UoL-FP algorithm. Every layer is considered to be optically homogeneous and plane-parallel to the surface. Sun and satellite icons are taken from www.openclipart.org.

thick, brown line. For each level l at pressure p_l , the corresponding temperatures T_l and gas dry-air volume mixing ratios $q_{g,l}$ for species g are obtained through the pre-processing. The main reason for extending the bottom pressure level past the surface pressure is that the UoL-FP algorithm may retrieve surface pressure as well, which translates to a shift in the surface, usually between the two bottom layers N-1 and N. Should the surface pressure p_{surf} shift to lower pressures during the retrieval such that $p_{surf} < p_{N-1}$, the algorithm automatically drops the lowest level and reorganises the model atmosphere. The retrieved surface pressure can also increase (thus moving the surface level in Figure 4.2 "down"), which is the reason why p_N is deliberately chosen as $p_{surf} + 40$ hPa. Should p_{surf} become larger than p_N , the algorithm fails.

While the pressure grid and associated gas volume mixing ratios are defined on levels, the utilised radiative transfer models work exclusively on layers. This introduces some complexity as the model atmosphere has to be translated into layer space, and the resulting Jacobians (see Chapter 5) have to be translated back into level space. There are several reasons as to why the retrieval algorithm was chosen to operate on levels. First, one can argue that the notion of a volume mixing ratio is physically more meaningful at a level, where such a ratio can be unambiguously calculated. Secondly, when two levels are significantly far apart (as is the case for stratospheric levels), the layer representation is a less accurate one. Since absorption cross sections vary strongly and not linearly with temperature and pressure (see Figure 2.3), a large spacing of pressure levels will not capture the variation due to the temperature profile of the atmosphere. Another advantage is that many atmospheric models operate on σ -levels (or σ coordinate system), and thus a translation or comparison between them is easier achieved.

4.3 Extinction due to Gas Absorption and Rayleigh Scattering

The model atmosphere is passed to the radiative transfer solver in layer coordinates rather than levels. Every one of the plane-parallel, optically homogenous layers l is solely characterised by the total optical depth τ_l , the total single-scattering albedo ϖ_l , and the total composite phase function expansion coefficients $\beta_{l,m}$ (*m* being the expansion order). See Chapter 10, Page 141 for details on these quantities and how they enter the radiative transfer problem. For the calculation of gas optical depths, every layer l is subdivided into N_{sub} (in pressure) equidistant sub-layers in order to better capture the temperature and pressure dependence of gas absorption cross-sections. This is especially important at higher altitudes, where pressure levels are far apart. In every sub-layer *m*, the gas absorption optical depth for gas *g* is calculated as

$$\tau_{g,l,m}^{(\text{gas})}(\lambda) = \frac{q_{g,l,m}(1 - q_{(\text{H}_2\text{O}),l,m})}{g_{l,m}M_{\text{dry}}}\sigma(\lambda, p_{l,m}, T_{l,m})\Delta p.$$
(4.1)

 $M_{\rm dry} \approx 28.9644 \,{\rm g}\,{\rm mol}^{-1}$ is the molar mass of dry air, and $g_{l,m}$ is the altitude-dependent local acceleration due to gravity. The (dry-air) volume mixing ratios for sub-layer boundaries are obtained through linear interpolation between the adjacent levels (in log-space), and the mixing ratios q for the sub-layers are simply the average between the sub-layer boundaries. The wavelength-dependent gas absorption coefficient σ for sub-layer-centre temperatures $T_{l,m}$ and pressures $p_{l,m}$ are calculated through linear interpolation based on the tabulated cross section tables at fixed T and p. The total layer gas optical depth for a specific gas is simply the sum of all sub-layer contributions. Finally, the total layer gas optical depth is the sum of all contributions due to the considered gases:

$$\tau_l^{(\text{gas})} = \sum_g \sum_{m=1}^{N_{\text{sub}}} \tau_{g,l,m}^{(\text{gas})}.$$
(4.2)

For the UoL-FP algorithm, $N_{\rm sub} = 10$.

Rayleigh scattering is a direct result of light being scattered by air molecules in the Earth's atmosphere. The presence of these molecules gives rise to incoherent scattering. Strictly speaking, the Rayleigh regime is found where the mean free path between the scatterers is larger than the wavelength of the incident light $d > \lambda$, and under the condition that the scatterers themselves are of negligible size compared to d (Demtröder 2013, Chapter 10). Even though Rayleigh scattering is not a dominant factor in the considered wavelength ranges near ~1 µm, it is not negligible, especially in the O₂ A-band.

The scattering cross-section due to Rayleigh scattering can be fundamentally derived by considering an incident light wave that is scattered by an externally forced, harmonic oscillator with a finite damping term. This physically corresponds to an oscillating electron. The cross-section is heavily wavelength-dependent, see e.g. Bucholtz (1995):

$$\sigma^{(\text{Ray})} = \frac{24\pi^3 [n_s^2(\lambda) - 1]^2}{\lambda^4 N_s^2 [n_s^2(\lambda) + 2]^2} \frac{6 + 3\rho_n(\lambda)}{6 - 7\rho_n(\lambda)}.$$
(4.3)

In above equation, $n_s(\lambda)$ is the refractive index of air and N_s is the molecular number density of air in cm⁻³. $\rho_n(\lambda)$ is the so-called depolarisation factor, a quantity that reflects the (an)isotropy of the considered air molecules. Values for depolarisation factors are taken from Young (1980), whereas the value for $n_s(\lambda)$ is calculated following Allen et al. (1964). The optical depth due to Rayleigh scattering for layer l, $\tau_l^{(Ray)}$, is then

$$\tau_l^{(\text{Ray})} = \sigma^{(\text{Ray})} \frac{N_A \,\Delta p}{M_{\text{dry}} \,g_l}.\tag{4.4}$$

4.4 Aerosols and the Composite Phase Function

Directional scattering properties of an atmospheric layer are characterised by the socalled phase function P, and the fractional amount of scattering is determined by the single-scattering albedo ϖ . P is a function of an incoming and an outgoing angle, so in polar coordinates becomes a function of four parameters: $P(\theta, \phi, \theta', \phi')$, (θ, ϕ) denoting the outgoing, and (θ', ϕ') denoting the incoming direction. As explained in more detail in Section 10.2, phase functions are usually constructed to obey certain symmetries (homogeneity and isotropy of the optical medium) such that they are only dependent on the scattering angle Θ , where $\cos \Theta = \cos \theta' \cos \theta + \sin \theta' \sin \theta \cos(\phi - \phi')$.

The phase function for molecular Rayleigh scattering can be analytically approximated as $P^{(\text{Ray})}(\Theta) = \frac{3}{4}(1 + \cos^2 \Theta)$. Considering the same anisotropy as for the Rayleigh optical depth, the phase function becomes

$$P^{(\text{Ray})}(\Theta) = \frac{3}{4(1+2\gamma)} \left[(1+3\gamma) + (1-\gamma)\cos^2\Theta \right], \tag{4.5}$$

with $\gamma = \rho_n/(2-\rho_n)$. This formulation goes back to Chandrasekhar (1960). Figure 4.3 shows the Rayleigh phase function in a polar representation.

While this particular phase function has a simple analytic form, aerosol phase functions are, in general, much more complicated. Aerosols are large particles when compared to the considered wavelength, and scattering thus belongs to a different regime. For perfectly spherical aerosols, for example, analytic solutions may be found through Mie theory - an exact solution to the Maxwell equations considering the boundary conditions involving a particle with a given refractive index and geometry. Mie theory also extends to any other symmetry or coordinate system for which the Laplace operator can be separated, such as cylinder coordinates or



Figure 4.3: Rayleigh scattering phase function from Equation 4.5.

various other ellipsoid coordinates. For aspherical scatterers whose size is comparable or larger than the wavelength of the incident light, an analytic expression as Equation 4.5 is no longer available. If P is only dependent on the scattering angle Θ , a decomposition using Legendre polynomials P_m as basis functions is possible, where $P(\Theta) = \sum_m \beta_m P_m(\cos \Theta)$, with m being the phase function expansion index, and β_m being the phase function expansion coefficients or Legendre moments.

Deriving a set of β_m for a specific aerosol type using an appropriate Mie solver, will only result in a phase function that is valid for one specific orientation of the particle. For spherical particles, this is sufficient; but for non-spherical scatterers this would create an unrealistic configuration of aerosols in the atmosphere. It is a more realistic assumption to consider aerosols being oriented randomly in the atmosphere, and thus the phase function should reflect this. Accounting for this is achieved through *T-Matrix* methods, described in e.g. Mishchenko et al. (1996) where an analytical averaging over the orientation of particles is performed. One then must also consider the fact that aerosols rarely have the same exact size. Real aerosols are polydisperse, and their particle sizes follow certain distributions - usually some form of power law or log-normal distribution.

The aerosol scheme used utilised in the UoL-FP algorithm was created by Leif Vogel. It is based on the MACC-II/CAMS aerosol system, which incorporates five different tropospheric aerosol groups: sea salt (SS), dust (DU), black carbon (BC), sulphates (SU) and organic matter (OM). The sea salt and dust groups are subdivided into three different size aggregates - for sea salt they are: $[0.03 \,\mu\text{m}, 0.05 \,\mu\text{m}]$ (SS1), $[0.05 \,\mu\text{m}, 5 \,\mu\text{m}]$ (SS2) and $[5 \,\mu\text{m}, 20 \,\mu\text{m}]$ (SS3), for dust they are: $[0.03 \,\mu\text{m}, 0.55 \,\mu\text{m}]$ (DU1), $[0.55 \,\mu\text{m}, 0.9 \,\mu\text{m}]$ (DU2) and $[0.9 \,\mu\text{m}, 20 \,\mu\text{m}]$ (DU3). Organic matter and black carbon groups are subdivided into hydrophobic and hydrophilic subgroups. For the hydrophilic group, the size of the particles are a function of the specific humidity. The MACC-II/CAMS aerosol properties for dust are derived from spherically-shaped particles. As this was found to be an unrealistic approach, the nonspherical dust

types P19 and P21 from MISR² (R. A. Kahn et al. 2015) retrievals are incorporated in a replacement scheme (explained later in this section). A third aerosol mixture is added to represent high-altitude cirrus clouds, where the mean altitude is derived from climatology and a function of latitude (Eguchi et al. 2007).



(c) Black carbon and organic matter

(d) Sulphates and aspherical dust types P19 and P21.

Figure 4.4: Phase functions of the base types - radial components of the plot are logarithmically scaled. The hydrophilic black carbon type is not displayed, as it is almost overlapping the hydrophobic type.

Figure 4.4 displays the phase functions for the MACC-II/CAMS base types at 755 nm, and at 0% relative humidity. The shape of the phase functions reveal some of the fundamental properties of the scattering process itself. Larger aerosols (DU3, SS3, P19, P21) exhibit a prominent forward peak ($\Theta = 0^{\circ}$), compared to smaller aerosols.

Measured spectra in the three-band approach for both GOSAT and OCO-2 do not contain enough information to distinguish contributions from the different base types, so it is convenient to aggregate the base types into mixtures. Dependent on the

²Multi-angle Imaging SpectroRadiometer, a space-based mission dedicated to studying aerosols

relative abundance of each aerosol type, two different mixtures are constructed: one large, and one small type. The small type mixture consists of both black carbon types, both organic matter types, sulphates as well as a small dust type. The large type mixture consists of all sea salt types and a large dust type. The briefly mentioned replacement procedure for dust replaces the three MACC-II/CAMS dust types by two dust types: a medium (P19) and a coarse type (P21). The replacement is done in such a way that the scene-dependent total aerosol optical depth due to dust is retained. The aerosol scheme provides the β_m for the two mixtures separately, along with so-called aerosol fractions for scattering ($q^{(sca)}$) and extinction ($q^{(ext)}$) at several reference wavelengths near the edges of the three GOSAT/OCO-2 bands, as well as the aerosol optical depths ($\tau^{(aer,ref)}$) at (usually) another reference wavelength.

The radiative transfer models utilised by the UoL-FP algorithm expect total phase function expansion coefficients. Therefore, for a scene in which more than one aerosol mixture is present, the phase functions need to be converted into a composite one. To calculate the appropriate composite β_m , the layer-resolved optical depths of the aerosols need to be considered.

First, aerosol fractions $q_a^{(\text{ext})}$ and $q_a^{(\text{sca})}$, where the index *a* refers to one of N_a aerosol mixtures, need to be interpolated to every spectral point in the given band. This is achieved using the Ångström exponent α

$$\alpha = -\frac{\ln\left(q_a^{\text{(beg)}}/q_a^{\text{(end)}}\right)}{\ln\left(\lambda^{\text{(beg)}}/\lambda^{\text{(end)}}\right)},\tag{4.6}$$

where q_a is either $q_a^{(\text{ext})}$ or $q_a^{(\text{sca})}$ at the reference wavelengths $\lambda^{(\text{beg})}$ and $\lambda^{(\text{end})}$ for one band. To obtain q_a at any given wavelength λ_i in that band, the exponent is used:

$$q_{i,a} = q_a^{\text{(beg)}} \left(\frac{\lambda_i}{\lambda^{\text{(beg)}}}\right)^{\alpha} = q_a^{\text{(end)}} \left(\frac{\lambda_i}{\lambda^{\text{(end)}}}\right)^{\alpha}.$$
(4.7)

In a similar fashion, the $\beta_{m,a}$ for mixture *a* is originally given at the two reference wavelengths. Calculating the coefficients for every spectral point *i* is too expensive to be feasible for a retrieval scheme, hence why also here an interpolation is performed. Over such a small spectral range, linear interpolation is considered to be sufficient to capture the spectral dependency:

$$\beta_{i,m,a}^{(\text{aer})} = (1 - c_i) \cdot \beta_{m,a}^{(\text{aer,beg})} + c_i \cdot \beta_{m,a}^{(\text{aer,end})}, \qquad (4.8)$$

with

$$c_i = \frac{\lambda_i - \lambda^{(\text{beg})}}{\lambda^{(\text{end})} - \lambda^{(\text{beg})}}.$$
(4.9)

To obtain the extinction and scattering optical depth at any given wavelength, the reference profile is multiplied by the extinction and scattering coefficients, respectively:

$$\tau_{i,l,a}^{(\text{aer-ext})} = q_{i,a}^{(\text{ext})} \cdot \tau_{l,a}^{(\text{aer,ref})}, \qquad (4.10)$$

$$\tau_{i,l,a}^{(\text{aer-sca})} = q_{i,a}^{(\text{sca})} \cdot \tau_{l,a}^{(\text{aer,ref})}, \tag{4.11}$$

Note that the reference wavelength for the aerosol optical depth can be different from both $\lambda^{(beg)}$ and $\lambda^{(end)}$, and is calculated only once for the entire retrieval, as opposed to each band. The partial contributions due to each aerosol mixture *a* are summed up to obtain the total aerosol optical depth due to extinction and scattering in any given layer:

$$\tau_{i,l}^{(\text{aer-ext})} = \sum_{a=1}^{N_a} \tau_{i,l,a}^{(\text{aer-ext})},$$
(4.12)

$$\tau_{i,l}^{(\text{aer-sca})} = \sum_{a=1}^{N_a} \tau_{i,l,a}^{(\text{aer-sca})}.$$
(4.13)

The total composite phase function expansion coefficients are composed according to the fraction of scattering that each aerosol mixture contributes:

$$\beta_{i,l,m} = \frac{\tau_{i,l}^{(\text{Ray})} \cdot \beta_{i,l,m}^{(\text{Ray})} + \sum_{a=1}^{N_a} \tau_{a,i,l}^{(\text{aer}-\text{sca})} \cdot \beta_{i,m,a}^{(\text{aer})}}{\tau_{i,l}^{(\text{Ray})} + \tau_{i,l}^{(\text{aer}-\text{sca})}}.$$
(4.14)

So far, the expansion coefficients were discussed only for the phase function, rather than the phase matrix, which is the generalisation that takes into account the polarised state of light. The formalism, however, is analogous for the other elements of the phase matrix.

Finally, single-scattering albedos $\varpi_{i,l}$ are calculated using the aerosol scattering coefficient $q_{i,a}^{(sca)}$:

$$\varpi_{i,l} = \frac{\tau_{i,l}^{(\text{Ray})} + \sum_{a=1}^{N_a} q_{i,a}^{(\text{sca})} \cdot \tau_{l,a}^{(\text{aer,ref})}}{\tau_{i,l}},$$
(4.15)

with $\tau_{i,l}$ being the total layer optical depth

$$\tau_{i,j} = \tau_{i,j}^{(\text{gas})} + \tau_{i,j}^{(\text{aer-ext})} + \tau_{i,j}^{(\text{Ray})}.$$
(4.16)

With these quantities, the model atmosphere is fully defined for the radiative transfer calculations.

Fluorescence Contributions to the O₂ A-band

As an additive contribution to the TOA radiance, solar-induced chlorophyll fluorescence (SIF, see Part II) emanating from the surface is added *after* the radiative transfer calculations have taken place. Only the O₂ A-band is considered, as the fluorescence radiances are negligible in the other two bands. Over the small range of the spectral band, the fluorescence radiance is assumed to vary only linearly, hence a two-parameter characterisation is chosen for the signal at the bottom of the atmosphere $F_{\rm s}^{\rm BOA}(\lambda)$:

$$F_{\rm s}^{\rm BOA}(\lambda) = f_{\lambda_2} + f_{\lambda_2} \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_1} (1 - \frac{1}{f_{\rm ratio}}), \qquad (4.17)$$

where f_{λ_2} is the fluorescence radiance at a reference wavelength ($\lambda_2 = 755 \text{ nm}$) and f_{ratio} is the ratio between f_{λ_2} and f_{λ_1} ($\lambda_1 = 772 \text{ nm}$), and usually around ~1.8. To propagate the surface radiance to the top of the atmosphere, a simple non-scattering approach is utilised, where the signal is modified by gas absorption:

$$F_{\rm s}(\lambda) = F_{\rm s}^{\rm BOA}(\lambda) \exp\left(-\frac{\tau^{\rm (gas)}(\lambda)}{\cos\theta}\right),\tag{4.18}$$

with θ being the satellite viewing zenith angle. This is a novel addition to the UoL-FP algorithm and was added as part of the research for this thesis. It was done in particular to facilitate the investigation of the effect of fluorescence on XCO₂ retrievals (Chapter 8, Page 102).

Radiative transfer calculations yield sun-normalised transmittances, which have to be multiplied by a solar spectrum. In the UoL-FP algorithm, the solar model consists of a Planckian black-body spectrum with atmoic spectral lines overlaid. The UoL-FP algorithm uses an empirical line list, details are described in Boesch et al. (2006) and sources therein.

4.5 Instrument Model

The radiative transfer calculations yield high-resolution top-of-atmosphere radiances at a wavelength or wavenumber grid matching the native resolution of the tabulated absorption coefficients. The currently used tables feature a spectral resolution of 0.01 cm^{-1} , so the number of spectral points per band is between ~30 000 for the O₂ A-band and ~16 000 for the strong CO₂ band. Incorporating the finite resolution of the instrument in question is done via so-called *instrument line shape functions* (ILSs). Within the remote sensing community, the ILS is also often referred to as the *instrument spectral response function* (ISRF), *instrument transfer function* (ITF) or *slit function*. It is conceptually the same idea that is behind the optical transfer function in optics, or the *impulse response* in electrical engineering. The ILS characterises the response of the instrument to an ideal (δ -like), monochromatic input signal. Given a high-resolution intensity $I(\lambda)$, the intensity measured by the instrument $I^{\text{conv}}(\lambda)$ is obtained through convolution with the ILS $h(\lambda')$:

$$I_{\rm conv}(\lambda) = \int_{-\infty}^{\infty} I(\lambda') \ h(\lambda - \lambda') \, \mathrm{d}\lambda', \qquad (4.19)$$

which is translated for discrete intensities at spectral indices i as

$$I_{\text{conv},i} = \frac{\sum_{i'} I_{i'} h_{i-i'} \Delta \lambda_{i'}}{\sum_{i'} h_{i'} \Delta \lambda_{i'}}.$$
(4.20)

ILSs for GOSAT and OCO-2 look fundamentally different due to the nature of the instruments (see Section 3.2). TANSO-FTS is a Fourier-transform type instrument, hence the ILS takes the shape of a $\sin(x)/x$ function as the Fourier transform of a boxcar function. The ILS for OCO-2 is closer to a Gaussian shape. While they are generally the best-characterised aspect of an instrument, deficiencies in the application or mis-characterisation can give rise to large biases in the retrieved concentrations, as pointed out in a study by Connor et al. (2016). There are several ways to include the ILS into the UoL-FP algorithm - either as a parametrised function (Gaussian), or tabulated for specific wavelengths that can be interpolated between to obtain the ILS at a given wavelength. The L1B data packages for OCO-2 include a tabulated ILS for each spectral index.



Figure 4.5: Example of the GOSAT and OCO-2 instrument line shape functions. The GOSAT ILS (left) is taken for the reference wavenumber $13\,200 \,\mathrm{cm}^{-1}$ (O₂ A-band), whereas the one for OCO-2 (right) is taken from the same band, for footprint 1, and the first detector sample.

Before the convolution takes place in the algorithm, optional corrections to the spectrum can be performed, such as adding an additive radiance signal (*zero-level offset*), or scaling the entire spectrum by a constant factor (*continuum scaling*).

After the corrections and convolution, the radiances on the instrument spectral grid can be further altered by employing the residual fitting technique. Mentioned in Frankenberg et al. (2015), it is a method to mitigate systematic deficiencies of the forward model. Analysing the spectral residuals³ of a large ensemble of retrievals, it was found that several systematic spectral shapes were common for many scenes. It became clear that these shapes can be related to imperfections in the modelling of the gas absorption coefficients. A technique to mitigate these effects is based on these systematic shapes, which can be extracted from the residuals using e.g. principal component analysis. N residual waveforms $R_k(\lambda_i)$ are added back to the radiances

$$I_{\text{conv}}(\lambda_i) = I_{\text{conv}}^*(\lambda_i) + \sum_{k=1}^N r_k R_k(\lambda_i), \qquad (4.21)$$

where $I_{\text{conv}}^*(\lambda_i)$ are the convolved radiances *before* the addition of the residual waveforms. The scaling coefficients r_k are then part of the state vector to be retrieved and will generally scale linearly with the brightness of the scene.

Finally, mapping the spectral points i onto a wavelength or wavenumber grid is achieved through a polynomial function

$$\lambda_i = \sum_{j=0}^N i^j d_j \tag{4.22}$$

and dispersion coefficients d_j . For most retrievals, a good first guess of the d_j is vital to the success of the algorithm - if the dispersion relation is off by more than a pixel⁴, the algorithm will most likely not be able to correct for that discrepancy. This is mainly due to the strong non-linearity of the forward model with respect to the dispersion relation. Sub-pixel deviations, however, can usually be accounted for by retrieving a dispersion shift (d_0) and/or a dispersion stretch (d_1).

4.6 Derivatives with Respect to Atmospheric and Surface Parameters

In Chapter 5, the inverse scheme to solve the retrieval problem is introduced. In addition to modelling the radiances I as they would be measured by the satellite, the retrieval algorithm also requires partial derivatives with respect to certain atmospheric, surface or instrumental parameters. There is a potential confusion regarding the terminology - the partial derivatives of the radiances with respect to atmospheric or surface parameters are usually referred to as *weighting functions*, whereas the partial derivatives with respect to state vector elements (see Section 5.1) are called *Jacobians*.

³The difference between final modelled and convolved radiances, and the radiances measured by the instrument, $I^{\text{conv}} - I^{\text{meas}}$.

⁴Here, "pixel" refers to a discrete unit at which a radiance value is given in L1B data, which is not necessarily the smallest discrete unit on the detector.

However, this distinction is not always made and the two terms are sometimes used interchangeably.

The UoL-FP algorithm employs monochromatic radiative transfer models, which produce Stokes vectors⁵ given a model atmosphere configuration: $\mathbf{S} = \mathbf{S}(\tau, \varpi, \beta, \mathbf{s})$. τ, ϖ, β are the layer-resolved total optical depths, total single-scattering albedos and composite phase function expansion coefficients. The vector \mathbf{s} contains the set of surface parameters required by the surface model. In the simplest case, and regularly employed for XCO₂ retrievals, this is the Lambertian albedo ρ . The utilised RT models are all linearised, meaning that not only the Stokes vector is calculated by an RT model call, but upon request, the partial derivatives with respect to the (layerresolved) atmospheric and surface quantities are returned. Applying the chain rule using these derivatives, the partial derivatives with respect to state vector elements can be analytically calculated.

To calculate the Jacobians for gas concentration g, first the derivatives for gas optical depth at layer l are written as

$$\frac{\partial \mathbf{S}}{\partial \tau_{g,l}^{(\text{gas})}} = \frac{\partial \mathbf{S}}{\partial \tau_l} \frac{\partial \tau_l}{\partial \tau_{g,l}^{(\text{gas})}} + \frac{\partial \mathbf{S}}{\partial \overline{\omega}_l} \frac{\partial \overline{\omega}_l}{\partial \tau_{g,l}}.$$
(4.23)

The red-coloured derivatives of the Stokes vector with respect to the total optical depths in layer l are provided by the radiative transfer model. From Equation 4.15, $\partial \varpi_l / \partial \tau_{g,l} = -\varpi_l / \tau_l$, and $\partial \tau_l / \partial \tau_{g,l}^{(\text{gas})} = 1$:

$$\frac{\partial \mathbf{S}}{\partial \tau_{\sigma l}^{(\text{gas})}} = \frac{\partial \mathbf{S}}{\partial \tau_l} - \frac{\partial \mathbf{S}}{\partial \varpi_l} \frac{\varpi_l}{\tau_l}.$$
(4.24)

From the derivatives of the layer-dependent optical depths, the derivatives with respect to level concentrations $q_{g,l}$ are obtained through the formalism detailed in Spurr et al. (2014). Similar considerations are made for the temperature Jacobian. Calculating the surface pressure Jacobian is more intricate, as all contributions to the optical depth depend on it, however the general approach is the same: establishing partial derivatives $\partial \tau_{g,l}^{(\text{gas})} / \partial p_{\text{surf}}$ and $\partial \varpi_l / \partial p_{\text{surf}}$ and applying the chain rule. Details are found in Crisp et al. (2010).

Jacobians for the layer-resolved aerosol optical depths for mixture a follow the same approach, however the phase function contributions need to be taken into account:

$$\frac{\partial \mathbf{S}}{\partial \tau_{l,a}^{(\text{aer-ext})}} = \frac{\partial \mathbf{S}}{\partial \tau_l} - \frac{\partial \mathbf{S}}{\partial \varpi_l} \frac{\varpi_l}{\tau_l} + \sum_{s=1}^{6} \sum_{m=0}^{LM} \frac{\partial \mathbf{S}}{\partial \beta_{l,m,(s)}} \underbrace{\frac{\sigma_{l,a}^{(\text{aer})} \left(\beta_{m,a} - \beta_{l,m,(s)}\right)}{\tau_l^{(\text{Ray})} + \tau_l^{(\text{aer-sca})}}}_{\frac{\partial \beta_{l,m,(s)}}{\partial \tau_l^{(\text{aer-ext})}}}, \quad (4.25)$$

⁵See Section 10.1, Page 141 for the definition of the Stokes vector.

with (s) being one of the six independent elements⁶ of the phase matrix. Derivatives with respect to the phase function expansion moments $\partial S / \partial \beta_{l,m,(s)}$ can be calculated by the radiative transfer models for all l,m and (s), however it would be computationally very costly to do so. Each requested partial derivative adds to the computation time, and the total number of Legendre moments LM is usually of the order several dozen or hundreds. As a consequence of the linearised aspect of the used radiative transfer models, the double sum is calculated internally, if the input $\partial \beta_{l,m,(s)} / \partial \tau_{l,a}^{(aer-ext)}$ is provided to the models. From here on, the conversion from layer to level derivatives (Spurr et al. 2014) is applied to obtain the aerosol concentration derivatives at the layer boundaries.

Surface parameter Jacobians are straightforward for the case of a Lambertian surface, which is defined solely by the albedo ρ . The radiative transfer models return $\partial S/\partial \rho$ directly, so no further calculations are necessary. In the UoL-FP algorithm, the surface albedo can be defined as a general polynomial of order N_{ρ} , as a function of wavelength λ :

$$\rho(\lambda) = \sum_{k=0}^{N_{\rho}} \rho^{(k)} (\lambda - \lambda_0)^k, \qquad (4.26)$$

 λ_0 being a reference wavelength within the band, and $\rho^{(k)}$ being the coefficients, where $\rho^{(0)}$ is the albedo, $\rho^{(1)}$ is the albedo slope and so on. Applying the chain rule again, the derivatives with respect to the coefficients are obtained as

$$\frac{\partial \mathbf{S}}{\partial \rho^{(k)}} = \frac{\partial \mathbf{S}}{\partial \rho} \frac{\partial \rho}{\partial \rho^{(k)}} = \frac{\partial \mathbf{S}}{\partial \rho} (\lambda - \lambda_0)^k.$$
(4.27)

Other Jacobians which can not be easily computed using atmospheric or surface weighting functions are obtained through finite differencing

$$\frac{\partial \mathbf{S}(x)}{\partial x} \approx \frac{\mathbf{S}(x + \Delta x) - \mathbf{S}(x)}{\Delta x}.$$
(4.28)

This approach is feasible if the radiative transfer calculations are not required at this point anymore, but relate to any manipulation to the high-resolution spectra. The Stokes vector has to be evaluated twice for every finite-difference Jacobian. Examples for which this is the preferred method are the instrument dispersion and the parameters of a parametrised instrument line shape function. Both are highly non-linear modifications to the final convolved spectrum and thus the change in the radiances are best characterised by calculating \mathbf{S} after a small perturbation of the relevant state vector element.

⁶As a result of the symmetry and homogeneity of the aerosol mixtures. See Thomas et al. (2002) and Mishchenko et al. (2006).

Chapter Five

Inverse Method



FTER establishing the forward model of the UoL full-physics algorithm, a scheme is needed to extract information from the measurement. A task like this falls under the general class of *inverse problems*, a well-studied area

of research. The UoL-FP algorithm employs a Bayesian optimal estimation technique - the motivation, and details are discussed following Rodgers (2000).

5.1 Inverse Problem Formulation

The relationship between a forward model $\mathbf{F}(\mathbf{x}, \mathbf{b})$, which is a function of the *state* vector \mathbf{x} and *auxiliary parameters* \mathbf{b} , and the measurement \mathbf{y} is postulated as

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon}, \tag{5.1}$$

where ϵ is the random measurement noise. The state vector **x** is a set of parameters that are unknown and are to be informed by the measurement, for example a trace gas concentration or surface reflectivity. Auxiliary parameters **b** are required by the forward model, but are considered either known or not relevant in the current retrieval problem, such as the viewing geometry or the surface altitude. As a shorthand notation, the explicit mention of the auxiliary parameters can be dropped to write $\mathbf{F}(\mathbf{x})$.

Given some assumptions on the measurement noise, the equation above would require an inverse function \mathbf{F}^{-1} in order to be solved trivially. For most applications, such an inverse does not exist. The non-existence of a general inverse to the forward model \mathbf{F} gives rise to the many techniques to infer \mathbf{x} from a given \mathbf{y} .

Most strategies for solving Equation 5.1 require knowledge about the sensitivity of the forward model with respect to a change in the state vector. Mathematically, this corresponds to the Jacobian matrix, which is the first order derivative

$$\mathbf{K} = \frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}}.$$
 (5.2)

To demonstrate the usage of the Jacobian matrix, one can picture a case in which the forward model is fully linear in \mathbf{x} . Following (first-order) Taylor expansion around
any state vector \mathbf{x}_0 is then exact:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + \frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{x}_0).$$
(5.3)

The above then reduces the inverse problem (dropping the measurement noise for the time being) to following form

$$\mathbf{y} = \mathbf{F}(\mathbf{x}_0) + \frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{x}_0) = \mathbf{F}(\mathbf{x}_0) + \mathbf{K}_0(\mathbf{x} - \mathbf{x}_0),$$
(5.4)

where \mathbf{K}_0 is the Jacobian matrix evaluated at \mathbf{x}_0 . The final state vector can then be obtained as

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{K}_0^{-1} \left[\mathbf{y} - \mathbf{F}(\mathbf{x}_0) \right].$$
 (5.5)

Thus, for the case of a fully linear retrieval problem, both the forward model and the Jacobian matrix need to be evaluated only once for one given state vector \mathbf{x}_0 .

In practice, the linearity assumption does not hold for many forward models and state vectors. Formally, however, any forward model can be linearised at a given \mathbf{x} ,

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + \mathbf{K}_0(\mathbf{x} - \mathbf{x}_0)$$
(5.6)

where the linearity assumption approximately holds in a (small) region around \mathbf{x}_0 .

When **F** is not linear, solving Equation 5.1 for **x** is often achieved through an iterative technique: starting from a *first guess* for the state vector, \mathbf{x}_0 , the output of the forward model $\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0)$ is compared to the actual measurement **y**. After assessing the difference between **y** and \mathbf{y}_0 as well as between **x** and \mathbf{x}_0 , the state vector is updated (\mathbf{x}_1) and the process is repeated. Once the difference between the forward model output and the measurement is below a certain threshold, usually involving the expected measurement noise, the iterative process is considered to be converged. The state vector after the last update is then the *retrieved* or *final* state vector and represents a possible solution of Equation 5.1.

The inverse problem (Equation 5.1) is turned into an optimisation problem, for which there are again many methods. For a full-physics retrieval algorithm, there are some criteria which the optimisation scheme has to fulfil. For one, the number of forward model evaluations should (on average) not exceed more than e.g. 10. As each evaluation of \mathbf{F} (including Jacobian calculation) takes roughly a minute, it is not feasible to utilise optimisation schemes related to Markov chain Monte Carlo methods, which generally require many evaluations per iteration. Due to the necessity of an initial guess, it is also important that the optimisation scheme is wrapped by a statistical inference method, to allow for an interpretation of the results. A commonly used technique, and the one incorporated into the UoL-FP algorithm, is Bayesian optimisation in combination with the Gauss-Newton minimisation method.

5.2 Bayesian Optimal Estimation

The premise of Bayesian optimal estimation (OE) is Bayes' theorem, which provides a way of calculating and interpreting the probability of an event E, given that H is known to be true. P. M. Lee (2012) uses the term "degree of belief" to describe the meaning of Bayesian probability. The notion of H, *prior knowledge*, is what distinctly differentiates Bayesian statistics from frequentist statistics, in which probability is related to the relative occurrence of an outcome given a large number of trials. The theorem can be formally stated as

$$P(E|H) = \frac{P(H|E)P(E)}{P(H)}.$$
(5.7)

The probability of event *E* occurring, while *H* is known P(E|H), is computed using P(H|E), the conditional probability of event *H* assuming that *E* is true, P(E) and P(H) being the unconditional probabilities of *E* and *H* occurring independently of each other.

Applying Equation 5.7 for the inverse retrieval problem is very easily done. The event E is replaced by the state vector \mathbf{x} , and H is replaced by the measurement \mathbf{y} :

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}.$$
(5.8)

The probability P now takes the more concrete meaning of a *probability density function* (PDF). To solve an inverse retrieval problem, the desired quantity is the *posterior* distribution $P(\mathbf{x}|\mathbf{y})$, which is the probability of a state vector \mathbf{x} resulting in a measurement \mathbf{y} . Similarly, $P(\mathbf{y}|\mathbf{x})$ is the PDF describing knowledge about the measurement given a certain state, and is thus straightfowardly related to measurement noise. $P(\mathbf{x})$ is the probability distribution of the state vector without the measurement, it represents the *prior* knowledge about the state. Finally $P(\mathbf{y})$ is related to our prior knowledge about the measurement itself and is formally obtained by integration of $P(\mathbf{y}|\mathbf{x})$ over all \mathbf{x} . This integration is in practice difficult to do, and serves as a normalisation factor which in this case is not necessary as a specific PDF type will be chosen. It is characteristic of the Bayesian approach that every variable, be it an element from either \mathbf{x} or \mathbf{y} is represented by a PDF.

Rodgers (2000) emphasises that $P(\mathbf{x}|\mathbf{y})$ does not yield a result in terms of an inversion of **F**, unless a specific **x** from the distribution is picked. The Bayesian framework produces some insight as to how the measurement informs and potentially improves on the prior knowledge of **x**.

At this stage, Equation 5.8 is a formulation independent of the type of PDF that describes either $P(\mathbf{x})$, $P(\mathbf{y})$ or $P(\mathbf{y}|\mathbf{x})$. For practical reasons, a specific kind of PDF has to be picked. A Gaussian distribution is the algebraically simplest choice and

has some physical footing - e.g. measurement noise is often of Gaussian nature. The multivariate Gaussian distribution function in n dimensions for a random **u** is

$$P(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^n \mathbf{S}_u}} \exp\left[-\frac{1}{2}(\mathbf{u} - \overline{\mathbf{u}})^{\mathrm{T}} \mathbf{S}_u^{-1}(\mathbf{u} - \overline{\mathbf{u}})\right],$$
(5.9)

where $\overline{\mathbf{u}}$ is the mean of the PDF, and \mathbf{S}_u is the symmetric covariance matrix. Using the explicit representation of a PDF in Equation 5.9, Bayes' theorem can be applied by writing $P(\mathbf{x})$, $P(\mathbf{x}|\mathbf{y})$ and $P(\mathbf{y}|\mathbf{x})$ down explicitly as well.

After some elementary algebraic manipulations (see Rodgers (2000) for details), the two characteristic parameters of the Gaussian posterior PDF $P(\mathbf{x}|\mathbf{y})$ can be calculated

$$\hat{\mathbf{S}} = (\mathbf{K}^{\mathrm{T}} \mathbf{S}_{\epsilon}^{-1} \mathbf{K} + \mathbf{S}_{a}^{-1})^{-1},$$

$$\hat{\mathbf{x}} = \mathbf{x}_{a} + \mathbf{S}_{a} \mathbf{K}^{\mathrm{T}} (\mathbf{K} \mathbf{S}_{a} \mathbf{K}^{\mathrm{T}} + \mathbf{S}_{\epsilon})^{-1} (\mathbf{y} - \mathbf{K} \mathbf{x}_{a}).$$
(5.10)

Understanding that the result of the Bayesian approach is $P(\mathbf{x}|\mathbf{y})$, rather than a particular \mathbf{x} , raises the question of how to represent the result. For practical reasons, it is more convenient to pick a state vector from $P(\mathbf{x}|\mathbf{y})$ - but which one? An obvious choice is to pick out the state \mathbf{x} for which the PDF exhibits a maximum, another one would be the mean state, which includes averaging all states, weighted by the PDF. If the PDFs are reduced to Gaussian distributions, then the two are mathematically the same. The former solution is called *maximum a posteriori* (MAP) state, the latter solution is the *expected value state*. If $P(\mathbf{x}|\mathbf{y})$ is Gaussian, then the mean state of the PDF, $\hat{\mathbf{x}}$ from Equation 5.10, is the MAP solution, with the posteriori covariance $\hat{\mathbf{S}}$ describing the "width" of the PDF.

5.3 Iterative Approach to the Non-Linear Problem

Usually, the forward model is not exactly linear, which means that either $\mathbf{y} \neq \mathbf{K}\mathbf{x}$ or the prior PDF is not Gaussian (or both). Rodgers (2000) classifies non-linear problems into four classes, the most relevant here being the *moderately non-linear* one, describing "problems where linearisation is adequate for the error analysis, but not for finding a solution". If the problem is non-linear, then the MAP solution $\hat{\mathbf{x}}$ from Equation 5.10 will not be a satisfactory inversion of the problem $\mathbf{y} = \mathbf{F}(\hat{\mathbf{x}}) + \boldsymbol{\epsilon}$.

To tackle the non-linear problem, a standard strategy is an iterative approach. Starting from a first-guess state vector \mathbf{x}_0 and the evaluation of the forward model and the Jacobian matrix at this state, an update to the state vector is calculated to yield a new \mathbf{x}_1 . Ideally, the difference between the measurement \mathbf{y} and the forward model at the new state $\mathbf{F}(\mathbf{x}_1)$ has decreased (compared to $|\mathbf{y} - \mathbf{F}(\mathbf{x}_0)|$). This process is repeated until a quantitative convergence criterion is reached, at which the state vector is accepted as the MAP solution. A suitable convergence criterion involves a *cost function*, which is sought to be minimised. For the non-linear problem, the chi-square statistic of the distribution $P(\mathbf{y}|\mathbf{x})$ provides a cost function

$$\chi^{2} = \underbrace{\left(\mathbf{y} - \mathbf{F}(\mathbf{x})\right)^{\mathrm{T}} \mathbf{S}_{\epsilon}^{-1} \left(\mathbf{y} - \mathbf{F}(\mathbf{x})\right)}_{\mathrm{Measurement}} + \underbrace{\left(\mathbf{x}_{\mathrm{a}} - \mathbf{x}\right)^{\mathrm{T}} \mathbf{S}_{\mathrm{a}}^{-1} (\mathbf{x}_{\mathrm{a}} - \mathbf{x})}_{\mathrm{Prior}}$$
(5.11)

that illustrates the Bayesian optimal estimation approach very well. The OE method will attempt to find an \mathbf{x} which minimises the difference between the forward model and the actual measurement while at the same time weighing in the information about the prior state through \mathbf{x}_a and \mathbf{S}_a . If the measurement is characterised by high levels of measurement noise, then $\hat{\mathbf{x}}$ will be closer to the prior, as the retrieval considers the prior information "more trustworthy". Similarly, for a situation where the measurement noise is small compared to the prior covariance, the measurement will play a bigger role and the retrieval will seek to minimise the difference between \mathbf{y} and $\mathbf{F}(\mathbf{x})$, even if that means that $\hat{\mathbf{x}}$ ends up being far away from the prior state \mathbf{x}_a .

UoL-FP utilises the Levenberg-Marquardt (LM) modification to the Gauss-Newton method for updating the state vector between iterations. It introduces a damping term $\gamma \mathbf{D}$, with $\gamma \in \mathcal{R}^+$. Using this damping constant, the LM method continuously scales between the Gauss-Newton method ($\gamma \rightarrow 0$) and the method of steepest descent ($\gamma \rightarrow \infty$). The state vector update from iteration *i* to iteration *i* + 1 is given by

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{S}_a^{-1} + \mathbf{K}_i^{\mathrm{T}} \mathbf{S}_{\epsilon}^{-1} \mathbf{K}_i + \gamma \mathbf{D})^{-1}$$
(5.12)

$$\times \left[\mathbf{K}_{i}^{\mathrm{T}} \mathbf{S}_{\epsilon}^{-1} \left(\mathbf{y} - \mathbf{F}(\mathbf{x}_{i}) \right) - \mathbf{S}_{a}^{-1} (\mathbf{x} - \mathbf{x}_{a}) \right].$$
(5.13)

Note that an increased γ will force the step size $|\mathbf{x}_{i+1} - \mathbf{x}_i|$ to be smaller. The Jacobian matrices \mathbf{K}_i have to be recalculated at every iteration step. For damping matrix \mathbf{D} , the choice $\mathbf{D} = \mathbf{S}_a^{-1}$ is made.

Part of the LM strategy is to update γ based on whether the state vector update results in an actual reduction of the cost function, or whether it ends up increasing it. In the UoL-FP algorithm, the adjustment to γ depends on the quantity *R*:

$$R = \frac{\chi_i^2 - \chi_{i+1,\text{FC}}^2}{\chi_i^2 - \chi_{i+1}^2},$$
(5.14)

where χ_i^2 and χ_{i+1}^2 are the χ^2 values corresponding to iterations *i* and *i* + 1, and $\chi_{i+1,\text{FC}}^2$ corresponds to the χ^2 value forecast computed through $\mathbf{K}_i \mathbf{x}_i$. Clearly, if the forward model is exactly linear, then $\chi_{i+1,\text{FC}}^2 = \chi_{i+1}^2$, and thus R = 1. For any $0 \le R \le 1$, the cost function has decreased as $\chi_{i+1,\text{FC}}^2 < \chi_{i+1}^2$, signifying that the retrieval is still exploring the forward model in a non-linear region. When R < 0, the cost function increased. For the UoL-FP algorithm, an iteration step is considered

divergent, when R < 0.25. When R > 0.75, the value of γ is updated by $\gamma_{i+1} = \gamma_i/2$, and when R < 0, $\gamma_{i+1} = 10 \gamma_i$. The initial value for γ is a matter of empirical knowledge, and there does not seem to be a consistent and simple way of picking γ_0 . The first-guess state vector \mathbf{x}_0 (iteration i = 0) does not have to be the prior mean \mathbf{x}_a , but tends to be the obvious choice.

Finally, the convergence criterion is established using the posterior covariance matrix:

$$\mathrm{d}\sigma_i^2 = \mathrm{d}\mathbf{x}_{i+1}^{\mathrm{T}}\,\hat{\mathbf{S}}^{-1}\,\mathrm{d}\mathbf{x}_{i+1}\,,\tag{5.15}$$

 $d\mathbf{x}_{i+1}$ is the state vector update $\mathbf{x}_{i+1} - \mathbf{x}_i$, where γ has been set to 0. The quantity $d\sigma_i^2$ is the change in the state vector as a fraction of the posterior covariance matrix, and once $d\sigma_i^2 < n \cdot f$, the retrieval is considered converged. n is the length of the state vector, and f is a real-valued scaling factor to adjust the convergence criterion.

5.4 Retrieval Characterisation and Errors

The Bayesian approach provides not only a means of finding a state vector to solve the original problem in Equation 5.1, but also offers a number of valuable and quantitative ways to characterise various aspects of the retrieval itself. A quantity that already appears in Equation 5.10 (for the linear case) is

$$\mathbf{G} = \mathbf{S}_{a}\mathbf{K}^{\mathrm{T}}(\mathbf{K}\mathbf{S}_{a}\mathbf{K}^{\mathrm{T}} + \mathbf{S}_{\epsilon})^{-1}, \qquad (5.16)$$

and is called *gain matrix*. It can be thought of as the derivative of the MAP solution with respect to the measurement (and roughly to the forward model), as can be seen by calculating the partial derivative using Equation 5.10 explicitly:

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}} = \mathbf{G} \approx \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{F}}.$$
(5.17)

The gain matrix reveals the contribution to each state vector element with respect to the measurement vector \mathbf{y} and can reveal whether certain spectral ranges carry information about the retrieved quantity or not, and how much.

Related to G, another quantity can be calculated as

$$\mathbf{A} = \mathbf{G}\mathbf{K} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} \left(= \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{F}} \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right), \tag{5.18}$$

which is known as the *averaging kernel matrix*. It describes the sensitivity of the retrieved state vector $\hat{\mathbf{x}}$ to the *true* state vector \mathbf{x} . This notion seems particularly confusing when talking about retrievals from real measurements. Strictly speaking, this interpretation is only valid when the true state is known. For simulation studies,

the true state is obviously known, and as such the averaging kernel provides information about the sensitivity of the retrieval. It can be used to design instruments, choose spectral ranges, and determine where in the atmosphere the sensitivity of the measurement to i.e. concentration of a certain trace gas is highest. When the true state is not known, as in most retrievals involving real measurements, the Jacobian matrix **K** in Equation 5.18 is not evaluated at the true state but rather at the final or penultimate iteration. However, as long as those states are in the vicinity of $\hat{\mathbf{x}}$, an argument can be made that the Jacobian matrix should be sufficiently similar when evaluated at the MAP solution state.

Using the pressure weighting operator **h** (O'Dell et al. 2012), the elements of the state vector that correspond to the CO_2 profile can be converted into the column-averaged XCO_2 via

$$XCO_2 = \mathbf{h}^T \hat{\mathbf{x}},\tag{5.19}$$

and the associated column averaging kernel is given by $\mathbf{a} = \mathbf{h}^{T} \mathbf{A}$. The corresponding variance is calculated from the posterior covariance matrix as

$$\sigma_{\rm XCO_2}^2 = \mathbf{h}^{\rm T} \hat{\mathbf{S}} \, \mathbf{h}. \tag{5.20}$$

As an example, Figure 5.1 shows the column averaging kernels for GOSAT and OCO-2 for ocean glint measurements from May 2016, divided into three solar zenith angle bins. As per design, the kernels peak at lower altitude levels, showing that the



Figure 5.1: GOSAT and OCO-2 column averaging kernels showing the near-surface sensitivity for both instruments. The retrievals are taken from May 2016, and the three lines correspond to solar zenith angle bins (in degrees) to which the retrievals were grouped.

TANSO-FTS and OCO-2 instruments are most sensitive to concentrations in the

troposphere. Since the values are close to 1 until the bottom of the atmosphere, the instruments are capable of providing information about near-surface concentrations.

Using the Bayesian approach, a variety of error characterisation metrics can be calculated to assist in diagnosing the result of a retrieval. One major component is the posterior covariance matrix itself, which straightforwardly represents the width of the posterior distribution. It thus provides information about whether the measurement added information compared to the prior. If for example $\mathbf{h}^T \hat{\mathbf{S}} \mathbf{h} < \mathbf{h}^T \mathbf{S}_a \mathbf{h}$, then the measurement has indeed reduced the uncertainty with which XCO₂ was known. The aforementioned gain matrix can be used to map model errors into state vector space.

In Part III (Section 11.5) of the thesis, a fast radiative transfer method is analysed regarding the XCO_2 error that the approximation error induces. If **F** is a reference forward model and **f** is another forward model with a known deficiency, such as an approximation to the radiative transfer calculations, then the error to the state vector is calculated as

$$\Delta \mathbf{x} = \mathbf{G} \left[\mathbf{F}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) \right]. \tag{5.21}$$

This equation is part of the linear analysis framework, which again formally requires linearity over the range of change induced by \mathbf{f} . Both \mathbf{F} and \mathbf{f} have to be evaluated at the true state \mathbf{x} and with auxiliary parameters \mathbf{b} (not explicitly written) being the same.

 $\hat{\mathbf{S}}$ is also extensively utilised in Part II (Section 7.6). While the diagonal elements are the familiar variances of all elements within the state vector, off-diagonal elements can carry useful information about the retrieved state as well. Correlation coefficients

$$C_{ij} = \frac{\hat{S}_{ij}}{\sqrt{\hat{S}_{ii}\hat{S}_{jj}}},\tag{5.22}$$

defined between -1 and 1 reveal whether state vector elements are correlated in the sense that the Jacobians associated with those elements are similar. If two state vector elements are highly (anti)correlated, the forward model produces very similar results, regardless of which of the two state vector elements are changed. Such correlations in retrievals can potentially give rise to systematic biases.

There are many more aspects to error characterisation within the optimal estimation approach, such as decomposing $\hat{\mathbf{S}}$ into smoothing (errors due to the shape of the column averaging kernel) and measurement error (contribution due to measurement noise). The scope of this thesis, however, was not the development of an XCO₂ retrieval, and apart from the mentioned concepts, no other error characterisation was employed.

5.5 General Retrieval Setup

In the retrievals presented in this thesis, the number of atmospheric levels (see Figure 4.2) is 20 (19 layers) with the pressure grid described in Section 4.2. The Lambertian surface model including an albedo slope is employed for simplicity, prior values are estimated from the measured continuum level radiance $\overline{I}_{\text{meas}}$, the solar zenith angle θ , and the solar irradiance (intensity) I_0 :

$$\rho_{\rm a} = \frac{\pi \overline{I}_{\rm meas}}{I_0 \cos \theta}.\tag{5.23}$$

Priors for albedo slopes are set to zero. Surface albedo is practically unconstrained, and the associated prior covariance is $\sigma_{\rho}^2 = 1$. Dispersion coefficient priors are either taken directly from the L1B data package (OCO-2) or estimated using the known location of a strong solar line at ~12 985.163 cm⁻¹ (~770.1097 nm).

The CO₂ volume mixing ratio profile at the 20 levels are part of the state vector. The corresponding prior covariance matrix $\mathbf{S}_{a}^{CO_{2}}$ is designed to reflect the variability of CO₂ in the atmosphere with a total prior uncertainty of

$$\sqrt{\mathbf{h}^{\mathrm{T}} \mathbf{S}_{\mathrm{a}}^{\mathrm{CO}_{2}} \mathbf{h}} \sim 12 \,\mathrm{ppm},$$
 (5.24)

and contains off-diagonal elements to incorporate the correlations between levels (O'Dell et al. 2012). The same $S_a^{CO_2}$ (see Figure 5.2) is used for every scene.



Figure 5.2: CO_2 prior covariance matrix S_a (left) and the corresponding correlation matrix **C** (right). The axes labels indicate the level (0 is the top of the atmosphere). For the relationship between level index and pressure, see Section 4.2.

Apart from CO_2 , the gases H_2O and CH_4 (for GOSAT) are retrieved, as several associated spectral lines appear in the GOSAT or OCO-2 windows. For the XCO_2

retrievals, the actual vertical distribution of these gases are assumed to be well-represented by models, and hence a scaling approach is used. The prior profiles are scaled by real-valued scaling factors $c_{\rm H_2O}$ and $c_{\rm CH_4}$, which are retrieved (prior values: 1).

To account for uncertainties in the meteorological data and real elevation of the satellite footprint, the surface pressure p_{surf} is retrieved. For most operational retrievals for GOSAT, the UoL-FP algorithm uses a prior uncertainty equivalent of 4 hPa. A strong deviation from the prior value generally results in a bias in XCO₂. It indicates that the true surface pressure is significantly far away from the prior value, and the retrieved value still can be, and thus the wrong reference dry air mass will lead to a biased XCO₂ value.

In the state vector, aerosols are defined by their logarithmic extinction profiles per pressure difference. The logarithmic transform is a common method to deal with quantities that are physically bounded from below by zero. Performing the aerosol retrieval in log-space therefore ensures that the back-transformed value never turns negative. This choice comes with a number of drawbacks. A very small prior will almost never move away from that initial value, unless a very large prior covariance is employed. The current UoL-FP aerosol scheme calculates a prior covariance matrix for each scene independently, where covariances are larger for low-AOD scenes, and more restrictive for scenes with a high aerosol load. As can be seen on an example of the aerosol prior covariance matrix for the small-type aerosol (Figure 5.3), the profile is allowed to vary only in the troposphere, which is where the MACC-II/CAMS models inform the aerosol scheme. The right panel of Figure 5.3 shows the high correlations between the various layers.

It is important to realise that the GOSAT or OCO-2 spectra do not provide 20 individual pieces of information about every aerosol mixture. The actual *degrees of freedom* for the aerosols in the retrieval is generally between ~0.5 and ~2.0 for the small and large type aerosol mixtures, and between ~0.5 and ~1.0 for the cirrus type mixture.

Fluorescence radiances are accounted for in land scenes, where priors come from the fluorescence retrievals from Part II. Early tests have shown that due to the large single-sounding uncertainty of the fluorescence retrieval, the XCO₂ retrievals exhibited some additional scatter related to the fluorescence priors. Therefore, a smoothing operation is performed on the fluorescence values in both time and space, before the actual value is used in the XCO₂ retrievals. The smoothing procedure is based on a simple Gaussian weighting: all fluorescence measurements within d < D = 250 km and t < T = 30 day are collected. Spatial and temporal weights are then calculated



Figure 5.3: Example of a prior covariance matrix for the small type aerosol mixture (left). The values for stratospheric levels are very small, and will now allow the retrieval to adjust the aerosol in those levels. Correlations (right) show that the aerosol extinction at a given level is coupled to the levels below and above.

as

$$w_d(d) = \exp\left(-\frac{d^2}{2(D/4)^2}\right),$$

$$w_t(t) = \exp\left(-\frac{t^2}{2(T/4)^2}\right),$$
(5.25)

and multiplied to yield $w(d, t) = w_d(d) \cdot w_t(t)$, which is then used to calculate the weighted mean

$$F_{\rm s}^{\rm avg} = \frac{\sum_{i} w(d_{i}, t_{i}) F_{\rm s}(d_{i}, t_{i})}{\sum_{i} w(d_{i}, t_{i})}.$$
(5.26)

Instrument noise for GOSAT retrievals is estimated through the out-of-band portion of the spectra. These portions of the spectra are accessible through the L1B files, however have to be manually converted from V cm to $W \text{ cm}^{-2} \text{ sr}^{-1} \text{ cm}^{-1}$. Noise in FTS-type devices is typically white noise, so the noise level inferred at the out-of-band regions can be assumed to be valid for the entire band.

Part II



Remote Sensing of Chlorophyll Fluorescence

Overview

Joiner et al. (2011) were first to retrieve fluorescence from GOSAT measurements on a global scale. Using an additional spectral window, but still employing the same basic retrieval principle, Frankenberg, Fisher, et al. (2011) showed that a simple bias correction can rid the retrieved fluorescence of unphysical values over non-vegetated areas. Soon, global fluorescence retrievals were performed using measurements from SCIAMACHY (Joiner et al. 2012) and the moderate-resolution spectrometer instrument GOME-2 (Joiner et al. 2013). The newest additions of space-based fluorescence measurements are the 2014-launched OCO-2 (Sun et al. 2017), as well as the recently launched Sentinel 5-Precursor whose TROPOMI instrument is capable of measuring fluorescence as well (Guanter et al. 2015). These various instruments capture chlorophyll fluorescence on different scales (footprints), from $80 \times 40 \text{ km}^2$ (GOSAT) down to $1.29 \times 2.25 \text{ km}^2$ (OCO-2) and repetition times.

J.-E. Lee et al. (2013) analysed GOSAT-derived fluorescence data over the challenging Amazon region and found that it better predicts water stress as quantified through vapour pressure deficit (VPD), than the enhanced vegetation index (EVI). A very comprehensive study using fluorescence data retrieved from GOME-2 measurements was carried out in Joiner et al. (2014). They assessed the seasonal cycles of fluorescence sampled at flux tower sites across the globe and concluded that fluorescence tracks the spring onset of photosynthesis well. Seeing as the absorbed photosynthetically active radiation tends to stay high even after the growing season, fluorescence provides better information about photosynthetic activity and is better suited to constrain carbon uptake in models. The onset of photosynthetic activity in high-latitude regions was studied by Luus et al. (2017). Similarly, they find that fluorescence is a better predictor of spring onset than EVI, and thus highlight the advantage of fluorescence over greenness indices to improve the modelling of the seasonal cycle in the tundra. In a more global approach, Parazoo et al. (2014) have used fluorescence in combination with an ensemble of dynamical global vegetation models (DGVMs) to constrain carbon uptake. Their results suggest that some models are over-sensitive to drivers of the growing season in the extra-tropics, while at the same time underestimating the seasonal amplitude in the Amazon region. Regardless of the value of fluorescence for the carbon science community, they also directly affect

the quality of XCO_2 retrievals, as shown by Frankenberg et al. (2012).

Fluorescence measurements from space will enter a new era in few years' time, due to the Fluorescence Explorer (FLEX) having been selected for a future Earth Explorer mission by the European Space Agency - a detailed overview of the mission is published in Moreno et al. (2015). FLEX will be solely dedicated to measuring chlorophyll fluorescence using a push-broom-type imaging spectrometer (FLORIS) to achieve the research objectives, which include the better understanding of the dynamics of plant photosynthesis and its interactions in dynamic vegetation models. The FLORIS instrument will measure both at low (2nm) and very high (0.1nm) resolutions - low resolution is sufficient to capture the photochemical reflectance index (PRI) as well as chlorophyll absorption, high resolution is needed to capture both the O₂-A and the O₂-B bands in which the fluorescence signal lies. In terms of spatial sampling, the FLEX mission will significantly reduce the footprint size compared to current observations, and will deliver soundings at $300 \text{ m} \times 300 \text{ m}$ resolution. The small footprint size is needed to achieve the science objective related to the identification of optimal management strategies for crop production. FLEX will be flying in tandem with Sentinel-3 (Donlon et al. 2012) and use its assessment of the atmospheric conditions to provide further information for the fluorescence retrieval. The measurement swath is stated as 150 km for the intended sun-synchronous orbit at 815 km, which then results in a revisit time of up to 19 days. FLEX is scheduled to be launched in 2022.

In this part of the thesis I first introduce chlorophyll fluorescence in the context of photosynthesis and explain how it arises on a microscopic level. I then lay out the concept behind retrieving fluorescence from high-resolution spectroscopic measurements and assess the result of processing a 7-year GOSAT record. Using this comprehensive fluorescence record, I study a large continental-scale event, in order to demonstrate the capabilities of fluorescence measurements to capture the vegetation's response to disruptions of the biosphere. Finally, I investigate the impact of introducing well-calibrated a priori fluorescence radiances to full-physics XCO₂ retrievals. Chapter Six

Solar-Induced Chlorophyll Fluorescence

OLAR-Induced chlorophyll fluorescence (SIF) is a byproduct of a chain of reactions that occur simultaneously alongside photosynthesis. To understand how SIF arises in the first place, a detailed description of photosynthesis is required. This chapter is heavily derived from the textbooks Taiz et al. (2010), and Cooper et al. (2013).

6.1 Chloroplasts, Chlorophyll, and Photosystems

Photosynthesis occurs entirely within plant and algal cells, more precisely in organelles (sub-structures of cells, dedicated to a specific task) called *chloroplasts*. Chloroplasts are found in a variety of plant cells, mostly in parts of the plant that appear green, such as leaves or the near the surface of the plant stem. They are between $5 \,\mu\text{m}$ and $10 \,\mu\text{m}$ long, and the number of chloroplasts per cell can vary with plant type as well as with the surrounding environmental conditions.

Like plant cells, chloroplasts themselves have an inner structure, illustrated in Figure 6.1. Surrounded by a double membrane, the objects of interest are the so-called *thylakoids*. Inside the double membrane, the actual thylakoids are found as interconnected stacks of flat discs (*grana lamellae*) surrounded by a colourless fluid (*stroma*). Thylakoids also feature a sub-structure, as they consist of the *thylakoid lumen* enveloped in the *thylakoid membrane*; the stacks are connected to each other through the same kind of membrane (however unstacked) called *stroma lamellae*.

The photochemical reactions of photosynthesis all take place in the thylakoid membrane and the stroma lamellae. Embedded in the thylakoid membrane are two types of large photochemical protein complexes, the so-called *photosystems* (I and II, abbreviated as PSI and PSII). For the understanding of SIF, the relevant elements of the photosystems are the antenna complexes, an array of chlorophyll molecules (and other carotenoid pigments) in which incoming light energy is absorbed, and the respective reaction centres. Both reaction centres of PSI and PSII are chlorophyll dimers, called P700 and P680, respectively. The numbers indicate the wavelength of maximal spectral absorption in nano metres.



Figure 6.1: Illustration of the inner structure of a chloroplast. The disk-shaped thylakoids are arranged in stacks (grana lamellae), interconnected via the stroma lamellae.



Figure 6.2: Chlorophylls a $(C_{55}H_{72}MgN_4O_5)$ and b $(C_{55}H_{70}MgN_4O_6)$. The two chloropylls are identical apart from the functional group **R**, which for chlorophyll a is another methyl (CH₃), and for chlorophyll b is an aldehyde (CHO).

The photosystems are physically separated from each other, PSI is found in the stroma lamellae as well as the edges of the grana lamellae, whereas PSII mostly resides in the grana lamellae. Both, however, are located in the thylakoid membrane. The roles of each mentioned component within the photosynthetic apparatus are explained in the following section.

6.2 Light-Dependent and Light-Independent Reactions

The process of photosynthesis is driven by two sets of reactions, one of them relying on incident light (*light-dependent reactions* or *thylakoid reactions*), the other one being independent on the presence of light energy (*light-independent reactions* or *carbon fixation reactions*). Since electron transport is a fundamental aspect of the lightdependent reactions, the preferred way to depict them is an energy diagram called Z scheme (Figure 6.3). The name derives from the fact that the redox diagram resembles a Z.



Figure 6.3: This illustration depicts a heavily simplified version of the Z scheme.

Following the Z scheme from left to right, the light-dependent reactions can be summarised in a compact manner.

Incident light is absorbed by the chlorophylls of the antenna-complex of a PSII, which leaves those chlorophylls excited for a short time. When an incident photon is absorbed by an electron, this photon-absorbing electron is raised onto a higher energy level and then subsequently returned to its former state. The excitation energy (the difference between the excited and former state) is then transferred from one chlorophyll molecule to a neighbouring one without any chemical reactions via *Förster* (or *fluorescence*) resonance energy transfer (FRET). This type of energy transfer occurs through dipole-dipole interactions, without any radiation being emitted, and is only effective over short distances ($\propto R^{-6}$).

Passed on between neighbouring chlorophylls, the excitation energy will eventually reach the reaction centre of the PSII, P680, which is subsequently excited (P680^{*}). Rather than passing on the excitation energy, the excited electron within P680^{*} is itself now ejected from the reaction centre. The oxidised (electron-deficient) P680^{*} regains its missing electron through the *oxygen-evolving complex* (or *water-splitting*)

complex), where water in the thylakoid lumen acts as electron donor:

$$2 H_2 O \rightarrow O_2 + 4 H^+ + 4 e^-.$$
 (6.1)

As seen in the equation above, a total of four photons (one per released hydrogen atom) is needed to oxidise two water molecules and provide four electrons to the electron-deficient reaction centres of PSII. This specific reaction seems to be unique amongst the currently known biochemical systems - practically all molecular oxygen in the Earth's atmosphere is a result of this particular reaction.

The ejected electron from the excited PSII reaction centre is carried through an electron transport chain (ETC) and loses energy at each reaction within the ETC (the details of the ETC are skipped here). At the end of the ETC, the electron arrives at an excited reaction centre of PSI (P700^{*}), which was excited through the same mechanism as P680. That is through transferred excitation energy coming from an incident photon absorbed by a chlorophyll molecule in the surrounding antenna complex. The electron that initially came from PSII is now used to replace the missing one from the excited P700^{*}, which passed its excited electron on to another ETC. As mentioned earlier, both P680 and P700 are named due to their absorption maximum being around 680 nm and 700 nm respectively. Chlorophylls in the antenna complexes exhibit a peak absorption at lower wavelengths, closer to ~450 nm.

The first ETC provides energy for the synthesis of *adenosine triphosphate* from *adenosine diphosphate* (ADP \rightarrow ATP), which in very simple terms can be described as an energy storage molecule that can carry chemical energy to be used for various metabolic functions within plant cells. At the end of the second ETC, the energised electron (originating from P700) is used in the reduction of a chemical compound called *nicotinamide adenine dinucleotide phosphate* (2 NADP⁺ + 2 H⁺ \rightarrow 2 NADPH).

Concluding the light reactions of photosynthesis, the relevant steps are as follows:

- Photons are absorbed by chlorophylls in the antenna complexes of the two photosystems, and the gained excitation energy is used to oxidise the photosystem reaction centres P680 and P700
- The electron deficient P680 receives an electron through the splitting of water, whereas the oxidised P700 obtains the electron ejected by P680
- Ejected electrons from both reaction centres pass through ETCs and are used to create ATP and reduce NADP⁺ to NADPH

At this point, the designation of these reactions make sense, as they cannot take place without photons that excite electrons in the reaction centres. The light-independent reactions follow up where the light reactions finished, they take place in the stroma. Known as the Calvin–Benson–Bassham (CBB) cycle, this set of reactions, in which both ATP and NADPH are utilised, ultimately leads to the fixation of atmospheric CO_2 into triose phosphate (G3P). The total (net) reaction of the CBB cycle is:

 $3 \operatorname{CO}_2 + 5 \operatorname{H}_2 \operatorname{O} + 6 \operatorname{NADPH} + 9 \operatorname{ATP} \rightarrow \operatorname{G3P} + 6 \operatorname{NADP}^+ + 3 \operatorname{H}^+ + 9 \operatorname{ADP} + 8 \operatorname{P}_i,$ (6.2)

with P_i being inorganic phosphates. The CBB cycle itself does not result in the formation of sugars, however two G3P molecules can form one glucose molecule ($C_6H_{12}O_6$). The particular details of the CBB cycle are far outside the scope of this thesis and do not directly relate to SIF, and are hence omitted here.

6.3 Non-Photochemical Quenching

In Section 6.1, photosystems were introduced as one of the fundamental building blocks of the photosynthetic apparatus. The individual carotenoid pigments of the surrounding antenna complexes only hold on to the incoming light energy for a very short time, on the order of picoseconds ($1 \text{ ps} = 1 \cdot 10^{-12} \text{ s}$). This quick dissipation of excitation energy is referred to as *quenching* and plays a vital role in the health of plants.

Since there is only a finite amount of photosystem reaction centres, not all excited pigments can pass on their excitation energy immediately. Once a reaction centre receives excitation energy and is in the process of passing it on to a mobile electron, it is *closed* and not able to receive any further excitation energy until it is *open* again. From receiving excitation energy to open up again, up to a millisecond can pass (Krause et al. 1991). Thus, it takes orders of magnitude longer for the reaction centres to open up again than the usual duration of a FRET process.

If a pigment is not able to pass on excitation energy, the excited state will be prolonged, especially in an environment of high light intensity. The longer a pigment stays in that singlet-excited state, the chances of a long-lived triplet excitation (${}^{3}Chl^{*}$) are increased, which in turn can cause the formation of the highly reactive and harmful singlet oxygen (${}^{1}O_{2}^{*}$) (Müller et al. 2001).

One mechanism to counter the formation of harmful singlet oxygen is the so-called *non-photochemical quenching* (NPQ), in which excess light energy is dissipated as heat. Again, NPQ itself is a complex topic and the details are skipped here.

6.4 Chlorophyll Fluorescence

Besides NPQ and photosynthesis, there is a third de-excitation pathway for pigments to dissipate their photon-induced excitation energy: chlorophyll fluorescence. Here, a small amount of the energy is turned into heat, whilst the majority is used to emit photons in the far-red range of the wavelength spectrum (roughly 650 nm to 800 nm). The spectral characteristic of the fluorescence signal should not be confused with the absorption profile of the chlorophylls in the antenna complexes.

Generally, the efficiency of photochemistry is above 90%, meaning that > 90% of the incident light energy is used for photosynthesis. The efficiency of SIF ranges between 0.6% and 3.0%, the remaining 7.0% to 9.4% is turned into heat via NPQ. As the level of illumination increases, the efficiency of the photochemistry pathway decreases as both NPQ and SIF ramp up accordingly.

Spectrally, the SIF signal (see an example in Figure 6.4) consists of two distinct peaks at around 685 nm and 750 nm. PSII contributes to both peaks, whereas PSI mostly contributes to the peak at 750 nm.



Figure 6.4: This figure illustrates a typical broadband SIF emission spectrum. The shaded area between 755 nm and 772 nm marks the location of the O_2 A-band as measured by GOSAT/TANSO-FTS and other instruments. Note that the SIF magnitude at 755 nm is about 1.8 times higher compared to 772 nm. The emission spectrum was calculated using FluorMODgui (Zarco-Tejada et al. 2006).

The SIF yield does not only change with varying illumination scenario, it also directly reflects other environmental parameters that affect the plant physiology. Willits et al. (2001) have shown that SIF yield (as a fraction of the maximally obtainable fluorescence signal) decreased significantly with leaf temperature, both in a laboratory and a greenhouse environment. Another example is the study performed by Souza et al. (2004), in which the SIF signal from cowpea plants (*Vigna unguiculata*) was studied while the plants were experiencing water stress. The plants were withheld water until the measured CO_2 assimilation rate vanished, and were then watered again. Several experiments within that study show similar results: water-stressed plants exhibit lower photochemical quenching (qP, quenching due to photochemistry), higher NPQ, and lower SIF yield.

The intricate relationship between SIF and actual CO_2 uptake has been recently investigated by Magney et al. (2017). They employed a new leaf-level experimental setup, which allows for the simultaneous measurement of SIF radiances and gas exchange. The fluorescence signal was measured via a *pulse-amplitude modulation* (PAM) fluorometer and a spectrometer, the gas exchange was measured using an infra-red gas analyser (IRGA). In their Figure 4 (bottom row) they show again how stressed vegetation emits lower-magnitude SIF radiances, while also showing that the SIF signal increases with increasing incident light. Finally, they show relationships between net CO₂ assimilation (as derived from IRGA measurements) and steadystate SIF. While roughly speaking, the SIF radiances increase with CO_2 assimilation rates (A_{net}) for non-stressed samples, the relationship is non-linear. Looking at the stressed (increased ambient temperatures, reduced humidity) samples, $A_{\rm net}$ saturates for both investigated plant species, however the magnitude of the SIF radiances increases almost to unstressed levels. From the results shown in that figure, it does not seem easily possible to distinguish a stressed from a unstressed sample based on the leaf-level SIF radiances alone, using only one measurement.

6.5 SIF and Primary Production

SIF and primary production are linked on a microscopic scale, as explained in Section 6.2. Describing the link between SIF as a leaf-level phenomenon, and SIF as recorded as an average over several square kilometres of vegetation, is not an easy task. Porcar-Castell et al. (2014) have published an overview on the topic and highlighted some of the current challenges, most of them pertain to the seasonality of SIF and GPP - both a decoupling of SIF and GPP is needed on that time scale, as are leaf-level measurements of SIF via PAM. The macroscopic relationship between primary production and SIF is fairly straightforward. Going back to Monteith (1972), primary production can be written using what is now called the *light use efficiency* (LUE) model:

$$GPP = PAR \cdot fPAR \cdot LUE_{p}. \tag{6.3}$$

GPP is simply a product of three terms. PAR is the *photosynthetically active radiation* and describes the amount of down-welling radiation in the wavelength band between 400 to 700 nm. This particular band is the main spectral window in which Chlorophylls absorb light (see Section 6.2), thus PAR represents the radiation which can be harvested for photosynthesis. The second term, $fPAR^1$, is the *fraction* of PAR that is actually absorbed by a plant, so that the product of the first and second term is the absolute radiation that is available for the plant for use in photosynthesis. The last

¹Sometimes also written as fAPAR.

term is the light use efficiency itself (LUE_p) , which again is a fractional value and describes how much of the available radiation is actually used for photosynthesis.

PAR is relatively easily calculated and depends mostly on the location and the time of year, as well as cloud cover. *f*PAR is usually derived from measured surface reflectance, taking into account specific biome types. LUE_p is a more intricate parameter, as it depends highly on the season, the plant type and the overall current physiological condition of the plant, for example whether it is stressed or not. Plant conditions are of course highly dependent on many environmental factors. The LUE model is commonly used, such as in the MODIS GPP product (MOD17, Zhao et al. 2005).

Analogous to Equation 6.3, a similar relationship can be proposed for SIF radiances (Guanter et al. 2014):

$$SIF = PAR \cdot fPAR \cdot LUE_f \cdot f_{esc}. \tag{6.4}$$

Here, the first two terms are equal to those in Equation 6.3, however the light use efficiency term is different. LUE_f , is the fraction of absorbed radiation that is then re-emitted as SIF. The last term $f_{\rm esc}$ is the fraction of emitted SIF radiation that makes it through the canopy (if there is one) to reach the sensor on a space-based platform.

Ignoring f_{esc} , GPP and SIF, as modelled through Equation 6.3 and Equation 6.4, can be related easily as

$$GPP \sim SIF \cdot \frac{LUE_{p}}{LUE_{f}},\tag{6.5}$$

which states that SIF and GPP are linearly related. If the two light use efficiency terms are varying in the same manner throughout the season for a specific plant type, then the relationship between SIF and GPP is simply given by a scale factor, as LUE_p/LUE_f can be assumed to be near constant. Guanter et al. (2014) point out that the co-variation between LUE_p and LUE_f is shown through experimental evidence in e.g. Zarco-Tejada et al. (2013).

Even though the linearity between GPP and SIF has been observed through spacebased measurements, the relationship is still poorly understood in the context of other influential meteorological variables. In Section 9.1, the GOSAT-retrieved SIF is compared to GPP models and measurements. Chapter Seven

SIF Retrieval from GOSAT

7.1 Fraunhofer Line Discrimination Principle

RACTICALLY all retrieval concepts employed for measurements from SCIA-MACHY, GOME-2, GOSAT and OCO-2 are derived from the *Fraunhofer line discrimination* (FLD) method, described first in Plascyk (1975), and Plascyk et al. (1975). The FLD principle works on the basis that solar (Fraunhofer) lines will be altered when an additional radiance signal from the surface is added

to the TOA radiance. This can be illustrated in a very simple example: consider a single solar absorption line, centred around a wavelength λ_{in} (Figure 7.1). Assume



Figure 7.1: Illustration of the FLD principle, inspired by Alonso et al. (2008). The solar line, centred at λ_{in} , is filled in via the presence of the additive fluorescence radiance F_s .

that the solar irradiance at the top of the atmosphere I is known for that particular wavelength, as well as for another wavelength in the continuum, λ_{out} . Corresponding to the known solar irradiances I_{in} and I_{out} , the radiances, as measured by a satellite, L_{in} and L_{out} are also known. Irradiance and measured radiance can be related through

$$L_{\rm out} = I_{\rm out} \cdot R + f \cdot I_{\rm out},\tag{7.1}$$

$$L_{\rm in} = I_{\rm in} \cdot R + f \cdot I_{\rm out},\tag{7.2}$$

where R is a reflectance term (i.e. related to Lambertian albedo), and f is the additive fluorescence radiance F_s as a fraction of the continuum level solar irradiance

 $f = F_{\rm s}/I_{\rm out}$.

These two equations can be solved for R and f,

$$R = \frac{L_{\rm out} - L_{\rm in}}{I_{\rm out} - I_{\rm in}},\tag{7.3}$$

$$F_{\rm s} = f \cdot I_{\rm out} = \frac{L_{\rm in}I_{\rm out} - L_{\rm out}I_{\rm in}}{I_{\rm out} - I_{\rm in}}.$$
(7.4)

This demonstrates in a very simple way that reflectance can be decoupled from the fluorescence signal if the solar irradiance is known and if two measurements, inside and out of the solar line, are available. The characteristic change of the solar line in the radiance signal (compared to the solar irradiance) is known as the *in-filling* of the solar line. While the absolute depth of the solar line $L_{out} - L_{in} = R \cdot (I_{out} - I_{in})$ only changes through the reflectance term R, the *fractional line depth* (relative to the continuum level radiance) of the solar line

$$\frac{L_{\text{out}} - L_{\text{in}}}{L_{\text{out}}} = \frac{I_{\text{out}} - I_{\text{in}}}{I_{\text{out}}} \cdot \frac{R}{R + f}$$
(7.5)

is a function of f.

7.2 Physically-based Retrieval

The *physically-based* SIF retrieval method, first discussed in Frankenberg, Butz, et al. (2011) and employed in Frankenberg, Fisher, et al. (2011), is an implementation of the FLD method discussed in Section 7.1. In this thesis, this approach was used to produce the SIF data analysed in the later sections.

TANSO-FTS on board GOSAT provides high-resolution spectroscopic measurements of the O_2 A-band (see Section 3.2). Near the two edges of the band, several solar lines are situated, as shown in Figure 7.2. Frankenberg, Fisher, et al. (2011) utilise two separate spectral micro-windows, roughly at 755 nm and 772 nm and retrieve the fluorescence signal independently.

The 755 nm window contains two strong solar lines, as well as half a solar line at the very edge of the band. In previous versions (prior to V201) of the GOSAT L1B data, the radiances extended to lower wavelengths, and thus included the entire solar line at 757.61 nm. Throughout this thesis, the GOSAT L1B V201 and V202 were used, which introduced a new data field radiance_BestEstimated. These radiances come already calibrated and corrected for instrument degradation, requiring no further treatment and is input into the retrieval algorithm.

The 772 nm window contains several solar lines, as well as two weak O_2 doublettes (red arrows in Figure 7.2, see Section 2.1, Page 13 for details). While there are more



Figure 7.2: A modelled O_2 A-band spectrum with shaded regions marking the two micro-windows at 755 nm and 772 nm that are utilised in the retrieval scheme. The two red arrows mark the positions of two weak O_2 doublettes in the 772 nm micro-window.



Figure 7.3: Close-up of the two micro-windows used in the physically-based retrieval scheme.

solar lines in this window than in the 755 nm window, the SIF signal is only about half as strong here. Figure 7.3 shows the two micro-windows close-up. Apart from the weak O_2 lines in the 772 nm micro-window, there are no absorbers in these two spectral windows that can be resolved using GOSAT or OCO-2 measurements. This means that SIF can be retrieved from these two windows without having to constrain or retrieve additional gases, as no (or negligible) atmospheric absorption takes place in those two small wavelength ranges - apart from the mentioned oxygen lines.

SIF can be represented in a straightforward fashion as the absolute radiance value, or as a fraction of the continuum level radiance. The continuum level radiance is very close to the maximum radiance value in the O₂ A-band, $I_{\text{max}}^{(1)}$, so the relative fluorescence can be calculated as

$$F_{\rm s}^{\rm (rel)} = \frac{F_{\rm s}}{I_{\rm max}^{(1)} - F_{\rm s}},\tag{7.6}$$

since $I_{\text{max}}^{(1)}$ already contains the fluorescence radiance.

7.3 Retrieval Set-Up

The state vectors for the two independent retrievals are as follows:

$$\mathbf{x}_{755} = (c_0, c_1, d_0, d_1, z_0), \tag{7.7}$$

$$\mathbf{x}_{772} = (c_0, c_1, d_0, d_1, p_{\text{surf}}, z_0).$$
(7.8)

For both windows, two continuum scaling parameters c_0 and c_1 are retrieved. c_0 is a scaling factor with which the convolved spectrum is multiplied by, and c_1 is the corresponding gradient. These two parameters allow for an adjustment of surface reflectance if the first guess albedo of the assumed Lambertian surface does not match the measurement. The two dispersion parameters d_0 (dispersion shift) and d_1 (dispersion stretch) similarly are retrieved in case the first guess of the instrument dispersion does not align the absorption lines fully. The zero-level offset z_0 is essentially the SIF signal, which is assumed to be constant across the micro-window. Due to the presence of oxygen absorption lines in the 772 nm, \mathbf{x}_{772} also contains surface pressure in order to match those lines to the measurement.

Rather than using a DOAS set-up to retrieve SIF from GOSAT, the UoL-FP algorithm was utilised. While the computational effort is probably orders of magnitude higher, the already available meteorological data, a working solar model, as well as compatible inputs to the algorithm made it the obvious choice.

In order to speed up the retrieval, a new RT model was added to the UoL-FP algorithm. Prior to this addition, the only radiative transfer option was the LIDORT model - a general purpose multiple-scattering code; however SIF retrievals do not need a scattering RT model and would not benefit from a sophisticated MS code. The new addition allows running the algorithm while skipping any calls to the sophisticated RT codes, and use a simple non-scattering Beer-Lambert approach instead. Given the layer-resolved ($i = 1 \dots N_{lay}$) total optical depths τ_i and the Lambertian surface albedo ρ , the scalar TOA intensity can be written as:

$$L_{\text{TOA}} = \underbrace{I_0 \cdot \exp\left[-\frac{\sum_{i}^{N_{\text{lay}}} \tau_i}{\mu_0}\right]}_{\text{incoming solar beam}} \cdot \underbrace{\frac{\rho \cdot \mu_0}{\pi}}_{\text{surface reflectance}} \cdot \underbrace{\exp\left[-\frac{\sum_{i}^{N_{\text{lay}}} \tau_i}{\mu}\right]}_{\text{reflected beam}}, \quad (7.9)$$

where μ_0 and μ are the cosines of the solar zenith and viewing zenith angles, respectively, and I_0 is the incident solar irradiance. As no aerosols are considered in this retrieval, and only oxygen contributes to molecular absorption, the optical depths τ_i are a result of the extinction due to O₂ absorption and Rayleigh scattering.

Since the surface pressure has to be retrieved for the larger-wavelength spectral window, the layer-wise weighting functions (see Section 4.6, Page 44) for both optical depth and single-scatter albedo have to be supplied to the algorithm for the computation of the surface pressure Jacobians:

$$\frac{\partial L_{\text{TOA}}}{\partial \tau_i} = -L_{\text{TOA}} \cdot \left(\mu^{-1} + \mu_0^{-1}\right),\tag{7.10}$$

$$\frac{\partial L_{\text{TOA}}}{\partial \varpi_i} = 0. \tag{7.11}$$

The first guess for the state vectors are obtained exactly as for the UoL-FP CO₂ retrievals: surface pressure is derived from ECMWF ERA-Interim and the SRTM digital elevation map, the instrument dispersion is estimated through fitting of a single, known solar line near 770.11 nm. Surface albedo is estimated though the measured signal itself, and the first guess zero-level offset is $1 \cdot 10^{-8}$ W cm⁻² sr⁻¹ cm⁻¹.

Regarding the prior covariances of the retrieval, they are set equally for every sounding. z_0 is very loosely constrained with $\sigma = 1 \cdot 10^{-4} \text{ W cm}^{-2} \text{ sr}^{-1} \text{ cm}^{-1}$, as is surface pressure for \mathbf{x}_{772} with $\sigma = 20$ hPa. Both S- and P-polarised spectra, as well as the averaged, unpolarised PS spectra were processed independently (see Section 3.2, Page 23).

7.4 Collocation of Auxiliary Data

Since SIF is a surface parameter rather than an atmospheric one, it is convenient to have additional surface-related data available to study the retrieval results. To this end, several data sets have been sampled at the GOSAT measurement locations:

- 1. ESA CCI Land Cover v2.0.7 (global, 300 m resolution, 2009 2015)
- 2. ESA CCI Soil Moisture v3.2 (global, 0.25° resolution, 2009 2015)
- MODIS/Aqua MYD13C1 V006 Vegetation Indices (global, 0.05° resolution, 2009 – 2016)
- 4. GOME-2 SIF (global, $80 \text{ km} \times 40 \text{ km}$ footprint, $2009 2016)^1$
- 5. TRMM (3B43, multi-satellite) precipitation (global between -50° and 50° latitude, 0.25° resolution, 2009 2016)

For all products, apart from the ESA CCI soil moisture data, the spatial extent of the GOSAT footprint was fully taken into account. The elliptical footprint is given by the longitude and latitude of the vertices and co-vertices. From these points, the ellipse is fully defined and a polygon representing the instantaneous field of view (IFOV) is constructed.

Figure 7.4 illustrates the idea of the collocation of the elliptical GOSAT footprint with higher-resolution data, such as the ESA-CCI Land Cover product. First, both the GOSAT footprint polygon as well as a subset of the LC data are transformed into a local map coordinate system. The chosen coordinate system is the Lambert conformal projection. Conformal projections are advantageous here, as the shape of the ellipse is not distorted at high latitudes. Within that new local map coordinate system, the pixels that lie fully within the elliptical footprint are collected, and the fractional contribution of each land cover class, relative to the total number of pixels considered, is stored.



Figure 7.4: Example of the collocation/aggregation of land cover data. This (almost circular) GOSAT footprint, located in Manitoba (Canada), contains about 65% needleleaf evergreen tree cover, about 17% shrubland and about 17% water body LC classes (total of ~1800 pixels). Hatched pixels are not collocated, as the areal overlap is < 100%.

MYD13C1 data are collocated in the same manner. However, due to the coarser spatial resolution of the MYD13C1 data(0.05° climate model grid), the restriction

¹Level 2 data produced by GFZ Potsdam / Philipp Köhler

of fully enveloped pixels has to be loosened. Otherwise, the number of matches would be greatly reduced. In the collocation calculation, the relative areal overlap between the GOSAT footprint polygon and every rectangular pixel of the target data is calculated as a percentage of the GOSAT footprint area (A_f) . If the pixel is fully within the GOSAT footprint, that overlap is 100%, when located entirely outside, it is 0%, and so forth. For MYD13C1 collocations, the relative overlap is required to be >50% (see Figure 7.5a).

Collocation of the ESA CCI soil moisture data is handled slightly simpler. Located on a 0.25° grid, the data is bi-linearly interpolated at the location of the GOSAT footprint centre, and the spatial extent of the footprint is ignored (not pictured).

To collocate GOME-2 Level 2 SIF data, there is no sub-pixel treatment necessary, as the GOME-2 footprint is much larger than the GOSAT footprint. All GOME-2 SIF retrievals, for which the centre of a GOSAT footprint lies within the GOME-2 footprint, are collocated, and the relative areal overlap is also stored (see Figure 7.5b).



(a) MYD13C1 / EVI variable $(0.05^{\circ}), A_f > 50\%$. (Same for MYD11C2).



(b) GOME-2 ($80 \text{ km} \times 40 \text{ km}$) footprint (yellow rectangle). The base map was taken from the Esri ArcGIS REST Services using the World_Imagery layer[†].

Figure 7.5: Collocation example for the various target data sets at different resolutions to illustrate how the spatial extent of the GOSAT footprint was taken into account. Coordinate labels are suppressed here since the location of the footprint is the same as in Figure 7.4. Hatched pixels are not aggregated. For (b), the projection was switched to a cylindrical equal area map projection. [†]Sources: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community.

Additionally, every GOSAT measurement is also assigned to one of the eleven TransCom regions over land (see Figure 7.6):

1. North American Boreal



Figure 7.6: TransCom basis function map ("The Transcom Experiment" 2018), which divides the landmasses (excluding Greenland and Antarctica) into 11 separate regions. Ocean regions have been ignored in this visualisation.

- 2. North American Temperate
- 3. South American Tropical
- 4. South American Temperate
- 5. Northern Africa
- 6. Southern Africa
- 7. Eurasian Boreal
- 8. Eurasian Temperate
- 9. Tropical Asia
- 10. Australia (and New Zealand)
- 11. Europe

Unless stated otherwise, all maps and other visualisations using these auxiliary data were created using the above methods and are thus sampled at GOSAT locations. While this may produce different results than taking the source data directly, this makes sure that potential sampling biases are kept to a minimum and do not interfere with the interpretation of the results.

7.5 Pre-selection and Initial Filtering of the Retrieval Results

While Frankenberg, Fisher, et al. (2011) use the retrieval of the 772 nm window to screen for cloudy scenes, here the pre-processing results of the UoL-FP pipeline were used to select scenes with low cloud contamination. All scenes for which the apparent, retrieved surface pressure deviated from the ECMWF-informed prior by more than 100 hPa were disregarded.

For the analysis of the results, all retrievals (per sounding) were collected for which the algorithm successfully² finished for both micro-windows, and each of the three polarisation states (P, S, PS) was treated independently.

Generally, retrievals for the 755 nm window show higher success rate than those for 772 nm. Depending on the month, sometimes up to 5% of retrievals of the 772 nm micro-window failed due to the surface pressure exceeding the lowest level of the model atmosphere.

The fit quality is assessed via the normalised χ^2 statistic, which takes into account instrument noise, as exemplified in Figure 7.7. The distributions of the χ^2 are centred



Figure 7.7: An example of a successful fit and corresponding residual. Radiances and residuals are in units of $1 \cdot 10^{-7} \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$.

around 1, however there is a large variability associated with each distribution ($\sigma \sim 0.5$) - which is likely due to a lack of re-scaling the instrument noise levels as described in Crisp et al. (2012). The retrievals using the S-polarised spectra show systematically lower χ^2 , when compared to the P- and PS-polarised retrievals (Figure 7.8). The difference is believed to be of instrumental nature, but the causes were not identified closer here.

 $^{^{2}}$ A retrieval is considered successful (not necessarily converged or "good") if the algorithm finishes the calculation, regardless of the quality or outcome of the retrieval. It is considered to have failed if the algorithm aborts the calculation, for example when unphysical values for the surface albedo (> 1.0) are encountered.



Figure 7.8: Two box plots showing the distributions of the χ^2 for both micro-windows and the three polarisations separately (only for the year 2010). The S-polarised retrievals are closer to 1, and also exhibit lower spread than the P- or PS-polarised ones.

The initial filtering is tabulated in Table 7.1. Due to a lack of truth data, the initial filtering is performed only to filter out large outliers and non-converged retrievals to provide a better baseline for the bias correction.

	Threshold(s) / Criteria
Fit quality	$0.5 \le \chi^2 \le 2.0$
Number of iterations	$N_{\rm iter} = 2$
Signal-to-Noise ratio	SNR > 50
Solar zenith angle	$SZA < 65^{\circ}$
Mean radiance in Band 1	$1.5 \cdot 10^{-7} < \bar{I}^{(1)} < 1.0 \cdot 10^{-6}$
Relative Δp_{surf}	$90\% < \Delta p_{ m surf}^{ m rel} < 110\%$

Table 7.1: Initial filter thresholds used to select retrievals for the further bias correction procedure. Mean radiances are in units of $W \text{ cm}^{-2} \text{ sr}^{-1} \text{ cm}^{-1}$.

7.6 Scale-Offset Correlation

The retrieval results can be further analysed by investigating the correlations between the state vector elements. The correlation matrix can be obtained from the posterior covariance matrix $\hat{\mathbf{S}}$:

$$C_{ij} = \frac{\hat{S}_{ij}}{\sqrt{\hat{S}_{ii}\hat{S}_{jj}}}.$$
(7.12)

Correlation coefficients C_{ij} can reveal whether state vector elements are coupled due to the fact that the corresponding Jacobians are similar. The correlation matrix in Figure 7.9 highlights two pairs of state vector elements that are very highly anticorrelated. The elements d_0 and d_1 , the dispersion shift and dispersion stretch, are unsurprisingly highly anti-correlated. For very narrow spectral windows, the Jacobians are similar enough to give rise to such a correlation coefficient close to -1.



Figure 7.9: Mean correlation matrices for the 755 nm (left) and 772 nm (right) microwindows (S-polarised spectra depicted, but P and PS correlations are very similar).

More concerning is the high anti-correlation between c_0 and z_0 , the scaling factor and the zero-level offset. An argument was presented earlier (see Section 7.1) explaining that reflectance can be decoupled from fluorescence, as the fractional in-filling of the solar lines is a result of fluorescence only. The correlation matrix shown in Figure 7.9, however, suggests that this decoupling cannot be achieved, as the retrieval algorithm can not distinguish between a reflectance scaling factor c_0 and an additive component z_0 .

High (anti-)correlations between these two state vectors are a result of the Jacobians $\partial \mathbf{F}(\mathbf{x})/\partial c_0$ and $\partial \mathbf{F}(\mathbf{x})/\partial z_0$ being very similar. The retrieval algorithm can "choose" to adjust either z_0 or c_0 with almost the same effect to the forward model (taking into account prior covariances for both).

The Jacobian of the zero-level offset d_0 is trivially just a constant 1, and the Jacobian for the reflectance scaling factor c_0 is the spectrum itself (see e.g. Figure 7.7). With $\mathbf{f}(\mathbf{x})$ being the forward model *without* the applied scaling and zero-level offset correction, the full forward model can be written as:

$$\mathbf{F}(\mathbf{x}) = z_0 + c_0 \mathbf{f}(\mathbf{x}),\tag{7.13}$$

and the Jacobians are

$$\frac{\partial \mathbf{F}(\mathbf{x})}{\partial c_0} = \mathbf{f}(\mathbf{x}),\tag{7.14}$$

$$\frac{\partial \mathbf{F}(\mathbf{x})}{\partial z_0} = 1. \tag{7.15}$$

Further following the argument, the similarity between the modelled spectra $\mathbf{f}(\mathbf{x})$ and 1, can be assessed through the cosine similarity between two vectors, defined as

$$\cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|^2 |\mathbf{b}|^2}.$$
(7.16)

Evaluating $\cos(\theta)$ for the measured spectra and a unit vector of the same length results in a similarity of $\cos(\theta) \sim 0.997686 \pm 0.00024$ for about 15 000 spectra taken from May 2013. While the above considerations are not to be read as a sophisticated analysis of one similarity metric between the Jacobians associated with c_0 and z_0 , it is meant to show the following: high anti-correlation naturally arises due to

- the narrow spectral range of the micro-windows,
- around 90% of the spectral points belonging to the continuum, thus making the spectra almost flat.

How is a retrieval of SIF then still viable if reflectance and zero-level offset are so highly anti-correlated? The prior information about the Lambertian surface albedo is inferred directly from the spectra themselves. Surface reflectance is therefore informed through the measurement, such that c_0 is expected not to deviate from 1 by more than a few percent. So if the prior information about the Lambertian albedo is sufficiently close to the true value, c_0 will be close to 1 and z_0 in the same manner will only lightly deviate from the true value as well. Since the first guess (equal to prior) albedo is calculated as

$$\rho_{a} = \frac{\bar{I}^{(1)}\pi}{I_{0}\cos(\theta_{0})},\tag{7.17}$$

the retrieved albedo $\hat{\rho}$ will always differ from the prior, since Rayleigh scattering is considered in the forward model, but not in the calculation of the prior value (θ_0 being the solar zenith angle, $\bar{I}^{(1)}$ the estimated continuum level radiance, and L being the solar irradiance).

To estimate the magnitude of this potentially systematic bias, simulations were performed for a set of discrete levels of surface albedos, SIF magnitudes and solar zenith angles. Then, retrievals were run for each of these scenarios using the 755 nm window, where the prior surface albedo was varied between 90% and 110% of the *true* surface albedo of the simulation; the prior SIF magnitude was kept constant at $1 \cdot 10^{-8} \,\mathrm{W \, cm^{-2} \, sr^{-1} \, cm^{-1}}$.

Figure 7.10 depicts this systematic bias. First, the nature of the bias is almost linear in the studied range. As expected from the negative correlation coefficient (Figure 7.9), the slopes of the fitted lines are negative as well. The figure reveals the consequences



Figure 7.10: A systematic retrieval bias for two scenarios where only the true albedo differed (SZA: 5°, true fluorescence: $3 \cdot 10^{-8} \,\mathrm{W \, cm^{-2} \, sr^{-1} \, cm^{-1}}$) with $\rho = 0.25$ and $\rho = 0.35$. The markers are filled according to the prior albedo as a percentage of the true albedo.

of the anti-correlation between c_0 and z_0 . If the prior albedo is lower than the true albedo, the retrieved albedo will be slightly underestimating the true value $(\hat{c}_0 - c_0 < 0)$ and thus results in an overestimation of SIF $(\hat{z}_0 - z_0 > 0)$. In the opposite case, where the prior albedo is larger than the true albedo, the retrieved scaling factor will be slightly larger than the true value $(\hat{c}_0 - c_0 > 0)$, which finally results in an underestimation of the SIF $(\hat{z}_0 - z_0 < 0)$.

The results of the full set of 200 scenarios (2 SZAs, 10 albedos, 10 SIF magnitudes) is seen in Figure 7.11. It shows the maximal extent of the deviation of \hat{z}_{0} - z_{0} when the prior albedo is varied between 90% and 110% of the true albedo. The figure makes it clear that the extent $\max(\hat{z}_{0}-z_{0}) - \min(\hat{z}_{0}-z_{0})$ is mostly dependent on the true albedo and the solar zenith angle, and largely independent on the true SIF. The extent of the deviation is roughly between $0.5 \cdot 10^{-11}$ to $2 \cdot 10^{-11}$ W cm⁻² sr⁻¹ cm⁻¹, which is about 2 orders of magnitude below the reported uncertainty of about $5 \cdot 10^{-9}$ W cm⁻² sr⁻¹ cm⁻¹.

7.6.1 Conclusions

To summarise and conclude this section, the c_0 - z_0 anti-correlation is explained as a result of the corresponding Jacobians being very similar - which again is mainly due to the fact that the retrieval micro-windows are very narrow, and as such the spectra consist largely of the continuum. To estimate the magnitude of this potentially systematic bias, a number of simulations and retrievals spanning a range of albedos and SIF magnitudes were performed. The simulations and retrievals confirmed that

0.25 -	1.440	1.440	1.440	1.440	1.440	1.440	1.440	1.440	1.440	1.440
. ංචු 0.3 -	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206
eque 0.35 -	1.038	1.038	1.038	1.038	1.038	1.038	1.038	1.038	1.038	1.038
- 4.0 E	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911	0.911
ung 0.45 -	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.812
- 5.0 rtian	0.732	0.732	0.732	0.732	0.732	0.732	0.732	0.732	0.732	0.732
on dia management of the second secon	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667
- 6.0 F	0.612	0.612	0.612	0.612	0.612	0.612	0.612	0.612	0.612	0.612
₽ ^E 0.65 -	0.565	0.565	0.565	0.565	0.565	0.565	0.565	0.565	0.565	0.565
0.7 -	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525
	- 0.0	5e-09 -	1e-08 -	1.5e-08 -	2e-08 -	2.5e-08 -	3e-08 -	3.5e-08 -	4e-08 -	4.5e-08 -

 $SZA = 5^{\circ}$

	0.25 -	2.010	2.010	2.010	2.010	2.010	2.010	2.010	2.010	2.010	2.010	
rtian Surface Albedo	0.3 -	1.682	1.682	1.682	1.682	1.682	1.682	1.682	1.682	1.682	1.682	
	0.35 -	1.448	1.448	1.448	1.448	1.448	1.448	1.448	1.448	1.448	1.448	
	0.4 -	1.272	1.272	1.272	1.272	1.272	1.272	1.272	1.272	1.272	1.272	
	0.45 -	1.134	1.134	1.134	1.134	1.134	1.134	1.134	1.134	1.134	1.134	
	0.5 -	1.023	1.023	1.023	1.023	1.023	1.023	1.023	1.023	1.023	1.023	
umbe	0.55 -	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	
ie La	0.6 -	0.856	0.856	0.856	0.856	0.856	0.856	0.856	0.856	0.856	0.856	
Tri	0.65 -	0.792	0.792	0.792	0.792	0.792	0.792	0.792	0.792	0.792	0.792	
	0.7 -	0.736	0.736	0.736	0.736	0.736	0.736	0.736	0.736	0.736	0.736	
		0.0 -	5e-09 -	1e-08 -	1.5e-08 -	2e-08 -	2.5e-08 -	- 3e-08 -	3.5e-08 -	4e-08 -	4.5e-08 -	
			Irue SIF Magnitude									

 $SZA = 65^{\circ}$

Figure 7.11: The difference between largest and smallest value of the deviations $\hat{z}_0 - z_0$ (in units of $1 \cdot 10^{-11}$ W cm⁻² sr⁻¹ cm⁻¹) when the prior albedo is perturbed between 90 % and 110 % of the true albedo.

if the prior albedo is different from the true albedo, both albedo and SIF cannot be fully recovered. As predicted from the sign of the correlation coefficient, an underestimation of the scaling (or albedo) will result in an overestimation of SIF, and vice versa.

For the simulations, the variation of the prior albedo was restricted to ± 10 %, and the absolute deviation of the retrieved SIF from the true SIF was less than three orders of magnitude smaller than the true SIF value itself ($2 \cdot 10^{-11}$ W cm⁻² sr⁻¹ cm⁻¹), and two orders of magnitude less than the expected uncertainty.

Since the calculation of the a-priori albedo is performed without accounting for Rayleigh scattering, the prior value will generally be higher than the retrieved value. As a consequence, the retrieved scaling factor c_0 will be overestimated, which leads to a systematic underestimation of SIF. As the results of the simulations indicate, this systematic underestimation is less than a percent of the typical SIF magnitude. Considering these numbers, the conclusion is that this systematic underestimation of SIF is not significant when compared to noise levels.

7.7 Bias Correction

After the initial filtering of the results, the retrieved z_0 (*raw* SIF) can be plotted on a global map, as done in Figure 7.12. While the map shows some of the expected



Figure 7.12: Gridded $(2^{\circ} \times 2^{\circ})$ retrieved z_0 for the year 2012 (P-polarised, 755 nm window).

features, such as the strong signals at the locations of tropical forests, there is a clear negative bias throughout the rest of the globe. There are striking and unphysical
negative values over the Sahara desert, the Arabian peninsula and central Australia. As pointed out by Frankenberg, Fisher, et al. (2011), this is a result of an instrumental artefact - a non-linearity issue originating in the analogue-to-digital conversion of the FTS interferograms. If the conversion suffers from unaccounted non-linearities, then the resulting effect on the measured spectra will be very similar to the infilling of spectral lines, and thus cannot be distinguished from SIF. This section explores the nature of these biases as well as how they are compensated for.

Looking at the strong negative retrieved z_0 over desert areas in Figure 7.12, they coincide with the location of medium-gain measurements (see Figure 3.1, Page 25 for an explanation of the gain modes). Since the source of the bias is instrumental, it makes sense to investigate and apply any bias correction to the different gain modes as well as the different micro-windows and polarisations separately.

The general treatment as laid out in Frankenberg, Fisher, et al. (2011) was followed to characterise this instrumental bias. For every single observation, the mean radiance in the O₂ A-band $\bar{I}^{(1)}$ is used as a proxy for the scene brightness. All scenes for which there should not be any contribution due to SIF were then collected. While Frankenberg, Fisher, et al. (2011) were using all of Antarctica for high-gain, and a region in the Sahara desert for medium-gain spectra for this purpose, a more advanced approached was used. All soundings for which over 95% of the pixels within the GOSAT footprint belong to the general land-cover classes (see Section 7.4) urban (index: 190), bare area (indices: 200, 201, 202), or snow and ice (index: 220), are considered to be devoid of vegetation. For these measurements, a true SIF of $0 \,\mathrm{W}\,\mathrm{cm}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{cm}^{-1}$ can be assumed, even though some areas might experience nonzero SIF due to transient vegetation activity such as desert blooming. SIF measurements would then be mis-attributed if the land cover map does not account for these phenomena. The number of such SIF measurements are assumed to be small, so the overall calibration should not be affected by effects like these. Whilst the vast majority of selected measurements are indeed located over Antarctica in the winter season, this approach allows for sufficiently enough measurements to fulfil the selection criteria during other seasons. These sounding locations represent areas that are *permanently* bare, such as ice shields or deserts, while areas that experience a seasonal cycle are not considered. This guarantees that the approach is not a self-fulfilling prophecy in which the SIF values for non-permanently vegetated areas (e.g. forests during local winters) are calibrated to zero by default.

Figure 7.13 shows this bias for the P-polarised spectra, 755 nm micro-window, for 2012. The two gain modes behave very differently, as was already anticipated from Figure 7.12. The bias shows a non-trivial dependency on the mean band radiance for the high-gain measurements, with two kinks near $4 \cdot 10^{-7}$ W cm⁻² sr⁻¹ cm⁻¹ and



Figure 7.13: Radiance-dependent bias for P-polarised spectra, split up into the two micro-windows and the two gain modes. The solid, black line (derived using the LOWESS technique) is used to highlight the shape of the bias.

 $5.5 \cdot 10^{-7}$ W cm⁻² sr⁻¹ cm⁻¹. These kinks were observed already for older versions of the GOSAT L1B, however their positions are shifted compared to Figure 7.13. The medium-gain measurements span a smaller range of scene brightnesses, and do not exhibit any strong radiance-dependent shape, but rather a near-constant negative bias of approximately $1.6 \cdot 10^{-8}$ W cm⁻² sr⁻¹ cm⁻¹. The ordinates in Figure 7.13 are scaled the same for both micro-windows. It shows that the spread of raw SIF values is systematically larger for the 755 nm window. This is explained by the posterior uncertainty of the retrieved z_0 , which is, on average, twice as large as for the 772 nm window.

The impact of measurement polarisation is shown in Figure 7.14. The bias in absolute numbers is strongest for S-polarised spectra, and also systematically larger for the 772 nm retrievals. This is not surprising, as the S-polarised spectra are roughly 7% brighter than the P-polarised ones.

Figures 7.13 and 7.14 show the instrumental bias for all measurements in the year 2012. The procedure to correct the raw SIF is to sample calibration curves (Figure 7.14)



Figure 7.14: For easier visualisation, the point clouds as seen in Figure 7.13 are replaced by LOWESS-smoothed curves. For both gain modes, the same observation is made: S-polarised correction curves are shifted up (compared to P-polarised ones), as are the curves for the 772 nm window.

and subtract the radiance-dependent value,

$$F_{\rm s} = z_0 - z_0^{\rm (corr)},\tag{7.18}$$

from the raw SIF. This leads to the selected soundings to be corrected such that they are, on average, around zero. Mapping the same measurements as in Figure 7.12, but displaying the bias-corrected values, results in a much more plausible distribution and overall picture, as seen in Figure 7.15. The negative SIF values that cover large areas in Figure 7.12 are largely gone and shifted towards zero. After bias correction, large-scale features like tropical forests (e.g. Amazon, central Africa) and areas with intense agricultural use (e.g. eastern US) stand out.

These maps and curves have been derived for selected retrievals from the year 2012. There is, however, no reason to assume that the calibration curves did not change throughout GOSAT's operational lifetime from April 2009 onward. After all, the source of the bias is instrumental, and both instrumental degradation as well as their attempted correction in the L1B data might affect the radiance dependence of these curves.

To show the time dependence of the bias, roughly half the time series (2009-2013) is shown in Figure 7.16. Biases for both high- and medium-gain measurements exhibit a time-dependence, most prominently as a strong initial decrease of the bias from 2009 to 2010 and then a slower decrease in the subsequent years. In the next chapter, time



Figure 7.15: Same data as Figure 7.12, but using bias corrected values. Regions that are considered permanently free of vegetation (Sahara, Antarctica, Greenland, etc.) now exhibit roughly 0 fluorescence.



Figure 7.16: The correction value $z_0^{(\text{corr})}$ as a function of $\overline{I}^{(1)}$ and time. The left-hand (right-hand) side plot shows the values for high-gain (low-gain) measurements. As an example, a trapezoidal area for a winter season is highlighted where measurements over Antarctica dominate the derivation of the correction.

series of the retrieved SIF are investigated. The bias corrected SIF, F_s , is directly influenced by the correction term $z_0^{(\text{corr})}$ (Equation 7.18), and any time-dependent feature in the correction term will by mapped onto F_s . In order to account for the this dependence, three approaches will be utilised.

The first one, is essentially what has been used throughout this section so far: correction curves are derived on annual aggregates and applied to the raw SIF depending on the year of the measurement. The second approach is similar, but the time-base is reduced from annual aggregates to seasonal (3-monthly: DJF, MAM, JJA, SON) ones. Seasonal aggregates have the advantage that variations of the calibration curves within the year are captured better, with the obvious disadvantage that there are (depending on the season) potentially far less measurements to extract the curves from. As a third approach, the time dependence is taken into account explicitly. Here, $z_0^{(\text{corr})}$ in Figure 7.16 is split into bins along the $\bar{I}^{(1)}$ dimension. Bin boundaries are chosen such that the shape of the curves in Figure 7.14 is sufficiently well captured. For every one of these bins, the data is then averaged in monthly bins and a smoothing of the values is achieved through the LOWESS³ technique (see Figure 7.17). This allows for the time-dependence and the inter-annual variations to be accounted for. Other methods, such as fitting high-order polynomials or other compound functions involving Fourier series were less successful in describing the seasonal cycles.



Figure 7.17: For the noted bin in $\bar{I}^{(1)}$ -space, the values of $z_0^{(\text{corr})}$ (S-polarised, 755 nm, high-gain) are aggregated and the mean and standard deviation are shown in the above figure.

 $z_0^{(\text{corr})}$ is sampled at the centres of the radiance bins and at monthly grid points according to the LOWESS-smoothed curves (see Figure 7.17). To cover the full radiance range between the filter thresholds (Table 7.1), the grid points at $1.5 \cdot 10^{-7}$

³Locally weighted scatter plot smoothing - the implementation in the statsmodels Python package was used.

and $1 \cdot 10^{-6}$ W cm⁻² sr⁻¹ cm⁻¹ are obtained through linear extrapolation based on the closest two points, however the time dependence is described through the same curves as the closest grid point. Using the grid points, the function value at every point between the grid points is obtained through a third order Spline interpolation.

The bias correction procedure is performed for all three polarisations, the two microwindows, and both absolute and relative SIF. In the following sections and chapters, it is implied that only those measurements that passed the initial quality filtering are considered.

7.8 Assessment of Retrievals and Comparison to GOME-2

After applying the bias correction and initial filtering, the total number of measurements is around 4 million - less for P and PS, and a bit more for S-polarised measurements (Figure 7.18). The largest contributor for measurements being filtered out are low signal levels with a mean radiance below $1.5 \cdot 10^{-7}$ W cm⁻² sr⁻¹ cm⁻¹. Since the same filter threshold was used for all polarisations, and the P-spectra are generally slightly lower in radiance than the corresponding S-polarised ones, there are about hundred thousand more accepted SIF retrievals for the S-polarised measurements.



Figure 7.18: Out of roughly 4 million retrievals, about 3.5 million pass the quality filtering, with the S-polarised measurements surpassing the P-and PS-polarised ones.

With no truth data being available on a relevant scale, the quality of the bias correction can not be easily assessed. Its temporal stability, however, can be investigated. Taking all bias-corrected $F_{\rm s}$ over non-vegetated regions (excluding Antarctica) as indicated by the land cover map, time series can be constructed. Ideally, this series would be flat over time exhibiting zero fluorescence. These series are shown in Figure 7.19 for the two micro-windows and the three calibration methods separately. They have been re-sampled to semi-monthly periods, the solid line represents the mean of all considered measurements in the re-sampling period. The mean of the time series are unsurprisingly all below $0.015 \cdot 10^{-8}$ W cm⁻² sr⁻¹ cm⁻¹, and the root-mean-squares suggest that the Spline-based calibration technique results in the smallest seasonal

amplitude (listed in Table 7.2). Performing the same analysis for relative SIF, it becomes clear that the Spline correction does not work as well as for absolute SIF radiances.



Figure 7.19: Stability of the bias-corrected SIF over Antarctica. Solid lines represent means, while coloured and shaded areas are one standard deviation away from the mean for that time period. The dashed, grey lines are the mean state vector uncertainties for the entire time series. Absolute SIF are in units of mW m⁻² sr⁻¹ nm⁻¹.

	Me	ean	RMS		
	755 nm	772 nm	755 nm	772 nm	
Annual	-0.0024	0.0044	0.031	0.035	
Seasonal	-0.0005	0.00022	0.024	0.024	
Spline	0.0082	0.013	0.021	0.025	

Table 7.2: Time series means and root-mean-squares (RMS) of SIF for the three calibration modes as seen in Figure 7.19. Values are in units of 10^{-8} W cm⁻² sr⁻¹ cm⁻¹. (S-polarised measurements only)

Figure 7.19 also confirms the magnitude of the reported uncertainty (as the squareroot of the corresponding diagonal element of $\hat{\mathbf{S}}$), since the value is close to the standard deviation in any given temporal interval. The seasonal cycle that is seen in the standard deviations in the time series is a result of Antarctica coverage increasing during the winter seasons - and thus the large number of high-SNR (low-noise) measurements are reducing the average uncertainty.

	Me	ean	RM	мs
	755 nm	772 nm	755 nm	772 nm
Annual	-0.0043	0.0055	0.046	0.064
Seasonal	-0.00012	-0.00047	0.035	0.042
Spline	0.012	0.014	0.028	0.038

Table 7.3: Same as Table 7.2, but for relative fluorescence (values in %). Notably, the Spline-based correction procedure performs worse than the other two methods in terms of the mean, the RMS is however lower.

In the supplementary materials of Frankenberg, Fisher, et al. (2011), the SIF retrievals in the 772 nm window are shown to be expectedly lower than for the other microwindow (see Figure 6.4, Page 66). For various maps and figures, a constant scale factor is applied, such that $F_{\rm s}^{(755)} \sim 1.8 \cdot F_{\rm s}^{(772)}$. The performed retrievals here show a similar trend, which is illustrated in Figure 7.20. SIF at 755 nm is systematically higher than at 772 nm, and this enhancement is best seen in the eastern US and the Eurasian forests.



Figure 7.20: Full time-series averaged on 2° by 2° degree bins, P-polarised spectra. The intensity of the green shading for tropical Africa and the eastern US clearly show the stronger SIF signal for the 755 nm micro-window.

The mentioned scaling factor of roughly 1.8, however, could not be reproduced. By dividing the gridded results from Figure 7.20, the scaling between the two retrieval windows can be displayed on a map, as seen in Figure 7.21. This map reveals that the overall scaling between the two retrieval windows does not quite reach 1.8 on a global scale, and is far from being homogeneous. While a good portion of Europe actually has a scaling factor higher than 1.75 (as does north-eastern China and the Indian subcontinent), most of the eastern US shows a factor between 1.25 to 1.5, and the tropical forests in Africa and South America even lower than 1.25. Figure 7.21, however, resembles a map of plant functional types (PFTs) or biomes, suggesting that the ratio is a function of vegetation type. The two maps have been created using



Figure 7.21: Ratio $F_{\rm s}^{(755)}/F_{\rm s}^{(772)}$ of SIF for the two retrieval windows. Grid cells with values lower than $0.2 \,\mathrm{mW}\,\mathrm{m}^{-2}\,\mathrm{sr}^{-1}\,\mathrm{nm}^{-1}$ were masked and appear white in order to highlight areas with a strong SIF signal.

P-polarised spectra. While the same plots using the S-polarised measurements does not exactly yield the same results, the differences are localised and most apparent over Europe, where the scaling factors appear smaller for S-polarised SIF. Comparing the polarisations side-by-side is shown in Figure 7.22, where the difference of the two micro-windows to an assumed scaling factor of 1.8 is shown. What this representation of the retrieval results shows is that the scaling factor of 1.8 is roughly realised for the P-polarised spectra everywhere apart from the tropics.

Since Frankenberg, Fisher, et al. (2011) do not show the ratio (or its derivation) explicitly, but rather scaled maps, it is difficult to point out differences to the results obtained in their publication. Figure 7.22 also suggests the P-polarised measurements being more in line with the expected ratio.

Assessing the GOSAT SIF retrievals against fluorescence derived from GOME-2 measurements can provide an overall check on whether the magnitude and spatial distribution are comparable. Following the procedure in Köhler et al. (2015), all measurements with a latitude larger than 60° S (i.e. excluding Antarctica) from 2010 until the end of 2016 have been collected and gridded onto a regular 2° by 2° grid. This 7-year average is then compared against a 7-year average of GOME-2 SIF, where the data was re-gridded to the same regular grid (from 0.5° by 0.5°). After gridding, all grid cells in which more than 100 valid GOSAT SIF retrievals fall into, and which exhibit a grid cell standard deviation of less than $1.5 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ (~3000 grid cells for P-polarised, and ~2700 for S-polarised spectra), are then compared against GOME-2 SIF on a cell-by-cell basis for all overlapping cells. Figure 7.23 shows this cell-by-cell comparison such that it can be compared to Figure 15 from Köhler et al. (2015), although the multi-year average in that publication spans a shorter range (only 2 years). What can be immediately noted is that the overall correlation is significantly better (compare to $R^2 = 0.68$), however the best-fit linear function exhibits a smaller slope (compare to y = 0.26 + 2.19x).



(b) P-polarised

Figure 7.22: The maps show the difference between $1.8 \cdot F_s^{(772)}$ and $F_s^{(755)}$ for S- and P-polarised spectra. The tropical forests in South America, Africa, and Indonesia exhibit the largest differences. In these regions, $F_s^{(755)}$ is considerably larger than $1.8 \cdot F_s^{(772)}$.



Figure 7.23: Overall comparison of Spline-corrected GOSAT SIF against GOME-2 SIF (Köhler et al. 2015). The 1:1 line is shown in solid, black; the best-fit line is dashed, red.

While Figure 7.23 shows the comparison for one polarisation and one calibration mode, the coefficients of determination (R^2) are listed in Table 7.4. The good agreement between the GOSAT SIF and GOME-2 SIF is not highly surprising, as already Guanter et al. (2012) have shown that even using two different retrieval approaches for the same measurements will result in very similar SIF values. In their Figure 10, the fitted linear function has a slope of 1.07, suggesting that the SIF magnitude itself is independently captured in the same way by the two retrieval methods.

	Р		C L	5	PS		
	755 nm	772 nm	755 nm	772 nm	755 nm	772 nm	
Annual	0.826	0.781	0.797	0.628	0.818	0.734	
Seasonal	0.829	0.790	0.799	0.642	0.818	0.739	
Spline	0.832	0.795	0.802	0.658	0.830	0.762	

Table 7.4: R^2 values between retrieved GOSAT SIF and GOME-2 SIF, compared on a cell-by-cell basis. Retrievals in the 755 nm micro-window exhibit a higher correlation, as do measurements using P-polarised spectra.

GOME-2 SIF is consistently larger than the values retrieved from GOSAT (in both micro-windows), which is rooted in the GOME-2 retrieval using the wider spectral window. As described in Köhler et al. (2015), the retrieval uses a spectral window that spans the range from 720 to 758 nm and covering the second peak of the broadband SIF spectrum (see Figure 6.4, Page 66). Given this broadband shape, scale factors of ~1.7 (755 nm) and ~2.2 (772 nm) between GOSAT SIF and GOME-2 SIF seem consistent. Another contribution to the GOME-2/GOSAT differences is the fact that MetOp-A, which carries the GOME-2 instrument, is in an orbit with local equatorial crossing time of 9:30h, compared to GOSAT's 13:00h. The SIF emission scales approximately with the cosine of the solar zenith angle (Frankenberg, Fisher, et al. 2011), also shown in a study by Amoros-Lopez et al. (2008), in which laboratory measurements using artificial illumination were conducted. The SIF emission at 9:30h is further away from the maximum of the diurnal cycle than the emission at 13:00h.

A comparison between the retrieved F_s and GOME-2 SIF has been established: due to the different spectral windows and the different local measurement time, the values obtained from GOSAT are systematically lower, and the scale factor is about 1.7 (2.2 for the 772 nm window). This relationship was established using the gridded L3 data as well as the gridded retrieved F_s to comply with Köhler et al. (2015). Aggregating this way, however, ignores the GOSAT-specific sampling pattern. In order to explore this relationship for various regions independently, the collocation as described in Section 7.4 (Page 73) was used. Grouping this analysis into the different TransCom regions (Figure 7.6, Page 76), the slopes of the GOSAT to GOME-2 relationship can be investigated for several continental-scale regions. Note that for the following graphs, the axes were switched, such that GOME-2 SIF is now displayed on the abscissa rather than the ordinate.

For all 11 TransCom regions over land, the SIF data was first sub-selected to contain only those GOSAT measurements, which had a GOME-2 measurement on the same day, with less than 5 hours difference between the two measurements. The GOSAT footprint had to fulfil the criterion that 90% of the footprint area overlapped with the GOME-2 footprint, and the cloud fraction of the GOME-2 footprint (Wang et al. 2008) had to be less than 0.25. This subset was then gridded into 2° by 2° bins, and the entire time-series from April 2009 until December 2016 was considered.



Figure 7.24: GOSAT/GOME-2 comparisons for various Transcom regions after 2° by 2° binning. Only collocated GOME-2 measurements were used where the time difference to the GOSAT measurement was less than 5 hours. The standard error of the estimated slopes and intercepts is given as well. Note that the axes are switched when comparing to Figure 7.23. This figure illustrates the region-dependent differences between GOME-2 and GOSAT SIF.

A few examples are shown in Figure 7.24. These plots illustrate not only that the overall relationship still holds when the data is sub-divided into regions, but also

that the slope of the best-fit function can differ between them. The slope differences between e.g. North American Temperate and Europe are statistically significant, given the standard errors on the slopes (\pm) . Table 7.5 lists the determined slope for both windows and all calibration modes for every TransCom region. The largest slopes are seen in Southern Africa, the smallest ones in boreal North America. Significant differences between slopes in this case are most likely caused by the difference in overpass time, rather than being a result of the measurement approach.

		755 nm			772 nm	
TransCom region	Annual	Seasonal	Spline	Annual	Seasonal	Spline
NA Boreal	0.311	0.314	0.313	0.143	0.160	0.154
	(0.197)	(0.202)	(0.201)	(0.065)	(0.074)	(0.069)
NA Temperate	0.540	0.538	0.543	0.404	0.402	0.405
	(0.720)	(0.720)	(0.723)	(0.582)	(0.577)	(0.585)
SA Tropical	0.392	0.405	0.398	0.287	0.288	0.294
	(0.288)	(0.309)	(0.302)	(0.148)	(0.147)	(0.156)
SA Temperate	0.405	0.416	0.418	0.332	0.333	0.336
	(0.694)	(0.686)	(0.688)	(0.657)	(0.649)	(0.655)
Northern Africa	0.494	0.503	0.504	0.343	0.337	0.345
	(0.710)	(0.725)	(0.722)	(0.464)	(0.457)	(0.470)
Southern Africa	0.634	0.640	0.642	0.574	0.581	0.579
	(0.778)	(0.779)	(0.780)	(0.697)	(0.693)	(0.692)
Eurasian Boreal	0.409	0.413	0.415	0.243	0.257	0.254
	(0.413)	(0.420)	(0.420)	(0.256)	(0.267)	(0.263)
Eurasian Temperate	0.464	0.466	0.462	0.348	0.345	0.344
	(0.641)	(0.641)	(0.638)	(0.560)	(0.548)	(0.552)
Tropical Asia	0.596	0.586	0.596	0.664	0.646	0.659
	(0.355)	(0.349)	(0.355)	(0.407)	(0.401)	(0.409)
Australia	0.445	0.437	0.441	0.407	0.397	0.402
	(0.288)	(0.274)	(0.282)	(0.305)	(0.286)	(0.289)
Europe	0.411	0.410	0.414	0.237	0.251	0.254
	(0.501)	(0.501)	(0.503)	(0.336)	(0.359)	(0.367)

Table 7.5: Slopes of fitted linear functions between GOME-2 SIF and GOSAT SIF (see Figure 7.24, intercepts are not shown). The values in brackets indicate the R^2 value for the subset, and bold-faced slopes are those for which the R^2 is larger than 0.5.

7.9 Uncertainties of the SIF Retrieval

The retrieval uncertainty is characterised by the square root of the corresponding diagonal element of the a posteriori covariance matrix (see Equation 5.22, Page 54). For most SIF retrievals, this uncertainty is around 100% of the SIF value itself, making the SIF retrieval a very noisy one. Figure 7.19 already showed that over time, the

standard deviation of SIF values measured over bare surfaces is comparable to the mean uncertainty.

For the application of SIF data in studies, it is important to establish that these uncertainties truly reflect the random nature of measurement noise in the retrieval, as well as provide the right magnitude. To demonstrate that this is indeed the case, an analysis is performed on the bias corrected SIF retrievals over Antarctica. While the bias correction does change the SIF value, the associated uncertainties are not adjusted during the bias correction procedure. From the Antarctica SIF retrievals, N values are randomly picked and their statistics are plotted in Figure 7.25 (for one bias correction type). The error bar diagrams show both the SIF values of the subset, as well as the values of the uncertainties of the subset. Means of subsets are represented by round symbols, and the error bars extend by the standard deviation of the subset. Again (similar to Figure 7.19, Page 90), the standard deviations of the subsets are very close to the means of the retrieval uncertainties, if sufficiently many SIF measurements are aggregated into the subsets.



(a) Spline-based correction, 755 nm.

(b) Spline-based correction, 772 nm.

Figure 7.25: F_s (blue, circles) and associated uncertainties (orange, squares). Error bars of F_s are close to the mean single-sounding uncertainties. This confirms that the width of the posteriori distribution accurately represents the uncertainty of the retrievals.

Through visual examination, the reported uncertainties are indeed normally distributed for larger N. Underlying this assessment with the Kolmogorov-Smirnov test (KS-test) against a normal distribution (with μ and σ calculated from the subset) is shown in Table 7.6, where the *D*-values and *p*-values of the test are shown. The table shows that even for large N the hypothesis of the subset being drawn from a normal distribution cannot be rejected.

From these results the following conclusions can be drawn. Bias-corrected SIF does not show any systematic features, as the mean SIF value is close to zero. The remaining variability has the characteristics of purely random noise, and follows a

		755 nm			772 nm	
N	Annual	Seasonal	Spline	Annual	Seasonal	Spline
1	n/a	n/a	n/a	n/a	n/a	n/a
2	0.26 / 1.00	0.26 / 1.00	0.26 / 1.00	0.26 / 1.00	0.26 / 1.00	0.26 / 1.00
5	0.25 / 0.92	0.25 / 0.91	0.24 / 0.93	0.21 / 0.98	$0.25 \ / \ 0.92$	$0.22 \ / \ 0.97$
10	0.24 / 0.52	0.25 / 0.50	0.23 / 0.58	0.21 / 0.74	0.19 / 0.86	0.19 / 0.86
20	0.12 / 0.94	0.12 / 0.95	0.12 / 0.95	0.11 / 0.97	0.09 / 1.00	0.09 / 1.00
50	0.07 / 0.95	0.07 / 0.97	0.08 / 0.86	0.08 / 0.94	0.06 / 0.99	0.07 / 0.96
100	0.06 / 0.89	0.05 / 0.97	0.05 / 0.94	0.06 / 0.91	0.07 / 0.69	0.06 / 0.86
1000	$0.02 \ / \ 0.70$	0.02 / 0.69	0.02 / 0.72	$0.02 \ / \ 0.62$	0.03 / 0.41	0.03 / 0.52

Table 7.6: For the subsets, in which N random retrievals (absolute SIF) over the Antarctic are drawn, the *D*-values / *p*-values of the Kolmogorov-Smirnov test are listed. The *D*-values are smaller than the *p*-values for all N, all windows, and all calibration types, meaning that the test fails to reject the hypothesis of the SIF values being drawn from a normal distribution.

normal distribution whose standard deviation is close to the value expected from the a posteriori covariance matrix.

For a repeated set of N measurements of the same surface, the reported uncertainties thus represent uncorrelated errors. The mean of the set

$$\overline{F_{\rm s}} = \frac{1}{N} \sum_{i=0}^{N} F_{\rm s,i} \tag{7.19}$$

therefore comes with an error that scales with \sqrt{N} :

$$\sigma_{\overline{F_s}} = \frac{\sigma_{F_s}}{\sqrt{N}}.$$
(7.20)

Since every measurement $F_{s,i}$ has its own uncertainty $\sigma_{F_{s,i}}$, the uncertainty on the mean can also be calculated by applying Gaussian uncertainty propagation, which yields

$$\sigma_{\overline{F_s}} = \frac{1}{N} \sqrt{\sum_{i=0}^{N} \sigma_{F_{s,i}}^2}.$$
(7.21)

Clearly, if all $\sigma_{F_{si}}$ are the same, then Equation 7.21 reduces to Equation 7.20.

In later sections, regional means are computed where the region size varies from a few km^2 up to entire continents. For subsets like those, the above equations do not fully hold, as they are sampling different surfaces, and the truth is often not known. In the case of Antarctica, the true value can be assumed, and thus the uncertainty on the mean can be stated through Equation 7.20.

The variability of a continental-scale subset is a combination of the variability due to different surfaces (and true SIF) and the uncertainties due to measurement noise. If

both components are uncorrelated, the total uncertainty on a regional subset can be written as

$$\sigma_{\text{(tot)}} = \sqrt{\frac{1}{N^2} \sum \sigma_{F_{\text{s,i}}}^2 + \sigma_{\text{(Region)}}^2}.$$
(7.22)

Realistically, the two uncertainties will not be entirely uncorrelated. Regions of low photosynthetic activity, such as deserts, tend to exhibit a higher surface albedo, therefore a higher signal-to-noise ratio, and thus a lower single-sounding uncertainty. This, however, is a second-order effect and the regional uncertainties are expected to be the largest contribution to the total uncertainty.

 $\sigma_{(\text{Region})}$ is the standard deviation of the regional ensemble, and has to be estimated. If a large and heterogeneous region is chosen, such as various TransCom regions (e.g. Europe, temperate regions) then the true SIF already has an intrinsic variability for that region. To calculate the total uncertainty for large regions, following strategy is employed. SIF data for a TransCom region is subdivided into clusters using two different methods:

- Using the *k*-means algorithm according to their longitude and latitude,
- Grouping the soundings by their land cover class.

The number of clusters for the k-means method is chosen by dividing the total number of measurements per region by 100. For the land cover clustering, all soundings with the same majority land cover class are grouped into the same cluster. For each cluster, the standard errors of the mean SIF are calculated, and the mean of all cluster-standard errors is then the estimated $\sigma_{(\text{Region})}$ for that region. Regardless of the method, if there are less than 50 measurements in any given cluster, that cluster is disregarded. Initial cluster centroids for the k-means method are chosen randomly, so re-running the analysis will lead to slightly different values for that method. The two clustering methods are fundamentally different, as the land cover method groups together soundings that have similar vegetation type but could be significantly far away from each other, whereas the k-means procedure does the exact opposite (grouping based on locality, potentially very different vegetation types). Using all quality-filtered measurements from February 2014, the results are shown in Table 7.7; the results for July 2014 are listed in Table 7.8.

As expected, the values for the uncertainty on the means $(\sigma_{\overline{F_s}})$ are considerably lower than the estimated regional uncertainties $\sigma_{(\text{Region})}$. Both *k*-means and LCbased methods yield similar values for $\sigma_{(\text{Region})}$ regardless of the region. The only larger outlier is the LC-based aggregation for Antarctica - however this is mainly due to Antarctica being a large landmass with essentially only one land cover class

(February 2014)		All			k-means	6		LC	
Region	N	$\sigma_{\overline{F_{\mathrm{s}}}}$	σ	N	$\overline{F_{ m s}}$	σ	N	$\overline{F_{ m s}}$	σ
NA Boreal	1763	0.02	1.08	14	-0.04	0.10	7	-0.06	0.09
NA Temperate	2053	0.02	1.26	18	0.05	0.12	10	0.00	0.12
SA Tropical	602	0.05	1.40	6	0.94	0.14	2	0.78	0.12
SA Temperate	1385	0.04	1.45	11	0.73	0.13	7	0.73	0.12
Northern Africa	3972	0.03	1.74	34	0.04	0.17	8	0.24	0.11
Southern Africa	1052	0.04	1.40	9	0.79	0.13	6	0.81	0.13
Eurasian Boreal	2603	0.02	1.09	20	-0.01	0.10	11	-0.02	0.09
Eurasian Temperate	5074	0.02	1.32	43	0.04	0.13	12	0.10	0.08
Tropical Asia	903	0.04	1.28	7	0.69	0.13	5	0.68	0.13
Australia	1275	0.04	1.35	12	0.24	0.14	6	0.23	0.11
Europe	986	0.03	1.13	9	0.13	0.11	6	0.10	0.11
Antarctica	18662	0.01	1.15	138	-0.02	0.11	1	-0.02	0.01
Global Mean σ	1.31 ± 0.19		0.12 ± 0.02			0.10 ± 0.03			

Table 7.7: Estimation of regional uncertainties. In the first group, σ is the regional standard deviation including all measurements. For the *k*-means and land cover (LC) methods, *N* is the number of clusters with more than 50 soundings. For means and uncertainties, values are in units of mW m⁻² sr⁻¹ nm⁻¹. The last row is the mean (± standard deviation) of all $\sigma_{(\text{Region})}$.

(July 2014)		All			k-mea	ns		LC	
Region	N	$\sigma_{\overline{F_{\mathrm{s}}}}$	σ	N	$\overline{F_{ m s}}$	σ	N	$\overline{F_{ m s}}$	σ
NA Boreal	1955	0.02	1.17	17	0.26	0.11	8	0.20	0.09
NA Temperate	2688	0.02	1.42	24	0.79	0.12	10	0.69	0.11
SA Tropical	1762	0.03	1.37	15	1.08	0.13	5	1.00	0.14
SA Temperate	2309	0.02	1.23	22	0.31	0.12	9	0.34	0.09
Northern Africa	3466	0.03	1.87	31	0.29	0.18	8	0.50	0.11
Southern Africa	2411	0.02	1.09	21	0.21	0.11	9	0.18	0.10
Eurasian Boreal	3752	0.02	1.20	34	0.41	0.12	10	0.45	0.09
Eurasian Temperate	5479	0.02	1.45	47	0.32	0.13	14	0.52	0.11
Tropical Asia	608	0.05	1.44	6	1.16	0.15	3	1.14	0.13
Australia	2014	0.02	1.12	17	0.11	0.11	7	0.17	0.09
Europe	2692	0.02	1.24	23	0.72	0.12	11	0.63	0.10
Global Mean σ	1.3	32 ± 0.2	20	0	$.12 \pm 0$.02	0	$.10 \pm 0$.02

Table 7.8: Same as Table 7.7, but for July 2014. For this month, there are no soundings over Antarctica.

(permanent snow or ice). Otherwise, the global overages for both methods are $0.10 \text{ mW} \text{ m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ and $0.12 \text{ mW} \text{ m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ per region.

The presented assessment of regional uncertainties, while cumbersome, offers a way of calculating the uncertainties for continental-scale subsets.

Chapter Eight

Influence of SIF on XCO₂ Retrievals

T has been shown by Frankenberg, Butz, et al. (2011) that neglecting the fluorescence radiance in the O₂ A-band will ultimately lead to systematic biases in other retrieved quantities: surface albedo (~1%), aerosol optical depths ($\Delta \tau \sim 0.1$), aerosol layer height (~5 km) and surface pressure (~10 hPa). Investigating these interferences with full-physics XCO₂ retrievals was done by Frankenberg et al. (2012). They obtained artificial measurements from an orbital simulator and applied several retrieval configurations. These configurations consist of a test matrix in which aerosols, surface pressure as well as fluorescence were retrieved both separately, combined, or not at all. They found a substantial dependence of the retrieved XCO₂ on the true F_s with a slope of roughly ~1 ppm %⁻¹ (in units of relative SIF) if SIF is not part of the state vector. Fitting SIF compensates the systematic bias to a large degree.

For the generation of the ESA GHG-CCI Climate Research Data Package v4¹ fullphysics XCO_2 data set at the University of Leicester, the opportunity arose to process a large fraction of the GOSAT operational time frame both with and without a derived SIF prior (see Equation 4.18, Page 42). Unlike Frankenberg et al. (2012), the aim was not to make a set of controlled, simulated and noise-less observations, but rather to assess the impact of omitting the SIF prior through a large set of real retrievals. In this case, the true SIF value is unknown. It is important to note here that the SIF priors used in this exercise are not the same as presented in the previous sections, but the result of an earlier processing which used a monthly calibration routine. The omission of SIF will naturally have an effect on the spatial distribution of XCO_2 , as the Earth is not uniformly vegetated. Additionally, as SIF can have a strong seasonal cycle, the seasonality of the XCO_2 fields is expected to be altered.

The processed time-span ranges from January 2012 until December 2015, thus comprising four full years. Both retrieval sets feature a zero-level offset in Band 1, and the fluorescence radiance is in both state vectors - keeping in mind that the a priori covariance for the fluorescence ($\sigma = 1 \cdot 10^{-10} \,\mathrm{W \, cm^{-2} \, sr^{-1} \, cm^{-1}}$) is not large enough to allow for a significant departure from the prior value.

¹For v3 see Buchwitz, Reuter, Schneising, Hewson, et al. (2017).

8.1 Analysing Retrieval Results

In this section, the raw retrieval results for all soundings over land with a non-zero SIF prior are compared against the same retrievals where the a priori SIF is set to zero. The overall convergence² is essentially unchanged, out of 534 475 retrievals that satisfy the aforementioned criteria, 531 545 of the non-SIF set converged, whereas the number for the SIF set is 531 658. The number of iterations shows a significant change, as there are ~16% of retrievals that have less iterations using a non-zero SIF prior, as well as ~14% that have more iterations. The number of divergent steps are reduced for ~7.5% of retrievals, and increased for ~11%. Figure 8.1 shows the change in the number of iterations compared to the noSIF case, broken down by the number of iterations. For less than five iterations in the noSIF set, the change is small as the bulk of retrievals exhibits the same number of iterations. For more than five, a general reduction is seen.



Figure 8.1: The violin plot visualises the change of number of iterations (SIF - noSIF) when compared the set of retrievals with a zero SIF prior (noSIF). When more than four iterations were required for the noSIF case, the SIF case shows less needed iterations in general.

Fit quality for both sets is characterised by assessing the distribution of two related variables: the reduced χ^2 and the absolute root-mean-square radiance residuals for the three bands. The medians as well as the interquartile ranges of the two retrieval sets are listed in Table 8.1.

While the differences are not large, introducing the non-zero SIF prior results in a slightly worse fit quality for bands 1 and 3. The differences in residuals are negligible, and the number of iterations is reduced for cases in which the noSIF retrievals had

²See Section 5.3, Page 50 for convergence criteria.

	Median (SIF)	Median (noSIF)	IQR (SIF)	IQR (noSIF)
$\chi^2_{(\text{Band 1})}$	1.164	1.159	0.293	0.290
$\chi^2_{(\text{Band }2)}$	1.124	1.126	0.337	0.338
$\chi^2_{(\text{Band }3)}$	0.942	0.940	0.315	0.315
RMS _(Band 1)	$3.193\cdot10^{-9}$	$3.185\cdot10^{-9}$	$1.152\cdot10^{-9}$	$1.523\cdot10^{-9}$
RMS(Band 2)	$1.733\cdot10^{-9}$	$1.734\cdot10^{-9}$	$9.960 \cdot 10^{-10}$	$9.974\cdot10^{-10}$
RMS _(Band 3)	$9.288 \cdot 10^{-10}$	$9.273\cdot10^{-10}$	$6.926 \cdot 10^{-10}$	$6.916 \cdot 10^{-10}$

Table 8.1: A table summarising the fit quality indicators for both sets of retrievals. Using a physical SIF prior appears to lead to overall slightly worse fits.

originally more than four iterations. These cases, however, are less than 17% of all considered retrievals.

The dependence between F_s and the change in retrieved XCO₂ as predicted by Frankenberg et al. (2012) is not reproduced, as shown in Figure 8.2. Not only is the gradient half as steep (~0.5), but the correlation between $F_s^{(rel)}$ and Δ XCO₂ = XCO₂^(SIF) – XCO₂^(noSIF) is very low ($R^2 = 0.048$) as the data shows large scatter.



Figure 8.2: XCO_2 differences versus SIF prior on a sounding-by-sounding basis. The while, dashed line indicates the gradient of 1 ppm %⁻¹ (Frankenberg et al. 2012), and the solid, black line is the actual best-fit to the data.

Figure 8.2 also shows that using the physical SIF prior alone is not a strong predictor to predict the change in XCO₂, at a single-sounding level. An attempt was made to find a linear model using various physically relevant parameters that could predict the Δ XCO₂ better than using the a priori SIF. However, no combination of parameters could produce a better fit to the observed Δ XCO₂. Already in Section 7.8 (Page 89) the intercomparison between the SIF retrievals and GOME-2 has shown that due to the large uncertainty for single retrievals, spatial and temporal aggregation is necessary for a comparison to be successful. While the reported uncertainty of an XCO₂ retrieval is generally less than a percent of the retrieved value, it is of comparable size to the Δ XCO₂, which is expected to be on the order of ~1 ppm.

8.2 Spatial and Seasonal Patterns

Exploring the spatial and temporal patterns of the impact of SIF on the retrieved XCO_2 is done by aggregating the retrievals into 2° grid cells for each season, and keeping only grid cells with more than 50 soundings. The four maps in Figure 8.3 reveal a clearer picture on the reduction of retrieved XCO_2 as a consequence of a non-zero SIF prior.



Figure 8.3: Seasonal maps of $\Delta XCO_2 = XCO_2^{(SIF)} - XCO_2^{(noSIF)}$ showing the regional pattern which correspond to the vegetated areas.

The pattern of ΔXCO_2 closely follows the seasonal cycle of vegetation on a global scale (see e.g. Figure 9.1, Page 113). Tropical South America exhibits a negative bias throughout the year, and regions with large areas devoid of vegetation (Australia, South Africa) only show slight changes during the local growing season. Using the gridded data, rather than the soundings (Figure 8.2), the relationship between the SIF prior and ΔXCO_2 appears in a much clearer fashion. Although the R^2 value increases significantly for the gridded set (0.34 in DJF, 0.44 in MAM, 0.62 for JJA and SON), the relationship is still noisy, as shown in Figure 8.4.

Frankenberg et al. (2012) have already explained the mechanistic influence of SIF on the XCO_2 retrieval. In the O_2 A-band, the missing information about SIF is



Figure 8.4: ΔXCO_2 as a function of the SIF prior, like Figure 8.2, but with grid cell averages, based on seasonal aggregation.

incorporated into a change in surface pressure and aerosol concentration. The impact on aerosols is not straightforward, since in a three-band retrieval using the strong CO_2 band at 2.06 µm, the spectral properties of the aerosols are further constrained. A change in retrieved surface pressure has a very direct effect on the final retrieved XCO_2 by simply altering the dry airmass to which XCO_2 is normalised to. Figure 8.5 shows the change in XCO_2 as a percentage of the noSIF XCO_2 as a function of the relative change in surface pressure. The Δp_{surf} -to- ΔXCO_2 relationship exhibits a higher correlation (0.56 < R^2 < 0.73) than the relationship between ΔXCO_2 and the a priori SIF. The slopes in Figure 8.5 were derived using a robust linear model where no intercept was fitted. For all four seasons, the slope magnitudes are similar, and suggest that for every 1% of surface pressure change induced by the different SIF prior, XCO_2 is reduced by about 2%. One would naively assume that the change in p_{surf} fully maps into a change in XCO₂. The introduction of a non-zero SIF prior however also changes the retrieved aerosol optical depths. Using a linear model to predict the change in XCO_2 confirmed this assumption: when the changes in aerosol optical depths are introduced to the linear model as explanatory variables, the predictive capability of the model increased significantly.



Figure 8.5: Seasonal aggregation as in Figure 8.4, but showing the effect on surface pressure instead.

8.3 Time Series

As a consequence of Figure 8.2, the seasonal dependence of ΔXCO_2 is naturally stronger in regions with a strong vegetation cycle (boreal and temperate regions). At the same time, the overall magnitude of ΔXCO_2 in tropical regions (with higher overall F_s) is larger, with a smaller seasonal component. Analysing time series for region-averaged XCO₂ provides a way of quantifying the impact of SIF on the scales of TransCom regions.

The retrieved XCO_2 is collected and grouped by TransCom region and then averaged into semi-monthly bins. An example of such a resulting time series is shown in Figure 8.6, which shows the XCO_2 for the Tropical Asia region. The middle panel, showing the XCO_2 difference, exhibits a clear seasonal cycle that anti-correlates to the SIF cycle. The lowest panel shows the deviation in the retrieved aerosol optical depths for all three mixtures. All three mixtures are impacted by the presence of SIF, which explains why the surface pressure change does not map into XCO_2 difference 1:1.

From the ΔXCO_2 time series, characteristic features can be extracted. The mean magnitude $\overline{\Delta XCO_2}$ (the mean of ΔXCO_2) reflects the overall deviation compared to a zero SIF prior. The seasonal amplitude of ΔXCO_2 is calculated as the difference



Figure 8.6: XCO₂ for Tropical Asia. The middle panel shows the difference between using a SIF prior, and not using one, Δ XCO₂, along with the magnitude of the SIF prior. For this specific region, $F_{\rm s}$ and Δ XCO₂ are anti-correlated (R = -0.51). The bottom panel shows the relative change in retrieved AOD for the three mixtures separately.

between highest and lowest value of ΔXCO_2 in a given year, referred to as peak-to-peak amplitude. The mean and standard deviation of these 4 annual values are retained. Figure 8.7 shows the eleven regions in a space spanned by these two parameters.

According to Figure 8.7, the most affected regions are boreal regions with the large seasonal amplitudes, tropical South America with its high productivity throughout the year, as well as Europe and Tropical Asia. All these regions exhibit a mean seasonal amplitude (in ΔXCO_2) larger than 0.5 ppm.

Since the SIF seasonality is a periodic contribution that affects XCO_2 retrievals every year, the effect of SIF on the measured CO_2 growth rate is very small. For northern mid-latitudes (30°N - 60°N), the differences in annual growth rates between 2012 and 2015 is less than 0.01 ppm yr⁻¹. Tropical regions (20°S - 20°N) as well as southern mid-latitudes show less than 0.1 ppm yr⁻¹ difference. Growth rates were calculated as the difference of the concentration at a given month minus the concentration at the same month for the year before.



Figure 8.7: This figure maps each region into a space spanned by the mean of the ΔXCO_2 time series, and its seasonal amplitude. Error bars represent the standard deviation of the seasonal amplitudes.

Figure 8.7 (along with Figure 8.3) also reveals that the overall effect of the SIF prior is generally less than 1 ppm, or less than ~0.25%. The seasonal amplitudes of the XCO₂ time series themselves are several factors larger, ranging from ~3.2 ppm for Tropical South America, to almost 14 ppm for boreal North America. Table 8.2 summarises the mean seasonal amplitude (based on monthly aggregates) for both XCO₂ and Δ XCO₂ for each region as well as the fraction Δ XCO₂/XCO₂, which gives an indication about the SIF prior effect magnitude. The largest fraction is unsurprisingly seen in the most productive region, Tropical South America, which also has the smallest seasonal amplitude in general.

	Seasonal A	Amplitude [ppm]		
	ΔXCO_2	XCO_2	Fraction	R
NA Boreal	1.2 ± 0.4	13.9 ± 3.9	8.60 %	+0.31
NA Temperate	0.3 ± 0.0	5.3 ± 0.9	5.59~%	+0.11
SA Tropical	0.6 ± 0.1	3.2 ± 0.8	18.95~%	-0.07
SA Temperate	0.4 ± 0.1	4.0 ± 0.4	9.57~%	+0.18
Northern Africa	0.3 ± 0.0	4.6 ± 0.2	6.60~%	+0.56
Southern Africa	0.3 ± 0.1	7.4 ± 1.0	3.69~%	-0.60
Eurasian Boreal	0.8 ± 0.3	10.6 ± 0.6	7.13~%	+0.22
Eurasian Temperate	0.2 ± 0.0	6.0 ± 0.3	4.07~%	+0.67
Tropical Asia	0.5 ± 0.1	5.6 ± 0.5	9.57~%	-0.51
Australia	0.2 ± 0.1	4.3 ± 1.8	5.53~%	+0.05
Europe	0.6 ± 0.1	6.5 ± 0.8	9.20~%	+0.42

Table 8.2: Peak-to-peak seasonal amplitudes for both XCO_2 and ΔXCO_2 per region. The third column is the ratio of the mean $\Delta XCO_2/XCO_2$, expressed as a percentage, the last column is the correlation coefficient between F_s and ΔXCO_2 .

8.4 Effect on Flux Inversions

Inclusion of the SIF prior results in sub-ppm, regional biases across the globe. These systematic biases in the concentrations will cause differences in the result of flux inversions. Without actually performing the inversions, however, it is not straightforward to estimate the effect magnitude, given the differences in XCO_2^3 .

Chevallier et al. (2007) have investigated the sensitivity of an inversion system to observational biases. The biases assumed in their study are related to sub-micron aerosol particles, with an average magnitude of 0.29 ppm. As a result of the imposed total mass conversation, fluxes are not only modified in the regions with a pronounced XCO₂ bias, but also elsewhere. Significant annual biases in net fluxes are seen over Southern oceans (0.2 Pg C yr^{-1}) and Eurasian Temperate and Europe regions (> 0.5 Pg C yr^{-1}).

These values are considerably high when they are compared to continental-scale net fluxes - Reuter et al. (2014), for example, estimate the European terrestrial carbon sink to be 0.58 ± 0.37 Pg C yr⁻¹, which is roughly the same value as the bias mentioned in Chevallier et al. (2007).

The biases due to SIF shown in Figure 8.3 may be modified due to bias corrections applied to XCO_2 retrievals, especially if affected variables (e.g. p_{surf}) are used in the bias correction procedure. Since retrievals are usually adjusted towards a reference, the differences between SIF and noSIF data *after* bias correction will likely be smaller than the values presented in this section.

In Figure 8.8 the *difference* between XCO₂ differences (SIF vs. noSIF) are shown to illustrate the effect of bias correction for one specific bias correction procedure. The bias correction in this case was taken from the UoL CCI product version 7.0. This specific product is somewhat different from previous versions and other products, as the bias correction is only dependent on two variables: the retrieved zero-level offset (z_0) , and the CO₂ profile gradient. The explicit formula to obtain the bias-corrected XCO₂ is

$$XCO_{2}^{(corr)} = XCO_{2} - \left(-16.5754 + 100.72550 \cdot z_{0} + 17.84797 \cdot CO_{2}^{grad}\right).$$
(8.1)

Figure 8.8 shows that the difference between bias-corrected ΔXCO_2 , compared to non-corrected ΔXCO_2 , is slightly smaller (globally: 0.03 ± 0.24 ppm). The differences do not fully follow the patterns of global vegetation, areas like the Sahel zone and India exhibit a change where there was no noticeable difference before.

Of course, this assessment of bias correction is not fully genuine and comprehensive. If noSIF and SIF XCO₂ retrievals are retrieved independently, they also should go

³P. Palmer, private communication.



Figure 8.8: Differences of $\Delta XCO_2 = \Delta XCO_2^{(corr)} - \Delta XCO_2^{(nocorr)}$. This figure illustrates the net effect of one particular bias correction (see Equation 8.1) on the concentration. The differences tend to be lower than 0.1 ppm, however the spatial pattern still persists.

through the bias correction procedure independently. Since retrieved quantities, like the retrieved surface pressure and the retrieved CO_2 profile gradient will be different, the bias correction coefficients would thus be different for both sets. The final, corrected concentrations would show smaller differences than depicted in Figure 8.8. **Chapter Nine**

SIF Case Studies



EFORE investigating more detailed regional case studies, a global picture of SIF can be established first, which will also help identifying potential discrepancies. The global vegetation pattern and their seasonal variation is generally well-captured by greenness indices like NDVI and EVI.

The expectation is that SIF follows the seasonal cycle of forests, cropland and grassland turning green in the local summers, and reducing their photosynthetic activities during the local winters. Global maps of seasonally averaged SIF (gridded to 2° by 2° degree cells), along with sampled EVI data and the grid cell counts, are shown in Figure 9.1. Seasonal changes of the global patterns appear for the SIF as they do for EVI. The eastern US is blooming during the MAM and JJA seasons, as is Europe and eastern temperate Asia. The tropical forests in southern America, Africa and Asia do not exhibit the same seasonality. A fairly large discrepancy between the SIF and EVI is seen in northern boreal Asia. During the JJA season, EVI shows the region to be very green, whereas the corresponding SIF panel shows the northernmost areas to be close to zero. A look at the grid cell counts for the JJA season reveals that this is not merely an artefact of sparse sampling in that specific area. While not pictured, the maps for NDVI show similar patters with a different range of values.

Collapsing the maps in Figure 9.1 into latitudinal averages, the good correspondence between SIF and the vegetation indices is further illustrated in Figure 9.2. As before, SIF captures the overall latitudinal dependence of the two vegetation indices, and the discrepancy in the IIA season for northern latitudes between 50° and 70° is highlighted as well. The latitudinal aggregate of EVI correlates slightly better with SIF for all seasons than NDVI, however the difference is only marginal.

In the next two sections, case studies are presented to demonstrate that SIF is capable of detecting the response of the biosphere to events like drought or anomalously high precipitation. Since the full time series of the processed GOSAT-SIF record spans almost 8 years, anomalies occurring between April 2009 and December 2016 should be detectable, as well as the return to a steady state.



Figure 9.1: Seasonal averages for SIF (755 nm, P-polarised) and EVI. The rightmost column shows the number of measurements in each grid cell.



Figure 9.2: Same SIF data as in Figure 9.1, but averaged into 2° latitude bins. The left ordinate is for SIF in units of mW m⁻² sr⁻¹ nm⁻¹, the right ordinate is for EVI and NDVI. EVI was scaled by a factor of 1.6 to make it align with the NDVI curve.

9.1 SIF-GPP Relationship

The intrinsic relationship between primary production and SIF is explained in Section 9.1. Deriving the relationship empirically has been already performed both in the publication demonstrating the retrieval technique (Frankenberg, Fisher, et al. 2011), as well as dedicated studies, such as done by Parazoo et al. (2014). In their publication, they employed a Bayesian optimal estimation algorithm to relate GOSAT SIF to GPP as described through an ensemble of eight dynamical global vegetation models (DGVMs). Similar to Guanter et al. (2012), the relationship is assumed to be dependent on the biome, and this is also reflected in the results: the GPP~SIF scaling factor is significantly different for the various biomes; the largest values are found for cropland, whereas the lowest ones are seen over shrubland.

While such an extensive study is not within the scope of this section, a simple linear model can show whether the SIF data exhibits a similar relationship to primary production. Rather than working on a grid-cell basis, the relationships here are derived through the time-series themselves.

Primary production is assessed through the NASA-CASA (Carnegie-Ames-Stanford-Approach) (Potter et al. 2003) gross ecosystem exchange (GEE) model output (1° × 1.25°, global, 3-hourly). Matching GOSAT SIF measurements with the CASA data set is done by picking the value from the grid cell in which the GOSAT footprint lies, using the time coordinate closest to the GOSAT measurement time. Since GOSAT SIF measurements only cover land areas, the uptake value is divided by the land fraction $c^{(\text{land})}$ of that cell, in order to normalise the flux to the amount that is occurring over land surfaces.

GEE and GPP are synonymous terms used in different fields of research, both meaning the uptake of carbon into the ecosystem. For the rest of this section, when discussing the relationship between SIF and uptake, "GPP" is used despite the CASA model output having the label "GEE". SIF generally follows the CASA GPP for large portions of the globe, see Figure 9.3. Northern hemispheric temperate and boreal regions exhibit the same seasonal cycle for both data sets, as shown in the first three panels of the figure. For the Europe-subset, the NDVI time series exhibits an extended plateau before it drops to the seasonal minimum in winter. This can be interpreted as a sign that SIF captures the reduction of photosynthetic activity after the summer peak, while the greenness index NDVI shows the leaves of the plants still being green.

The last two panels in Figure 9.3 show two regions, in which the seasonal cycles of SIF and GPP do not line up. Over the South Asia region (essentially the Indian subcontinent), the GPP cycle is delayed by roughly three months, whereas over the Amazon, the cycle is shifted by six months, making it anti-cyclic. The mismatch in



season cycle over tropical South America was already observed through flux inversion studies using OCO-2 measurements, suggesting a deficiency in the CASA GPP.

Figure 9.3: CASA uptake and MODIS NDVI compared to SIF for a few TransCom regions. Temperate and boreal regions show that all time series are very much in sync in terms of the seasonal cycle. For tropical regions, like South Asia and the Amazon, the NDVI time series shows the same seasonal cycle as SIF, but the CASA uptake exhibits a significant phase shift.

Parazoo et al. (2014) use the same SIF~GPP relationship as the one investigated in Frankenberg, Fisher, et al. (2011), which is a linear slope GPP = $\beta \cdot$ SIF. Instead of the full-fledged optimisation scheme that Parazoo et al. (2014) employed, an attempt is made to find a generally valid scaling factor β that holds for the majority of regions. To that end, the monthly averaged SIF and the collocated CASA GPP are compared. Collocated model GPP is sampled at the time of measurement, which is not necessarily representative of the daily average due to local overpass times at higher latitudes deviating from the equatorial one. Frankenberg, Fisher, et al. 2011 propose an up-scaling based on the local solar zenith $\theta_0(t)$ at the footprint location and its value throughout the day. The daily averaged fluorescence $\overline{F_s}$ is approximated

as

$$\overline{F_{\rm s}} = \frac{F_{\rm s}}{\cos\theta_0(t_0)} \int_{t=t_0-0.5}^{t=t_0+0.5} \cos\theta_0(t) \,\mathrm{d}t, \tag{9.1}$$

where the integration is performed over the local fractional day t between times $t_0 - 0.5$ and $t_0 + 0.5^1$. Overall, $\overline{F_s}$ is roughly one third of F_s , but deviates more towards higher latitudes.

9.1.1 Instantaneous to Daily

The collocated model GPP data at GOSAT overpass time itself has to be related to a daily integrated value. For every grid cell, the 3-hourly fluxes are aggregated on a daily basis, taking into account the local timezone of the grid cell location. These instantaneous values correspond to the daily uptake maxima, and the total daily uptake is calculated by integrating the values at the nine time steps throughout the day using the trapezoidal rule. The up-scaling factor is then determined through a linear model, which relates the daily maximal value $\widehat{\text{GPP}}$ to the daily integrated value $\widehat{\text{GPP}}$. This calculation is based on the CASA GPP data for the year 2012, so every grid cell contains 366 values which the linear model regression uses.



Figure 9.4: Illustration of the global distribution of scale factors α (left), and as latitudinal aggregates $\alpha(\phi)$ (right).

The global distribution of scaling factors is shown in Figure 9.4. While there is still a clear topographic pattern to be seen (especially towards the Himalayan plateau), the scale factor is well-captured through a latitudinal dependence only, and the relationship is found hold for all seasons of the year. The right panel in that figure shows the mean value for the scaling factor, collated by latitudinal bin, and the standard deviation is illustrated as the (barely noticeable) shaded area. The red,

¹Note that Frankenberg, Fisher, et al. 2011 actually integrate from t_0 to $t_0 + 1$, which is mathematically the same since the cosine of the local solar zenith θ_0 is periodic with a period of one day.

dashed line on top of the mean line is a 7th order polynomial fitted through the latitudinal means,

$$\overline{\text{GPP}} = \alpha(\phi) \cdot \widehat{\text{GPP}},\tag{9.2}$$

$$\alpha(\phi) = \sum_{i=0}^{7} \phi^{i} \alpha_{i}, \qquad (9.3)$$

with ϕ being the latitude in degrees, and the coefficients α_i from Table 9.1.

	_	
l	α_i	
0	$3.0324\cdot10^{-1}$	
1	$-4.5488 \cdot 10^{-4}$	Table 0 1. Coefficients
2	$1.7228\cdot10^{-5}$	α_i for the polynomial
3	$7.8438 \cdot 10^{-7}$	representation of α (see
4	$1.0629\cdot10^{-9}$	Figure <mark>9.4</mark>).
5	$-2.8683 \cdot 10^{-10}$	
6	$-8.5659 \cdot 10^{-14}$	
7	$3.3928 \cdot 10^{-14}$	

9.1.2 GOSAT to CASA

The result of the comparison is shown in Figure 9.5, and can be directly compared to Figure 2 in Frankenberg, Fisher, et al. (2011). It is important to note that while the both figures aim to visualise the same concept, i.e. the linear relationship between scaled SIF and primary production, the underlying data is differently aggregated. Frankenberg, Fisher, et al. (2011) use $4^{\circ} \times 4^{\circ}$ grid cells where for each cell, the data has been aggregated for an entire year. Each filled circle in Figure 9.5 represents a monthly aggregate between April 2009 and December 2015, and the size of the circle is related to the number of SIF measurements in that given month. For every region (as labelled through the figure legend), the barycentre of the point cloud is additionally marked by either a filled square or a filled triangle with the same colour. Regions with less than 10 000 total measurements were skipped, and months with less than 150 for that specific region are also dropped in this analysis. Compared to Frankenberg, Fisher, et al. (2011), the spatial variability does not play a large role, since large regions are aggregated, however the seasonality is captured much better as there are monthly means from a 6-year time series available.

The red, dashed line in Figure 9.5 represents an ordinary least-square fit for the model GPP = $\beta \cdot SIF$, excluding the data pairs of regions which are labelled with a triangle, i.e. the more productive areas that have been identified to have a mismatch



Figure 9.5: SIF-GPP relationship for various TransCom regions, aggregated by season. Each circle represents a monthly mean, and the size of the circle is proportional to the standard deviation for that month. The value β is inferred through ordinary least-squares.

with the SIF seasonal cycle. DJF, MAM and SON seasons overall exhibit a highly linear SIF-GPP relationship throughout the various regions, as indicated by the R^2 . Only during the JJA months is a more significant departure from a linear trend seen, despite the still rather high value of $R^2 = 0.885$. As already noted earlier, the high correlation between SIF and GPP in this case is not surprising, nor should it be over-emphasised. Both time series exhibit a seasonal cycle with the same period, and as long as they are in phase and have a significant seasonal amplitude, R^2 will be large, regardless of the predictive capabilities of SIF.

Problematic regions where the SIF-GPP correspondence is less clear, some of which were identified already in Figure 9.3, are mostly areas of high productivity with evergreen vegetation, such as the Amazon or Tropical Asia. The scaling factor β , which lies between 20.67 and 23.99, is comparable to the values seen in Frankenberg,

Fisher, et al. 2011, Figure 2².

9.1.3 GOSAT to FluxNet

A second option to derive β independently of models is a comparison to directly measured fluxes from towers. The FluxNet network offers a compact set of GPP measurements from tower sites situated mostly in North America and Europe, East Asia and Australia. A smaller number of sites are located in South America and Africa. The study by Guanter et al. (2014) exploited GOME-2 measurements to compare SIF to GPP from a small subset of tower sites in the US Corn Belt and over European grasslands.

GOME-2 measurements possess an advantage when performing comparisons against flux towers: footprint size. With an average footprint area of 3200 km^2 , a single SIF measurement can cover a much larger fraction of the flux tower footprint and is less sensitive to surface in-homogeneity. GOSAT's average footprint size of $\sim 85 \text{ km}^2$ is not only much smaller than the one of GOME-2, but the distance between each pointing location is usually more than 250 km apart. Even after the sampling pattern change in 2011, there are at most three soundings near a specific pointing location for a single overpass. In short, collocating GOSAT and FluxNet is a challenging task, and ideally requires a more sophisticated method of interpolating between sounding locations to get a good estimate of the SIF magnitude around the tower site.

Nevertheless, a simple collocation can reveal whether there are actually GOSAT measurements close to tower sites, and whether a comparison to the measured GPP results in similar values that were shown by Guanter et al. (2014). The collocation procedure, which is repeated for each and every FluxNet site, is as follows: only quality-filtered GOSAT measurements that are less than 50 km from the tower location are considered. The SIF measurements are first aggregated into daily means, along with the daily averaged GPP measurements, and then again sub-setted to those days in which both a SIF measurement and a tower measurement have occurred. Then, further aggregation into calendar months is done, and months in which less than three GOSAT measurements occurred, are disregarded along with months that exhibit negative average SIF and a standard deviation for that month > $2 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$. Finally, all sites for which less than 10 monthly aggregates are available are thrown out. For the remaining data pairs, the R^2 values are calculated. A total of 30 tower sites pass the criteria, and with a final selection criterion, 13 sites are picked which exhibit an R^2 value larger than 0.7: two in Australia (AU-Tum, AU-Whr), one in Belgium (BE-Vie), three in Switzerland (CH-Cha, CH-Fru, CH-Lae), another three in

 $^{^{2}}$ No numbers are stated, but the best-fit slope appears to be close to 20.


Germany (DE-Obe, DE-RuR, DE-Spw), and finally four in the United States (US-Los, US-MMS, US-PFa, US-Syv).

Figure 9.6: SIF from GOME-2 and GOSAT, compared to collocated FluxNet GPP measurements as described in the text. Circles represent monthly average, and their areas scale with the GOSAT SIF standard deviation of the measurements in that month. Dashed, red lines are the best-fit linear models, whereas the dashed, blue lines represent the 95% confidence intervals for the model prediction.

Validating the simple collocation procedure is accomplished by using the GOME-2 SIF measurements that were aggregated on a per-sounding basis (see Section 7.4, Page 73) - shown in Figure 9.6. The left panel of Figure 9.6 shows the (GOSATcollocated) GOME-2 SIF values against the FluxNet GPP measurements, where each circle represents a monthly mean aggregate and the area of the circle is scaled to the GOSAT SIF standard deviation for that month. Fitting a linear model through the data results in a scaling coefficient of $\beta = 3.645 \pm 0.107$, which is close to the value published by Guanter et al. (2014) ($\beta = 3.72^3$). The right panel shows the same GOSAT sounding locations, but the abscissa now are the unscaled GOSAT SIF values $(F_{\rm s})$. Using the empirically derived relationship between GOME-2 SIF and GOSAT SIF (Figure 7.23), the value of $\beta = 6.087 \pm 0.221$ can be converted to a GOME-2 appropriate one: $\beta = 3.484 \pm 0.127$, which is lower than the value derived from GOME-2 SIF alone. Considering the simplicity of the collocation approach, both values are well within 10% of the published value. This shows that the selection of FluxNet sites as a result of the collocation procedure provides a similar relationship between midday SIF and measured fluxes as the established values. Visualising the

³Guanter et al. (2014) fit an intercept and do not provide a standard error on the linear fit.

same collocated pairs, but using $\overline{F_s}$ is shown in Figure 9.6. This figure also shows the mean slope from Figure 9.5 ($\beta = 22.987$) in addition to the best-fit slope. The slope derived from FluxNet measurements is significantly lower than the value inferred from CASA.



Figure 9.7: Similar as Figure 9.6, but with *scaled* GOSAT SIF. The black line is the linear function using the mean slope from Figure 9.5 (excluding DJF).

9.1.4 Sampling Issue

The results in Figures 9.5 to 9.7, and therefore the empirically derived results in Guanter et al. (2014) and Frankenberg, Fisher, et al. (2011) do not seem to agree at first glance. While Figure 9.5 agrees with Frankenberg, Fisher, et al. (2011) with a slope of $\beta = 22.987$, and Figure 9.6 agrees with Guanter et al. (2014) for unscaled SIF ($\beta = 3.484$), the two scale factors are inconsistent, as the scaled SIF to FluxNet relationship is roughly 30% lower than what is derived through CASA comparisons (Figure 9.7). Using the same procedure as done in Figure 7.23, the relationship between instantaneous GOME-2 SIF and scaled GOSAT SIF is derived as $y = (4.985 \pm 0.00163) \cdot x$, which allows the value of $\beta = 3.72$ from Guanter et al. (2014) to be re-scaled to a value appropriate to $\overline{F_s}$. The resulting value is $\beta = 18.54$, which is still below the value obtained from the continental-scale SIF-CASA comparisons.

The challenge of collocating GOSAT measurements with FluxNet towers was mentioned before, and is thought to be the underlying reason for the observed discrepancy. To confirm a potential sampling issue, a sampling procedure similar to the GOSAT-FluxNet collocation was performed. Thirteen locations that lie within the same TransCom regions as the selected towers from Figure 9.6 are randomly generated to represent artificial tower locations. Then, GOSAT measurements are selected that lie within 50 km, 250 km, 500 km or 1000 km of those locations - if less than ten measurements are found, that location is rejected and another latitude-longitude pair is drawn from a uniform distribution. These subsets are then aggregated into monthly means, for which the scaled GOSAT SIF and the sampled CASA GPP (up-scaled to daily averages) are compared like before. The R^2 value, the best-fit slope β are again stored. This entire procedure is repeated a hundred times for each collocation radius to obtain different sets of thirteen random locations.



Figure 9.8: Each point corresponds in the left panel to a linear fit for a set of collocation using the collocation radius indicated by the colour. The right panel shows histograms of the fitted slopes β .

The resulting set of β is plotted in Figure 9.8, which shows the slopes as a function of R^2 , as well as an overall histogram. For the three collocation radii, the standard deviation of each distribution is around 2, however the means of the distributions vary strongly with the radius. When a 500 km or 1000 km radius is used, the model fits are generally good ($R^2 \ge 0.7$) and the slope is similar to Figure 9.5, $\beta^{500} = 19.46 \pm 1.97$ and $\beta^{1000} = 21.87 \pm 1.72$. The total number of observations for a 500 km (1000 km) radius is generally above 120 000 (500 000). Reducing the radius to 250 km (~40 000 observations), also shifts the distribution down to $\beta^{250} = 14.53 \pm 2.11$, as well as the coefficients of determination, which now lie closer to $R^2 \sim 0.5$. Using the same 50 km collocation radius (~2000 observations) as was done for the FluxNet comparisons results in a set for which the fits are all $R^2 < 0.5$, and the mean of the distribution of the slopes is $\beta^{50} = 6.34 \pm 1.93$ - although these predictive power of the model fits are too low for these slopes to be meaningful.

This result clearly shows that the GOSAT-CASA comparison is heavily dependent on the collocation radius. The dependence on the model fit R^2 suggests further that the more fundamental reason is the number of measurements that are used to derive the relationship. To show this, the sampling procedure outlined in the above paragraph

(the results shown in Figure 9.8) is repeated, however instead of the sampled and up-scaled CASA GPP, another value is used: using β , the fluxes are converted into scaled SIF. To incorporate the large single-sounding precision that is associated with each measurement, rather than just dividing the CASA value by a β , a value is drawn from a normal distribution for each measurement, where $\mu = CASA/\beta$ and $\sigma = \sigma_{F_s}$. The same model fit is then performed, and the expectation is for this new obtained scaling factor to be $\beta' \sim 1$, the model to be fitted is

$$\overline{F_{\rm s}} = \beta' \cdot \frac{\rm CASA}{\beta}.$$
(9.4)



Figure 9.9: Illustrating the sampling bias, the fitted values for β' are shown like in Figure 9.8 - the true value of $\beta' = 1$ is reached only for larger collocation radii ≥ 500 km.

Figure 9.9 confirms the sampling issue to be result of GOSAT's measurement sparsity and the high *a posteriori* uncertainty, rather than the variation in the CASA fluxes. β' approaches 1 with larger collocation radii, and the explanatory power of the fitted models increases as well. With this result, an argument can be made that a scaling factor closer to the result of Figure 9.5 ($\beta = 22.987$) is a more realistic representation of the relationship between $\overline{F_s}$ and flux.

9.1.5 Total Annual Uptake

Using the SIF~GPP relationship β from Figure 9.5, as well as the latitude-dependent up-scaling factor $\alpha(\phi)$, an instantaneous or daily-averaged SIF value as measured by GOSAT can be converted into an integrated flux for the day of the measurement. To build a more complete picture on the terrestrial carbon uptake, spatial averaging of SIF measurements has to be performed. The quickest way to do this, is aggregation into grid cells. The total uptake in any given grid cell is calculated in the following way: all measurements ($i \dots N$) that fall into a given grid cell, as well as between two points in time, are collected and converted into daily integrated GPP:

$$\overline{\text{GPP}}_i = \overline{F_{s,i}} \cdot \beta, \tag{9.5}$$

where $\overline{F_{s,i}}$ is the *i*th scaled SIF measurement. Uncertainties for both SIF and β are propagated to obtain an uncertainty on $\overline{\text{GPP}}_i$.

Within that grid cell and the given time boundaries, the $\overline{\text{GPP}}_i$ represent a sample of repetitively measured values of daily fluxes (g C m⁻² day⁻¹). Assuming that all measurements are representative of the daily flux of the entire grid cell, the $\overline{\text{GPP}}_i$ can be averaged and multiplied by the number of days between the time boundaries to obtain the total flux between those boundaries. Finally the area A of a grid cell has to be taken into account, which is derived from elementary geometry:

$$A = \int_{\Omega_i} R_e^2 \, \mathrm{d}\Omega = R_e^2 \int_{\phi_{\min} + \pi/2}^{\phi_{\max} + \pi/2} \int_{\varphi_{\min}}^{\varphi_{\max}} \sin\phi \, \mathrm{d}\phi \, \mathrm{d}\varphi =$$
(9.6)

$$= R_e^2 \left(\varphi_{\max} - \varphi_{\min}\right) \left(\sin \phi_{\min} - \sin \phi_{\max}\right), \tag{9.7}$$

where φ is the longitude, *min* and *max* denote the extent of the grid cell in terms of latitude and longitude (in radians), and $R_e = 6371$ km is the Earth radius. Since grid cells are rather coarse units that will cover coastal areas, the grid cell area is then multiplied by the land fraction within that grid cell. The land fraction $c^{(\text{land})}$ is calculated from the 300 m CCI Water Bodies v4.0 product⁴ (Lamarche et al. 2017). While static, the land-sea mask should still give a reasonable value for the amount of land present in a grid cell.

Combining all of the above, the total uptake in a grid cell is

$$\overline{\text{GPP}}_{(\text{total})} = \left(\frac{1}{N} \sum_{i=0}^{N} \overline{\text{GPP}}_{i}\right) \cdot A \cdot c^{(\text{land})} \cdot N_{\text{days}}.$$
(9.8)

The above formulation is independent on the grid cell size, however due to GOSAT's sampling pattern, a strong dependency on any total regional or global flux will be changing with the size of the grid cell. To show this, as well as a sanity check of the formulation itself, the global annual carbon uptake is calculated for several grid cell sizes, from 2.0° to 5.0° in 0.25° steps. For this annual uptake value, the time boundaries where chosen to be the first and the last day of each year between 2010 and 2016. In this calculation, every grid cell represents the annual mean flux derived from the annual mean SIF. Grid cells with less than three measurements are skipped in the global sum.

The grid cell dependence curve in Figure 9.10 is explained by looking at the percentage of grid cells which contain SIF measurements, compared to the total number of grid cells with some amount of land in it (excluding Antarctica). For small grid cells, there will always be a number of empty grid cells, in which no GOSAT measurement is

⁴https://esa-landcover-cci.org



Figure 9.10: Global mean annual uptake between 2010 and 2016 and the corresponding fractional cover of the globe in terms of the grid cells. Full coverage (100%) is obtained if all grid cells with a land fraction > 10% contain valid SIF measurements.

taken. This can be seen easily in e.g. Figure 7.12 (Page 83), where there are stripe-like features in the Sahara and the Amazon. For the calculation of the total global uptake, the $\overline{\text{GPP}}_{(\text{total})}$ for all grid cells are simply added up. Any empty grid cells then are implicitly counted as grid cells with 0 Pg C yr⁻¹, which results in a lower total global value. Between grid cell sizes of 3.5° and 5°, the total global flux hovers around the 125 Pg C yr⁻¹ mark. The shaded area represents the mean uncertainty for the global uptake, calculated as follows. For every grid cell, the uncertainty on the uptake is a combination of the systematic uncertainty of β and the uncertainty on the mean SIF:

$$\sigma_{\overline{\text{GPP}}} = \overline{\text{GPP}} \sqrt{\left(\frac{\sigma_{\overline{F_s}}}{\overline{F_s}}\right)^2 + \left(\frac{\sigma_{\beta}}{\beta}\right)^2}.$$
(9.9)

 $\sigma_{\overline{F_s}}$ is the standard deviation of all SIF values in the grid cell, rather than a regional uncertainty as described in Section 7.9. In this case, the grid cells are not large enough to warrant the use of regional subsets within the grid cell. The annual uncertainty is then simply the sum of the uncertainties in all grid cells.

However, not only the spatial aggregation, but also the temporal aggregation will impact the estimated annual flux. To gain a rough estimate on the effect on temporal aggregation, the annual fluxes were calculated for a grid cell size of 3.5° using four different aggregation periods: yearly, quarterly, monthly and semi-monthly.

Mean annual fluxes in Figure 9.11 range from ~94 to ~127 Pg C yr⁻¹. The right panel in that figure shows the mean percentage of covered land grid cells (like in Figure 9.10). Clearly, shorter aggregation periods result in lesser coverage, and finally in a lesser value for the annual total uptake. To overcome the issue of coverage, a sophisticated gap-filling mechanism would be needed to account for missing grid-cells over land. The large variability seen in the total global uptake is not due to the



Figure 9.11: Annual global fluxes and global land coverage using a grid cell size of 3.5° and four different aggregation intervals. The significant coverage drop in the year 2014 for aggregation intervals smaller than a year is caused by a gap in the GOSAT data stream between December 2014 and January 2015. The dashed lines in the left panel show variability in the result due to $\sigma_{\beta} = 0.229$.

uncertainty on either β , nor on the SIF measurement uncertainty, but rather due to the methodology and lack of gap-filling.

The estimated total global uptake is comparable to the numbers reported by Zhang et al. (2017) for the MODIS GPP data set (125.66 to 129.42 Pg C yr⁻¹). The SIF-optimised GPP in Parazoo et al. (2014) is stated as 124.77 Pg C yr⁻¹. On the other hand, Anav et al. (2015) list global annual mean GPP from various sources, which range between 112 (MODIS) to 169 Pg C yr⁻¹ (IPSL-CM5A-MR Earth system model). The value derived using flux tower measurements in the work of Beer et al. (2010) is 123 ± 8 Pg C yr⁻¹ for 2010. Integrating the fluxes directly from the CASA GPP data set yields a mean annual uptake of 128.29 ± 1.06 Pg C yr⁻¹, making the SIF-derived annual uptake consistent with this value as well.

Using a model to derive a SIF-GPP relationship results in a fairly stable conversion, since the global coverage of the model allows for many pairs to be examined like in Figure 9.4. The error on the scaling factor is small at roughly ~ 1 %.

9.1.6 Summary

To summarise this section, the CASA GPP CO_2 flux has been assessed against GOSAT SIF retrievals for the duration between April 2009 and December 2015. For non-JJA months, the SIF-GPP relationship agrees across various regions, whether they are low (e.g. northern boreal regions during DFJ) or high (most tropical regions) in terms of productivity. The departure that is seen during the JJA months and the

phase shift for tropical seasons observed in Figure 9.3 give reason to assume that the CASA model does not capture the seasonal cycle in the tropics well.

The same relationship was then derived through comparisons of GOSAT SIF measurements and flux tower measurements from the FluxNet network. The conversion factor β as calculated from FluxNet collocations was found to be ~30% lower than the continental-scale comparison using the CASA model. It was shown that this difference is explained as a purely statistical effect, caused by GOSAT's sampling sparsity, rather than an effect of collocation scale and biome.

Estimating the total annual global uptake from SIF results in a reasonable value of $\sim 125 \text{ Pg C yr}^{-1}$. This value, however, does not arise spontaneously from the SIF data itself, but only through conversion of the SIF measurements to fluxes using a conversion factor that was derived on continental scales. As such, it is not surprising that the SIF-estimated global-scale integrated flux resembles the value of the model itself.

9.2 2012 North American Drought

During the summer of 2012, large parts of the northern American continent experienced a drought. The drought itself has been studied in literature, several publications have highlighted various regional aspects, for example Mallya G. et al. (2013) investigated the drought in the US Midwest, and Hoerling et al. (2013) have explored in great detail the scale of the drought in the Great Plains, as well as their causes: reduced atmospheric moisture combined with the lack of summertime thunderstorms.

In an initial step to reproduce the anomaly in the SIF time records, the collocated soil moisture and precipitation data is analysed to confirm that the drought is indeed captured through the GOSAT sampling pattern. In Figure 9.12, the precipitation (TRMM 3B43) (Z. Duan et al. 2013) and soil moisture (Dorigo et al. 2017) (active measurements) anomalies for the period between May and August clearly show the continental-scale drought. The baseline for the anomalies are the years 2010 until 2016. 2009 data was skipped as the are no GOSAT soundings before April 2009.

The rainfall anomalies for 2012 are strongest east of the Rocky Mountains, covering the Great Plains between Nebraska and Oklahoma, as well as extending eastwards as far as Kentucky; additional patches of low precipitation are located near the Northern Rockies as well as northern Mexico. Soil moisture anomalies appear roughly in the same regions.

Roughly 23 000 considered GOSAT measurements (S-polarised) fall into the bounding box shown in Figure 9.12. The bounding box 100° W-85°W, 35° N-45°N defines the



Figure 9.12: Precipitation and soil moisture anomalies for 2012 against a 2010 to 2016 baseline (gridded to $2^{\circ} \times 2^{\circ}$ cells). TRMM precipitation data does not extend further than 50°N. The bounding box shows the region of interest onto which the data is sub-set for the following time-series analysis.

region of interest which will be referred to throughout this section. The selection criteria were very basic: the soundings had to pass the quality filter (Table 7.1), and 60% of the GOSAT footprint had to cover a non-bare surface type, therefore excluding mostly urban areas. The sampling density of the full time series is shown in Figure 9.13.





The time-series extracted from the collocated TRMM data set is shown in Figure 9.14. Periods between May 1 and August 31 are explicitly marked by vertical grey bars, and the corresponding parts of the time-series are highlighted as well. Representing the precipitation measurements this way, it is clear to see that summer 2012 is indeed an anomaly when compared to the other summers between 2009 and the end of 2016. By 2014, precipitation seems to reach pre-2011 levels again. Interestingly, precipitation during summer 2011 was similarly low, compared to the 8-year average.

A similar picture is established for the soil moisture measurements, displayed in



Figure 9.14: TRMM precipitation time-series. The measurements are re-sampled to 15 day intervals, and for every interval the mean (thick, blue line) and the standard deviation (shaded area) are computed. For times between May and August of each year, the values are highlighted. The horizontal, red-dashed line marks the mean (standard deviation again shown shaded) of all values between May and August for all years, thus representing the climatology for the region.

Figure 9.15. While the anomaly is indeed seen for passive measurements and the combined data set, the most striking signal is observed using active measurements. Here too, the pre-2012 levels are recovered by 2014.



Figure 9.15: This figure is designed in the same way as Figure 9.14. Summer 2012 exhibits anomalously low soil moisture for all three time series, but is best seen for the active measurements (middle panel).

As one final additional perspective, the standardised precipitation evapotranspiration index (SPEI) is considered, which, like the well-known Palmer drought severity index (PDSI), includes temperature data to characterise droughts based on supply and demand. Constructing a similar time-series, the 2012 drought is clearly identified with the mean curve exceeding -1 during the summer.



Figure 9.16: The standardised precipitation-evapotranspiration index shows a significant departure from the average during the 2012 summer.

Figures 9.12 to 9.16 clearly show that the continental-scale drought in northern America is captured well using various auxiliary data, taking into account the GOSAT sampling pattern. The next step of the analysis is to find similar anomalies within the SIF data. It was mentioned before in Section 7.7 that is not inherently clear which polarisation, which retrieval window, and which calibration method provides the best quality SIF data. Even though Figure 7.19 suggests that using the Spline-based calibration is the least biased set of retrievals in terms of the stability of the seasonally-dependent calibration (for absolute SIF), there is no straightforward way of validating such a claim.

The three polarisations and three correction procedures form an ensemble of nine data sets for both relative and absolute SIF, and the two retrieval windows each. The variability of these nine distinct time series is shown in Figure 9.17. In this figure, the solid line represents the ensemble median, and the shaded area is bounded by the minimum and maximum value for that particular time interval. Clearly, the variability between the nine time series is small, apart from the first few months in the 772 nm data. The mean of the deviation $\max(F_s) - \min(F_s)$ for the time series is $< 0.12 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ for absolute, and < 0.15% for relative SIF. Considering the small differences between the calibrations and polarisations, the further analysis is performed using only one of the time series (S-polarised, Spline-based correction). It is to be noted here, this result does not imply a deviation on aforementioned level on a per-sounding basis between polarisations, considering the large uncertainty on every SIF retrieval. However, if SIF retrievals are averaged, as is done for the time-series in this section, all nine derived SIF data sets become comparable.

Red lines in Figure 9.17 are placed at the maximum SIF for the period May to August 2012. For all time-series apart from the relative SIF at 772 nm, the years after 2012 exhibit a higher maximum during the same season, and lower values compared to 2010. This result is again visualised in Figure 9.18, where the means of the May-August period of each year are plotted, compared to the mean of the



Figure 9.17: Ensemble medians (solid line) and ensemble maxima and minima (shaded) show the relative consistency between polarisations and calibration methods. The unphysical spikes near data gaps are a result of data sparsity.

same period in 2012 (red, dashed line). In this figure, the drop in SIF is seen for all time-series - the strength of the signal, however, is not equally strong for all four modes.



Figure 9.18: Summer SIF differences ΔF_s compared to the 2012 level. All four data sets (S-polarised) show similar trends with small deviations. For all sets, 2009 and 2012 summers are low compared to post-2012 levels. For relative SIF at 755 nm, the summer average for 2009 was even lower than for 2012.

Using the conversion from SIF to carbon uptake discussed in Section 9.1, the SIF

anomaly during 2012 can be expressed as a flux, tabulated in Table 9.2. Uncertainties for SIF-derived uptake have been assessed through Equation 9.9, for CASA the uncertainties are the standard deviation divided by the square root of soundings, and for CASA[†], they are the standard deviation divided by the square root of the number of grid cells. Due to the way the three sets of data are accumulated, the uncertainties represent a different variability. For the SIF-derived GPP, the uncertainties of both the SIF retrieval and the conversion factor β are propagated through. The GOSATcollocated CASA data carries uncertainties due to regional and temporal sampling of the region of interest. CASA GPP taken directly from the gridded data shows higher variability because of the full coverage of the region.

	SIF-derived		CASA		CASA^\dagger	
Year	ΔGPP	∆GPP [%]	ΔGPP	∆GPP [%]	∆GPP	∆GPP [%]
2011	-91 ± 161	-4.7 ± 8.4	-208 ± 41	-11.7 ± 2.3	-142 ± 177	-9.1 ± 11.3
2012	-228 ± 149	-11.8 ± 7.8	-69 ± 46	-3.9 ± 2.6	-33 ± 194	-2.1 ± 12.4
2013	-25 ± 195	-1.3 ± 10.1	-198 ± 51	-11.1 ± 2.9	-183 ± 198	-11.7 ± 12.7
2014	-203 ± 190	-10.6 ± 9.9	-140 ± 46	-7.8 ± 2.6	-116 ± 179	-7.4 ± 11.4
2015	-123 ± 96	-6.4 ± 5.0	-69 ± 41	-3.8 ± 2.3	-10 ± 196	-0.7 ± 12.5
2016	69 ± 190	3.6 ± 9.9	n/a	n/a	n/a	n/a

Table 9.2: Mean annual SIF-derived and CASA-sampled fluxes as absolute and relative differences to 2010 levels. Absolute fluxes are in units of g C m⁻² yr⁻¹. The third column pair (CASA[†]) are the annual flux differences to 2010 using non-collocated model data, where the entire region of interest was averaged, not taking into account the GOSAT sampling.

The comparison between CASA and CASA[†] directly reveals the effect of GOSAT sampling. While the trends and relative uptake differences are comparable, the absolute values are lower for CASA sampled at GOSAT locations - the 2010 annual uptake is ~20% lower in CASA[†] than in CASA.

The drought year 2012 is outstandingly low with a value of $-228 \pm 149 \,\mathrm{g}\,\mathrm{C}\,\mathrm{m}^{-2}\,\mathrm{yr}^{-1}$ (-11.8 ± 7.8%) compared to 2010. Schwalm et al. 2012 have studied the (western) North American (25°-50°N, 100°-125°W) drought between 2000 and 2004. Their Figure 3 shows carbon fluxes against a 1997-2007 baseline (excluding the drought years), and separated into three different biomes. Summing up the anomalous GPP during the drought period for all three biomes equals to a flux difference of $308 \,\mathrm{g}\,\mathrm{C}\,\mathrm{m}^{-2}\,\mathrm{yr}^{-1}$, or a relative drop of about 13% compared to the baseline.

CASA and SIF tell a different story, however. Evaluating the fluxes for the entire year reveals that the annual uptake seen by SIF for the year 2012 is almost 12% lower than the value for 2010. Only 2012 and 2014 show this drastic reduction. CASA's annual fluxes, again as differences to the 2010 value, drop lower in 2011 and 2013, whereas 2012 shows a smaller decrease. Considering these annual fluxes, CASA and

SIF are out of phase, and interestingly, the relative differences align well if the SIF values are considered to be delayed by one year.

Wolf et al. (2016) have already studied the 2012 drought, and have used MODIS GPP in combination with 22 flux tower sites to assess the impact on the North American carbon cycle. The main conclusion of their work is that the warm spring in 2012 reduced the impact of the drought by compensating for the low productivity in the summer. The warm spring is visible in both SIF-derived GPP and CASA fluxes (Figure 9.19). Due to the increased productivity in the 2012 spring, the total annual fluxes are less indicative of a large drought event, when compared to the preceding or succeeding years.



Figure 9.19: Monthly mean gross uptake from CASA (left) along with the SIF-derived uptake values (right). Both CASA and SIF show the same warm spring in 2012.

GPP measured via flux towers exhibit a similar response to the drought. The three FluxNet sites US-PFa, US-Los and US-WCr are within 50 km of each other (north Wisconsin), the station US-UMB is roughly 400 km to the east of this cluster, and a fifth site, US-MMS, is found further south in central Indiana (about 600 km south of US-UMB). A map indicating the locations of these five sites is shown in Figure 9.20.

Despite being very close to each other, US-PFa and US-WCr show different behaviour during the 2012 drought, as seen in Figure 9.21. The seasonal amplitude of US-PFa is about half of US-WCr. US-Oho, US-UMB, US-WCr and US-MMS show very similar behaviour, both in the seasonal cycle, and the relative drop during summer 2012. The differences in the time series demonstrate very clearly that flux tower measurements are representative of the local situation, which might not be in line with the response of the biosphere on a sub-continental scale.

The upper panel of Figure 9.21 shows the relationship between monthly means of SIF aggregated over the entire region of interest, against monthly means of GPP measured by the various flux towers. The three stations US-Oho, US-MMS, US-UMB



Figure 9.20: The location of five flux tower sites, which have a comparatively long record of measurements. Three sites (US-PFa, US-Los and US-WCr, slightly outside the region of interest outlined in Figure 9.12) are clustered very closely together, the remaining three sites, while still within the region of interest, are several hundred kilometres to the east and south.

and US-WCr exhibit slopes that are very close to the continental-scale value derived from CASA (Figure 9.5), whereas the US-PFa and US-Los sites deviate significantly. The Lost Creek station (US-Los) is situated in permanent deciduous shrub wetlands, and the Park Falls site (US-PFa) is situated in a clearing, where the tower footprint "encompasses a highly heterogeneous landscape of upland forests and wetlands (forested and non-forested)"⁵. As such, the behaviour of the time series at those sites is explained due to the fluxes being dominated by wetlands rather than vegetation.

Aggregating the four sites more representative of the Corn Belt vegetation (US-WCr, US-UMB, US-Oho, US-MMS) into one time series is shown in Figure 9.22, along with the mean scaled SIF for the large region of interest. Both time series exhibit a smallest seasonal maximum in the 2012 summer, however the relative anomaly is much larger for the SIF (~16.8%) signal, than it is for the flux tower aggregate (~9%). The tower measurements, however, do not exhibit a flux reduction in the 2009 summer.

Overlaying the monthly mean fluxes in Figure 9.23 shows that the 2012 summer decreases in productivity similar to SIF and CASA. Although the spring onset in the drought year is less pronounced than in Figure 9.19, 2010 exhibits an early onset as well. Integrating the monthly mean fluxes into annual values reveals that 2012 had less total uptake compared to 2011.

Combining the mean annual fluxes (by integrating the monthly mean fluxes) from Table 9.2 and Figure 9.22 is shown in Figure 9.24. Since the various data have different

⁵http://sites.fluxdata.org/US-PFa/



Figure 9.21: FluxNet GPP from the six selected locations compared to scaled SIF for the entire region of interest. Both panels show monthly means.



Figure 9.22: Monthly aggregated time series of the four stations US-WCr, US-UMB, US-Oho, and US-MMS (blue) and the region-averaged scaled SIF. The dashed, red line is the mean of the annual maxima of the full FluxNet time series from 1999 onward (not shown).



Figure 9.23: Monthly mean aggregates of the four FluxNet sites from Figure 9.22, but overlaid to show the early spring onset in 2012.

baselines, that figure shows the departure from the 2010 annual uptake in relative terms. The four different time series have in common that 2010 was a year of high productivity, and the following years were less so. SIF-derived annual uptakes agree better with FluxNet measurements than with CASA model calculations, despite there being a larger difference between SIF and FluxNet for 2014. The discrepancy in 2013 is one between model and measurements mainly.



Figure 9.24: Relative fluxes compared to the 2010 baseline (see Table 9.2).

The assessment so far was directed towards gross uptake. Flux inversions based on space-based measurements, however, can provide information on net carbon fluxes. Through a collaboration with the University of Edinburgh, in particular with Liang Feng and Paul Palmer, in which they have provided flux inversions using the GEOS-Chem transport model with UoL-FP derived XCO_2 data⁶. This flux inversion, however, has one specific modification made to the prior fluxes. As shown in Figure 9.25 (left panel), the prior fluxes already exhibit the warm spring onset.

 $^{^6\}mathrm{Version}$ 7.1, with revised bias correction terms, compiled by Jasdeep An and (University of Leicester).

In order to see if the early onset is actually seen in the CO_2 concentrations, the prior was modified to use the 2011 values for the entire year of 2012. The results of the inversion (right panel of Figure 9.25) show that the inversion process shifts the prior flux towards the early onset that is seen in the SIF retrievals as well as the models, confirming that the retrieved CO_2 concentrations have already tracked this large-scale phenomenon.



Figure 9.25: Net fluxes as given by the CASA prior (left) and the inversion performed by the University of Edinburgh, where prior fluxes for 2012 were set to be those of 2011 - thus showing the early onset is not merely a result of the prior.

As mentioned before, the early warm onset in 2012 over North America has been studied already by Wolf et al. (2016), and in particular from a SIF perspective by Sun Ying et al. (2015). The point of view from space-based observation of carbon fluxes, however, has not been studied to the thesis author's knowledge. Figure 9.25 already hints towards GOSAT having picked up this anomalous event in the North American carbon cycle.

Summary

N Part II of this thesis, I performed the retrieval of chlorophyll fluorescence from GOSAT measurements and analysed the results. While the retrieval scheme and the principle behind the bias correction was done following Frankenberg, Fisher, et al. (2011), I introduced a new correction method. The Spline-based correction explicitly takes the time-dependence of the bias into account. Additionally, I made use of collocated land cover data on a sub-pixel level, which allows regions other than Antarctica and Greenland to be used as reference soundings (Section 7.7). The retrieval results themselves exhibit a high anti-correlation between the retrieved reflectance and the retrieved fluorescence. At first glance, this suggests that the underlying principle of the fluorescence retrieval is not viable, which asserts a decoupling of reflectance from fluorescence. I conducted simulations and retrievals using simulated radiances to confirm that there is indeed a strong coupling between SIF and reflectance. The error due to this coupling, however, is on the order of 1% of the SIF radiances, and therefore has no significant effect in the case of GOSAT retrievals (Section 7.6). I compared the retrievals from GOSAT to SIF from GOME-2 (Section 7.8). The comparisons yield similar results as published by Köhler et al. (2015) and are in line with the expectation, considering the different spectral windows and the different overpass times. Uncertainty assessment is difficult, as there is no truth data available, apart from those regions in the world where zero fluorescence can safely be assumed (Antarctica, most of Greenland). I then introduced an approach based on regional clustering (Section 7.9) which allows for a more realistic value of the regional uncertainty to be calculated, compared to simply taking the standard deviation of SIF values of a large region.

SIF radiances have an impact on XCO_2 retrievals. Ignoring these contributions to the O_2 A-band results in a change in other quantities, such as surface pressure and aerosols. I retrieved four full years of GOSAT XCO_2 both with and without including the SIF radiances as priors. The differences in the retrieved CO_2 column are significant and can reach 1 ppm for some regions and seasons. Considering time series, the seasonal amplitudes are modified between 4 to 10%, with the South American Tropics being a significant outlier with almost 19% (Chapter 8). Finally, I analyse the SIF data itself to relate it to gross carbon uptake by vegetation. Even though the CASA model aligns well with the GOSAT SIF retrievals when looking at time series, some regions exhibit strong differences in the seasonal cycle. I attribute this to CASA not producing the right cycle in the tropics. Nevertheless, I established a global linear relationship between uptake and SIF through CASA (Section 9.1), which turned out to be comparable to already published values. Comparisons to flux tower sites, however, show a different scaling factor between SIF and fluxes. I explain this discrepancy through the sparse GOSAT sampling. Through a case study about the 2012 continental-scale drought in North America, I show that a large disruption of the photosynthetic activity can be tracked through the SIF time series over that region (Section 9.2).

Part III



FAST RADIATIVE TRANSFER FOR XCO₂ Retrievals

Chapter Ten

Atmospheric Radiative Transfer



OR most remote sensing applications in the Earth sciences, atmospheric radiative transfer is a fundamental and time consuming portion of the forward model. Briefly summarised, radiative transfer (RT) in this context describes the propagation of light through the atmosphere, and includes the various interactions of light with the constituents of the atmosphere as well as the surface.

In full-physics XCO₂ retrievals, RT produces radiances based on the model atmosphere and its properties (see Section 4.2, Page 34), which can then be used by the inverse method to fit the state vector elements to minimise the difference between measurement and model. In this part of the thesis I first introduce the radiative transfer equation (RTE) and common approaches to solve it. Then, I discuss various strategies to accelerate those calculations. A major focus of this thesis is the novel PCA-based approach, which I implemented and advanced to fully incorporate the spectral dependencies of aerosol scattering properties. The PCA-based approach has been used on retrievals from the NASA OCO-2 mission to demonstrate the advantage over other methods.

The Radiative Transfer Equation 10.1

In contemporary literature, Lommel (1889) is often cited to have written the first modern treatise on radiative transfer. He was motivated by experiments which showed that the application of Lambert's cosine law (Lambert 1760) alone does not agree with the measurements for diffusely reflecting surfaces, which had significant impact on the photometry of planets. Lommel's derivation is phenomenological, and describes the attenuation of light rays through scattering and absorption being linearly related to the incident ray intensity and the thickness of the medium through which it propagates. More than 120 years later, this way of introducing the RT equation is still often followed, as it is both compact and illustrative.

The book of Mishchenko et al. (2006) (Chapter 8) provides, amongst other related matters, a highly comprehensive treatment of the RTE. A full microscopic explanation of the equation is given. In their work, they make it very clear that the

phenomenological explanation relies on postulates - the aforementioned linearity between the attenuation and incident light intensity is never derived, even though it can be experimentally demonstrated. Mishchenko et al. (2006) derive the RT equation essentially from first principles - the Maxwell equations. The fundamental objects are the electromagnetic fields rather than intensities. Certain phenomenological postulates can be directly related to the corresponding phenomena that arise in the microphysical treatment. In this section, the RTE is introduced via the simpler, phenomenological approach.

The object of interest in RT is called *specific intensity* (or *intensity*), which upon integration along the wavelength, solid angle, time and area dimensions, yields the amount of transferred energy. Light energy is considered to travel along straight lines, which puts RT into the context of geometrical optics. Again, a note should be echoed from Mishchenko et al. (2006). Radiative transfer is not inherently tied to the domain of geometrical optics, however when applied to the atmosphere, a number of assumptions can be made which justify that simplification. Incident light coming from the sun is naturally incoherent and one can further assume that no other phenomena inside the atmosphere can give rise to coherence. As such, different contributions can be simply added incoherently without taking into account a potential relative phase between each contribution. Using Stokes calculus, a light ray of arbitrary polarisation state is represented through a real four-component Stokes vector:

$$\mathbf{S} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I \\ IP \cos 2\psi \cos 2\chi \\ IP \sin 2\psi \cos 2\chi \\ IP \sin 2\chi \end{pmatrix},$$
(10.1)

where each component of **S** is incoherent. ψ here is the orientation of the polarisation ellipse, $\tan 2\psi = U/Q$, and χ is the so-called ellipticity angle, where $\tan 2\chi = V^2/(\sqrt{Q^2 + U^2})$. *P* is known as the degree of polarisation, $P^2 = (Q^2 + U^2 + V^2)/I^2$, which for natural, unpolarised light is zero. *I* is the intensity of the light ray, and the three components *Q*, *U* and *V* describe the specific state of polarisation.

Chandrasekhar (1960) uses I rather than S to denote the Stokes vector. Note that the variable I is now used as general radiance and the distinction between solar irradiance and radiance is not made, as in Equation 7.9 (Page 73).

Superposing multiple light rays is then achieved by simply adding the various components $S^{(i)}$, $Q^{(i)}$, $U^{(i)}$ and $V^{(i)}$ independently to the total Stokes vector **S**. Light rays as well as each component of **S** are transversal, which again is not a general case. As Mishchenko et al. (2006) point out, the superposition between an incident transversal electromagnetic wave with a reflected spherical wave is neither transversal, nor does it have a specific direction of propagation. Coming back to the specific intensity, it is related to the transported energy in the following way:

$$d^{4}E(\lambda) = I(\mathbf{r}, \lambda) \cos \theta \, d\lambda \, d\sigma \, d\Omega \, dt, \qquad (10.2)$$

which describes the radiant energy *E* that passes through a surface element of area $d\sigma$ in a wavelength interval between λ and $\lambda + d\lambda$ that is emitted towards a direction of the unit vector **r** within a solid angle $d\Omega$ during the time dt. The angle θ is the angle between **r** and the surface normal of $d\sigma$.

Textbooks like Chandrasekhar (1960) or Thomas et al. (2002) introduce the phenomenological RTE in a similar manner: one considers a light ray passing through a homogeneous optical medium. The Stokes vector $\mathbf{S}(\mathbf{r}, \lambda)$ is then modified by an amount $d\mathbf{S}(\mathbf{r}, \lambda)$ while the ray is propagating an infinitesimal distance ds along the direction \mathbf{r} . Due to the incoherent nature of the considered light rays, the various contributions can be summed up, and the resulting RTE in its differential form is given by

$$d\mathbf{S}(\mathbf{r},\lambda) = \underbrace{-\mathbf{K}(\mathbf{r},\lambda) \mathbf{S}(\mathbf{r},\lambda) ds}_{\text{Extinction}} + \underbrace{\int_{4\pi} \mathbf{Z}(\mathbf{r},\mathbf{r}',\lambda) \mathbf{S}(\mathbf{r}',\lambda) d\Omega'}_{\text{Emission}}.$$
 (10.3)

Equation 10.3 is called the *vector radiative transfer equation*, referring to the fact that it deals with the full Stokes vector quantity, rather than just the intensity. The first term describes the change in the Stokes vector due to extinction in the medium. As already described by Lommel (1889), the ray is reduced in intensity by an amount that is proportional to the incident ray $\mathbf{S}(\mathbf{r}, \lambda)$ and the travelled distance ds. The proportionality is given by the *extinction matrix* $\mathbf{K}(\mathbf{r}, \lambda)$, which is a four-by-four matrix. The second term describes any emission that contributes to the radiation field into the direction \mathbf{r} . In Equation 10.3, thermal emission is not accounted for - in this thesis, all considered spectral ranges are below $2 \,\mu$ m wavelength, for which thermal radiation can be neglected. Depending on the investigated scenario, more contributions may be added which count as emission, such as fluorescence. The term *emission* in Equation 10.3 can lead to confusion, as this is distinct from thermal emission. Thomas et al. (2002) as well as other literature use this term to describe any contribution that increases the specific intensity pointing towards \mathbf{r} .

The quantity $\mathbf{Z}(\mathbf{r}, \mathbf{r}', \lambda)$ is called the *phase matrix*. It relates an incoming Stokes vector, coming from direction \mathbf{r}' to an outgoing one travelling towards direction \mathbf{r} . To account for the contributions to $\mathbf{S}(\mathbf{r}, \lambda)$ due to scattering, all possible incoming directions \mathbf{r}' have to be considered, hence why the second term is an integral over the full 4π solid angle $d\Omega'$, and the weighting is given by the phase matrix.

The emission term adds a significant amount of complexity to the RTE, so to demonstrate a simpler case, let's first consider a scenario without it. The vector RTE then simplifies to

$$\frac{\mathrm{d}\mathbf{S}(\mathbf{r},s,\lambda)}{\mathbf{S}(\mathbf{r},s,\lambda)} = -\mathbf{K}(\mathbf{r},\lambda)\,\mathrm{d}s\,,\tag{10.4}$$

where the location vector **x** has been replaced by the scalar path parameter s (**x** = $s \cdot \mathbf{r}$).

The solution to Equation 10.4 is trivially obtained through elementary integration, with the integration boundaries being s_0 and s:

$$\mathbf{S}(\mathbf{r}, s, \lambda) = \mathbf{S}(\mathbf{r}, s_0, \lambda) \cdot \exp\left(-\int_{s_0}^{s} \mathbf{K}(\mathbf{r}, \lambda) \,\mathrm{d}s'\right). \tag{10.5}$$

Unsurprisingly, this is exactly the Beer-Lambert law, and very intuitively replicates the exponential decrease of intensity in an optical medium in which extinction occurs. For the cases that are encountered in this thesis, $\mathbf{K}(\mathbf{r}, \lambda)$ is independent of the location within the medium - the medium is assumed to be homogeneous.

Extinction is caused by absorption or scattering, thus $\mathbf{K}(\mathbf{r}, \lambda)$ can be further decomposed into these two contributions:

$$\mathbf{K}(\mathbf{r},\lambda) = \mathbf{A}(\lambda) + \int_{4\pi} \mathbf{Z}(\mathbf{r},\mathbf{r}',\lambda) \,\mathrm{d}\Omega' \,. \tag{10.6}$$

 $\mathbf{A}(\lambda)$ is the absorption coefficient matrix, and can generally be assumed to be of diagonal form. Absorption only affects the intensity of a light ray, but does not change its state of polarisation. Scattering, on the other hand, may change the polarisation state of a ray. The second term in Equation 10.6 is similar to the last term in Equation 10.3, and represents the fraction of intensity that is *not* scattered away but still contributes to the Stokes field directed towards **r**.

The emission term in Equation 10.3,

$$\int_{4\pi} \mathbf{Z}(\mathbf{r}, \mathbf{r}', \lambda) \, \mathbf{I}(\mathbf{r}', \lambda) \, \mathrm{d}\Omega' \,, \tag{10.7}$$

is the implementation of the phenomenon called *multiple scattering*. As mentioned before, the integration is performed for all incident directions \mathbf{r}' where the phase matrix is essentially the weighting function which determines which fraction of the incident light is scattered towards direction \mathbf{r} . Simply worded: to calculate the Stokes vector pointing towards direction \mathbf{r} requires knowledge of the entire Stokes field. The appearance of the Stokes vector $\mathbf{S}(\mathbf{r}', \lambda)$ turns the RTE into an integro-differential equation. So far, there are no general solutions to the full RTE.

10.2 Solving the Scalar RT Equation

Since general solutions to the full RTE do not exist so far, it is common practice to reduce the RTE into less general forms and solve the simpler problem using very

specific boundary conditions. In this section, the steps towards solving the RTE will be done for a special case that can be applied to remote sensing. For the vast majority of remote sensing applications, RT calculations are reduced to one dimension: the position vector \mathbf{x} is replaced by a scalar position *s* to represent distance along the line of sight. This simplification significantly reduces the complexity of the problem already.

Considering a homogeneous and isotropic plane-parallel optical medium through which a light ray travels at an angle θ (measured towards the surface normal), the RTE in one dimension can be rewritten as follows:

$$\cos\theta \frac{\mathrm{d}}{\mathrm{d}s} \mathbf{S}(\theta,\phi,s) = -\mathbf{K} \, \mathbf{S}(\theta,\phi,s) + \int_{4\pi} \mathbf{Z}(\theta,\phi,\theta',\phi') \, \mathbf{S}(\theta',\phi',s) \, \mathrm{d}\Omega', \qquad (10.8)$$

where ϕ is the azimuthal angle of the ray. Compared to Equation 10.3, the extinction matrix **K** is no longer dependent on either position or direction due to the assumption of a stratified and homogeneous medium.

Solving the 1D vector RTE (Equation 10.8) numerically is still a computationally very expensive procedure, and not always necessary. While TANSO-FTS on GOSAT measures two linearly polarised radiances (P, S), exploiting its polarisation information is not common practice despite studies showing promising results (Kikuchi et al. 2016). They utilised a fully polarised BRDF and retrieved both P and S polarised radiances simultaneously. Their retrieval simulations show that errors in XCO_2 are reduced significantly for scenes with higher aerosol loadings. However, the measured radiances can be averaged to yield unpolarised radiances with loss of the polarisation information. For OCO-2, accounting for polarisation is necessary because the instrument only measures one polarisation direction, meaning that at least the first three components of the Stokes vector need to be computed.

Rather than tackling the full, polarised vector RTE, one strategy is to only solve the intensity component of the equation. When averaging the radiances from GOSAT/TANSO-FTS measurements, the intensity is sufficient. Extracting the first row of Equation 10.8 yields

$$\cos\theta \frac{\mathrm{d}}{\mathrm{d}s}I(\theta,\phi,s) = -K_{11}I(\theta,\phi,s) + \int_{4\pi} Z_{11}(\theta,\phi,\theta',\phi')I(\theta',\phi',s)\,\mathrm{d}\Omega'\,. \tag{10.9}$$

Note that in the above equation, the contributions from the phase matrix to the intensity component, through Z_{12} , Z_{13} and Z_{14} , have been dropped. This is equivalent to the notion of ignoring polarisation in the model, rather than just not calculating the Q, U, and V components.

In the next step, the equation is transformed into optical depth space. Optical depth, τ , is a quantity that arises from Equation 10.4, as the integral of the extinction

coefficient (or matrix) along the path:

$$\tau \equiv \int_{s_0}^s k_{\text{ext}} \, \mathrm{d}s' \,. \tag{10.10}$$

Using $d\tau = k_{ext} ds$ and $K_{11} = k_{ext}$, and with $\mu = \cos \theta$, the scalar RTE becomes

$$\mu \frac{\mathrm{d}}{\mathrm{d}\tau} I(\theta, \phi, \tau) = -I(\theta, \phi, \tau) + \frac{\varpi}{4\pi} \int_{4\pi} P(\theta, \phi, \theta', \phi') I(\theta', \phi', \tau) \,\mathrm{d}\Omega' \,. \tag{10.11}$$

The quantity ϖ is called *single-scattering albedo*, and is related to both absorption and scattering. Earlier (Equation 10.6) it was mentioned that extinction (and the corresponding coefficients) is fundamentally decomposed into a sum of two extinction sources: absorption and scattering, such that $k_{\text{ext}} = k_{\text{abs}} + k_{\text{sca}}$. When rewriting the RTE with the optical depth τ as the free coordinate, the (11)-component of the phase matrix **Z** can be written as $Z_{11}(\theta, \phi, \theta', \phi') = k_{\text{sca}}P(\theta, \phi, \theta', \phi')$, *P* being the scalar phase function. The single-scattering albedo ϖ is defined as

$$\varpi = \frac{k_{\rm sca}}{k_{\rm ext}} = \frac{k_{\rm sca}}{k_{\rm sca} + k_{\rm abs}}.$$
(10.12)

From a phenomenological point of view, ϖ describes the fraction of extinction that is due to scattering and will depend on the optical medium and the material properties of the scattering constituents itself.

An established method of solving the RTE, and the one utilised in the LIDORT and TWOSTR models, is called *discrete ordinates method*. Its basic premise is outlined in Appendix A, Page 213. A slightly different approach is called *successive orders of scattering*. Starting from the scalar RTE (Equation 10.11), the intensity can be represented as a power series in ϖ . This expansion is viable due to the single-scatter albedo ϖ being less than 1 by definition:

$$I(\theta, \phi, \tau) = \sum_{k=0}^{\infty} \varpi^k I^{(k)}(\theta, \phi, \tau).$$
(10.13)

Substituting the expression for the intensity from Equation 10.13 into and ordering the terms by powers of ϖ gives (Mobley 2018)

$$\begin{bmatrix} \mu \frac{d}{d\tau} I^{(0)}(\theta, \phi, \tau) + I^{(0)}(\theta, \phi, \tau) \end{bmatrix}$$

+ $\varpi \begin{bmatrix} \mu \frac{d}{d\tau} I^{(1)}(\theta, \phi, \tau) + I^{(1)}(\theta, \phi, \tau) - \int_{4\pi} I^{(0)}(\theta', \phi', \tau) d\Omega' P(\theta, \phi, \theta', \phi') \end{bmatrix}$
+ $\varpi^2 \begin{bmatrix} \mu \frac{d}{d\tau} I^{(2)}(\theta, \phi, \tau) + I^{(2)}(\theta, \phi, \tau) - \int_{4\pi} I^{(1)}(\theta', \phi', \tau) d\Omega' P(\theta, \phi, \theta', \phi') \end{bmatrix}$ (10.14)
:
+ $\varpi^N \begin{bmatrix} \mu \frac{d}{d\tau} I^{(N)}(\theta, \phi, \tau) + I^{(N)}(\theta, \phi, \tau) - \int_{4\pi} I^{(N-1)}(\theta', \phi', \tau) d\Omega' P(\theta, \phi, \theta', \phi') \end{bmatrix}$
= 0,

where $N \to \infty$. Phenomenologically, the terms $I^{(k)}(\theta, \phi, \tau)$ can be viewed as different contributions to the intensity, which have undergone k scattering events. Solving the above equation is performed by considering a finite number in the expansion series, e.g. N = 2. Since each bracket in Equation 10.14 has to be zero for the entire equation to be balanced, the terms $I^{(k)}(\theta, \phi, \tau)$ can be successively solved for each k, starting with k = 0 and the increasing until k = N. The k^{th} (for k > 0) equation depends on the intensity of order k - 1, hence why the successive solving of the equations is necessary.

10.3 Implementations and Performance Considerations

Implementing a full radiative transfer solver is a tremendous effort in itself, hence why a number of ready-to-use software packages exist and are very popular in the remote sensing community. For the UoL full-physics algorithm, several RT solvers are used: LIDORT¹ (Spurr et al. 2001), TWOSTR² (Spurr et al. 2011) and 2OS³ (Natraj et al. 2007).

LIDORT is a general-purpose code and solves the 1D scalar RTE using the discrete ordinates method, and adds several features that have not been discussed in the previous section. The LIDORT model can be used in pseudo-spherical approximation mode, in which for the incoming solar beam, the spherical curvature of Earth and the atmosphere is considered, rather than relying on the assumption of plane-parallel layers. In terms of surface models, a linear combination of up to three BRDFs can be used to represent the reflective properties of the surface. The number of stream directions (or streams) can be freely chosen in LIDORT.

TWOSTR (or 2Stream) is similar to LIDORT, as it is another implementation of the discrete ordinates method for the 1D scalar RTE, with the main difference being that the number of streams is n = 2 - one up-welling, and one down-welling stream. In the two-stream approximation, several steps in the solution finding procedure can be replaced by analytic solutions. Using TWOSTR is generally much faster than the same model calculations with LIDORT using n = 2. Whereas LIDORT computes the full intensity field, TWOSTR calculates the multiple-scattering contributions only.

2OS is a fast, linearised two orders of scattering model that computes the full Stokes vector and returns the single-scattering results separately.

Using a full vector RTE implementation like VLIDORT⁴ (the vector version of LIDORT) is computationally much more expensive then the scalar version. In the

¹LInearized Discrete Ordinates Radiative Rransfer

²**TWO STR**eam

³2 Orders of Scattering

⁴See Spurr (2008).

UoL-FP algorithm, the models can be combined to calculate the total Stokes vector, as laid out in Natraj et al. (2008):

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I_{\text{scat}} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} I_{\text{corr}} \\ Q_{2\text{OS}} \\ U_{2\text{OS}} \\ V_{2\text{OS}} \end{pmatrix}$$
(10.15)

 I_{scat} is the intensity calculated by LIDORT, not taking into account polarisation. The output of the 2OS model is then added to the scalar calculation. I_{corr} is the correction to the intensity due to the mixing of components induced by the \mathbf{P}_{1i} and \mathbf{P}_{i1} elements (i = 2...4). For n = 2 streams, the faster TWOSTR code is used to calculate the multiple-scattering contributions, and the single-scatter contributions are obtained through 2OS. In that case:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I_{\rm MS} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} I_{\rm SS} + I_{\rm corr} \\ Q_{2OS} \\ U_{2OS} \\ V_{2OS} \end{pmatrix}.$$
 (10.16)
LIDORT, TWOSTR 2OS

If polarisation is not considered at all, the I_{corr} contribution is neglected and the total intensity is given by

$$I = I_{\rm SS} + I_{\rm MS}.$$
 (10.17)

The aforementioned implementations are monochromatic, pure scattering models meaning that the only input apart from the viewing geometry and solar geometry, are the layer-resolved total optical properties for a given wavelength: optical depths, single-scatter albedos and phase function expansion coefficients. The models are blind to the various contributions to the optical properties.

Another crucial feature of the used models are their Jacobian capabilities. For retrieval problems, it is not only necessary to calculate the intensities (or Stokes vectors) as part of the forward model, but also the derivatives with respect to the optical property inputs. Without Jacobian capabilities, layer-dependent derivatives can be calculated through finite differencing. This method, however, is much slower - the models above return the derivatives using one single call.

Chapter Eleven

Fast Radiative Transfer Methods

11.1 General Overview and Motivation



EDUCING the overall computation time for RT calculations is the main objective of fast radiative transfer (fast RT). The necessity of such techniques arises due to either the forward model calculations taking long, or the number of retrievals being too large, or a combination of the two. RT calculations can easily be the single most computationally expensive portion in a retrieval, hence why optimising this particular portion for speed will significantly decrease the overall computation time for a retrieval. Without any acceleration technique, a typical UoL-FP iteration, including the calculation of Jacobians (but without polarisation), takes about 45 min.

From the 2009-launched GOSAT to the 2014-launched OCO-2, the number of measurements has increased 96-fold. Future missions, such as the geo-stationary GeoCarb (Polonsky et al. 2014), will potentially increase the number of measurements by another order of magnitude. Given the finite amount of computational resources available, fast RT is a necessary technique to make global and operational retrievals feasible.

When considering the accuracy requirement for the XCO_2 retrievals, which is stated as $\sim 0.3\%$ in Crisp et al. (2004), the RT computations have to be accurate to $\sim 0.1\%$ in order to still allow for errors from other components of the retrieval (O'Dell 2010). This puts another stringent requirement on fast RT methods - rather than merely speeding up the RT calculations, they also need to retain the accuracy. Chevallier et al. (2016) states even lower values on the required accuracy¹ with < 0.2 ppm, which is less than 0.1%.

11.1.1 Correlated k-distribution Method

A historically popular and arguably one of the first methods is the so-called *correlated k*-distribution technique, first described by Ambartzumian (1936). Like the methods

¹In Chevallier et al. (2016), the term "systematic error" is used instead of accuracy.

that will be discussed in the following sections, it belongs to the class of *band methods*, which are designed to calculate top-of-atmosphere intensities for an entire spectral band. Assuming a non-scattering atmosphere, the Beer-Lambert law can be used to write the TOA intensity as a function of the wavelength-dependent total optical depth $\tau(\lambda)$ or the equivalent formulation through the extinction coefficient k and optical path s, and the incident solar irradiance I_0^2 :

$$I = I[\tau(\lambda)] = I_0 \exp\left[-\tau(\lambda)\right] = I_0 \exp\left[-k(\lambda)s\right].$$
(11.1)

The above formulation can be reordered to make the TOA intensity independent of the incident solar irradiance and thus leads to the transmission \mathcal{T} written as:

$$\mathcal{T}(\lambda) = \frac{I[\tau(\lambda)]}{I_0} = \exp\left[-\tau(\lambda)\right],\tag{11.2}$$

and the corresponding *absorptance*

$$\alpha(\lambda) = 1 - \mathcal{T}(\lambda). \tag{11.3}$$

In a wavelength interval $\Delta \lambda = \lambda_1 - \lambda_2$, the transmission for a path *s* is obtained by integrating $\mathcal{T}(\lambda)$ over that interval:

$$\langle \mathcal{T}(s) \rangle = \frac{1}{\Delta \lambda} \int_{\lambda_1}^{\lambda_2} \exp\left[-k(\lambda)s\right] \mathrm{d}\lambda$$
 (11.4)

This calculation of the average transmission for an entire band is thus performed by repeatedly calculating Equation 11.4 for different $\Delta \lambda$ for which $k(\lambda)$ is close to constant. Additionally, the calculation has to be separated for the various parts of the optical path, pressures and temperatures, which all influence the extinction coefficient. The ansatz for the *k*-distribution method is to rearrange the extinction coefficients $k(\lambda)$ in increasing order. Since many spectral points will share the same $k(\lambda)$, like points in the continuum, or line centres, this redundancy can be exploited to reduce the number of calculations. The distribution of extinction coefficients is defined as

$$f(k) \equiv \frac{1}{\Delta \lambda} \sum_{i=0}^{N} \frac{\mathrm{d}\lambda}{\mathrm{d}k} W_i(k), \qquad (11.5)$$

where $W_i(k)$ is a window function that is equal to 1 when $k_{i,\min} \le k \le k_{i,\max}$ and otherwise 0. Using this distribution, the average transmission is rewritten as a sum rather than an integral:

$$\langle \mathcal{T}(s) \rangle \sim \sum_{i=1}^{N} \Delta k_i f(k_i) \exp\left[-k_i s\right].$$
 (11.6)

²The explanation in Thomas et al. (2002) is followed here.

The cumulative distribution function

$$g(k_M) \equiv \sum_{i=0}^{M} \Delta k_i f(k_i)$$
(11.7)

can now be used to rewrite the transmission as

$$\langle \mathcal{T}(s) \rangle \sim \sum_{i=1}^{N} \Delta g_i \exp\left[-k_i s\right] = \Delta g_1 e^{-k_1 s} + \Delta g_2 e^{-k_2 s} + \dots + \Delta g_N e^{-k_N s}.$$
(11.8)

If the *k*-distribution is known (through look up tables of laboratory measurements, or simulations), then Equation 11.8 can be calculated in a straightforward way.

While the fast RT methods presented in the next sections take a different approach and do not operate on the transmission quantity, the starting principle is the same. One could say that the *correlated-k* method belongs to a more general class of (fast) RT schemes that perform *spectral binning* of some variety. The main idea behind the following methods is to group spectral points which share similar optical depth, and then perform a small number of RT calculations that are representative for this group to then map the results back to each spectral point individually. Since some spectral groups can contain thousands of spectral points, the reduction in the number of RT computations can easily be more than 95 %.

The following fast RT methods have been developed for XCO_2 retrievals in the SWIR part of the spectrum. While one cannot assume that these techniques work equally well for other spectral ranges, there is nothing inherently restricting these methods to the SWIR. These methods have also in common that they do not change the underlying RT model calculations itself, but rather reduce the number of needed calculations. This makes the methods theoretically independent of the chosen RT models, however no such intercomparison has been done yet.

11.2 Low-Streams Interpolation

The Low-Streams Interpolation (LSI) technique was introduced by O'Dell (2010) to accelerate retrievals from GOSAT and OCO measurements. LSI is motivated by the same fact as the *correlated-k* method: many of the tens of thousands of spectral points will share similar total gas optical depths. Expanding on the idea, the difference between a high-accuracy and a low-accuracy RT calculation can be analysed. When talking about accuracy in the context of the discrete ordinates-based RT models, the term is synonymous with the number of quadrature streams $N_{\rm s}$.

Using one sample scenario for the TANSO-FTS instrument, the high-resolution intensities (no instrument line shape is applied) are calculated using $N_s = 2$ and

 $N_{\rm s} = 16$, and their point-by-point differences $\Delta I = I_{\rm low} - I_{\rm high}$ are shown in Figure 11.1 for each of the three bands. While the left panels show the intensity differences as a function of wavelength, the right panels show them as histograms and functions of the total optical depth due to gas absorption $\tau^{(\text{gas})}$. $\tau^{(\text{gas})}$ is simply calculated as the sum of all layer gas optical depths. When the ΔI are displayed as a function of total gas optical depth, they become smoothly varying, compared to the usual wavelength representation.



(c) Strong CO_2 band.

Figure 11.1: Intensity differences $\Delta I = I_{\text{low}} - I_{\text{high}}$ as functions of wavelength and total gas optical depth. The right panels are a histogram rather than point plots, and the point density is indicated by the hue (dark green = low, light yellow = high). Residuals for the right column are in the same units as on the left.

In Figure 11.1 lies the motivation of the LSI method: if the difference between a high- and a low-accuracy RT calculation is a slowly varying function of τ_{gas} , then coarsely sampling that function can be used to reconstruct a high-accuracy result from a low-accuracy calculation. A closer look at Figure 11.1c (and to a lesser extent Figure 11.1b) reveals that the this error function $\Delta I(\tau^{(\text{gas})})$ might take a smooth, but intricate form. The bifurcation seen for the strong CO₂ band is the result of two different absorbers, namely CO₂ and H₂O. However, the reason for the bifurcation

is not due to the gas species itself, but rather because of their vertical profiles. As $\tau_{\rm gas}$ is a sum of layered optical depths, profiles with different shapes can end up having the same total optical depth, despite their shapes being very different. A further complication is the fact that scattering layers can vary additionally to the gas absorption profiles.

The LSI technique makes use of a quantity introduced by M. Duan et al. (2005) to account for the different vertical profiles. Defining a new unit-less quantity

$$\xi = \frac{\tau^{(\text{gas})\prime}}{\tau^{(\text{gas})}},\tag{11.9}$$

where $\tau^{(\text{gas})'}$ is the gas optical depth of the column, integrated only up to the point where the cumulative gas optical depth equals half the total-column optical depth due to scattering (integrated from the top of the atmosphere downward). Figure 11.2 shows the error function (the same as in Figure 11.1c), but rather than coloured by density, the colour now represents the value of $\sqrt{\xi}$. It clearly shows that the two branches differ significantly in terms of their vertical profiles and their scattering behaviour.



Figure 11.2: The same intensity differences as Figure 11.1c, but colours representing $\sqrt{\xi}$. For larger values of $\sqrt{\xi}$, the spectral points lie mostly on the upper branch.

Based on Figure 11.1 and Figure 11.2, the LSI strategy can be formulated:

- 1. Every spectral point is assigned a bin in $\tau^{(\text{gas})}$ -space as well as one of two bins in $\sqrt{\xi}$ -space. Bins in $\tau^{(\text{gas})}$ -space are chosen to sample the error curves in Figure 11.1 properly and depend on the band. The bins for $\sqrt{\xi}$ are chosen such that the lowest 25% are assigned one bin, the rest of spectral points are assigned the other bin.
- 2. For every bin in $\tau^{(\text{gas})}$ - $\sqrt{\xi}$ -space, optical properties are constructed that are representative of the current bin. Binned gas optical depths are calculated as the average of all profiles in the bin, whereas scattering optical depths, single-scattering albedo as well as phase function expansion coefficients are taken from the band centre.

3. Given the binned profiles, errors between high- and low-accuracy RT model calculations are computed for a given Stokes component *S*:

$$\varepsilon_{S}(i,j) = \frac{S_{\text{low}}(i,j) - S_{\text{high}}(i,j)}{I_{\text{high}}(i,j)},$$
(11.10)

where *i* and *j* are the indices in τ_{gas} and $\sqrt{\xi}$ -space respectively signifying the bin. Using a bilinear interpolation scheme³, the error $\varepsilon_S(\tau^{(\text{gas})}, \sqrt{\xi})$ for any specific value of $\tau^{(\text{gas})}$ and $\sqrt{\xi}$ (and implicitly λ) is obtained from the errors related to the bin.

4. The approximation to the high-accuracy result on a point-by-point basis is finally calculated through the errors $\varepsilon_{S}(\lambda)$:

$$I(\lambda) = \frac{I_{\text{low}}(\lambda)}{1 + \varepsilon_I(\lambda)},$$
(11.11)

and

$$S(\lambda) = S_{\text{low}}(\lambda) - \varepsilon_S(\lambda) \cdot I(\lambda)$$
(11.12)

for Stokes components S other than I.

To show an example, the LSI method is applied for the sample scenario using $N_{\rm s,low} = 2$ and $N_{\rm s,high} = 16$, and the reconstruction accuracy is shown in Figure 11.3. While the accuracy is clearly below 1%, the residual between $I_{\rm LSI}$ and the full line-by-



Figure 11.3: Example of reconstructed radiance in the strong CO_2 band, normalised to the maximal value of I_{Full} to better highlight the residual slope.

line high-accuracy calculation I_{Full} exhibits a very prominent slope. This slope is a result of the spectrally-dependent aerosol scattering properties. Point 2 in the LSI strategy states that scattering properties are to be taken from the band centre. This step does not account for any spectral dependence, and thus affects the reconstruction accuracy. In the sample scenario, the aerosol optical depth was chosen to be rather

³For details, see O'Dell (2010).

high with $\tau^{(aer)} \sim 0.6$ which exacerbates the problem. In order to accommodate for this issue, one additional set of binned calculations for a single bin is performed, where the spectral properties are taken from one edge of the band, instead of the centre. This adds a wavelength-dependent term to the error:

$$\tilde{\varepsilon}_{\mathcal{S}}(\lambda) = \varepsilon_{\mathcal{S}}(\lambda) + \left[\varepsilon_{\mathcal{S},\text{edge}}(1,1) - \varepsilon_{\mathcal{S}}(1,1)\right] \frac{\lambda_c - \lambda}{\lambda_c - \lambda_{\text{edge}}}.$$
(11.13)

 $\varepsilon_{S,edge}(1,1)$ is the error (according to Equation 11.10) for the (1,1)-bin using scattering properties from wavelength λ_{edge} . This additional calculation therefore provides a slope with which the spectral dependence can be approximated linearly. For the considered bands and the typical applications, a linear correction is generally enough since the wavelength dependence is constructed to be linear (see Section 4.4, Page 37). Applying the slope correction reduces the residual slope, as seen in Figure 11.4.



Figure 11.4: LSI-reconstructed radiances with the applied slope correction (strong CO_2 band, same as in Figure 11.3). The residuals appear spectrally flat.

The reduction in computational speed comes through the reduction in the expensive multiple scattering calculations $I_{\rm MS,high}$, even though the corresponding low-accuracy calculations $I_{\rm MS,low}$ are performed line-by-line. Rather than calling the MS code for $N_{\rm s,high}$ up to 30 000 times for a line-by-line calculation, only $2 \times (N_{\rm bin} + 1)^4$ calls for high-accuracy calculations have to be made. In O'Dell (2010), the number of $N_{\rm bin}$ is 18, 8, and 14 for the three GOSAT or OCO bands. The positions were empirically determined. Also, not all $\tau^{(\text{gas})}$ bins are split into two $\sqrt{\xi}$ bins, as for some bands $\sqrt{\xi}$ does not vary significantly for either low- or high-optical depth bins.

An advantage of LSI is that while Jacobians are not required for the calculation of radiances, the Jacobians for any Stokes parameter can be calculated in a straightforward way by applying elemental differentiation rules to Equation 11.10, as well as Equations 11.11 and 11.12.

 $^{^4}N_{
m bin}$ being the number of bins in $au^{
m (gas)}$ -space.

I have implemented the LSI method described in this section in the UoL-FP algorithm based on a code that was provided by C.W. O'Dell. Prior to this, an older, unpublished experimental version of the method was utilised the UoL-FP algorithm, which treated absorption by water vapour as the "second coordinate", rather than making use of the dimensionless quantity ξ . All of the above figures were created through UoL-FP calculations.

11.3 Linear-k

The linear-k method formulated by Hasekamp et al. (2008) is another spectral binning technique in which the basic assumption is the representation of the multiple scattering Stokes vector \mathbf{S}_{MS} in terms of the absorption optical depth $\tau^{(\text{abs})}$ as well as a normalised vertical distribution \mathbf{n}^{5} :

$$\mathbf{S}_{\mathrm{MS}}(\lambda) = \mathbf{S}_{\mathrm{MS}}\left[\tau^{(\mathrm{abs})}(\lambda), \mathbf{n}(\lambda)\right],\tag{11.14}$$

with

$$\tau_i^{(\text{abs})}(\lambda) = \tau^{(\text{abs})} n_i(\lambda), \qquad (11.15)$$

where

$$\tau^{(\text{abs})} = \sum_{i=1}^{N_{\text{lay}}} \tau_i^{(\text{abs})}.$$
(11.16)

Similar to LSI, Hasekamp et al. (2008) have adopted a method to incorporate the vertical structure of the profiles, however in a more explicit way. As with other spectral binning schemes, a small number of reference calculations are used to approximate high-accuracy calculations. For linear-k, the assumption is that the difference between the reference vertical distribution \mathbf{n}_j for bin j and the actual distribution $\mathbf{n}(\lambda)$ can be used to map the reference calculation to each spectral point:

$$\mathbf{S}_{\mathrm{MS}}\left(\mathbf{n}(\lambda), \tau_{(j)}^{(\mathrm{abs})}\right) \approx \mathbf{S}_{\mathrm{MS}}\left(\mathbf{n}_{(j)}, \tau_{(j)}^{(\mathrm{abs})}\right) + \frac{\partial \mathbf{S}_{\mathrm{MS}}}{\partial \mathbf{n}}\left[\mathbf{n}(\lambda) - \mathbf{n}_{(j)}\right].$$
(11.17)

In the above formulation, MS Jacobians are only available at grid points, so linear interpolation is used to estimate the derivatives at non-grid points. Equation 11.17 is suitable for cases in which both scattering as well as surface properties do not depend on the wavelength, as otherwise residual slopes similar to Figure 11.3 remain.

 $^{^5 {\}rm In}$ Hasekamp et al. (2008), their Equation 7 is erroneous as confirmed through private communication with O. Hasekamp.

Correcting for this wavelength-dependence after the correction for the vertical profile has been performed, is done via

$$\begin{split} \mathbf{S}_{\mathrm{MS}}\left(\tau_{1}^{(\mathrm{sca})},\ldots,\tau_{N_{\mathrm{lay}}}^{(\mathrm{sca})},\mathbf{P}_{1},\ldots,\mathbf{P}_{N_{\mathrm{lay}}},\rho\right) &= \mathbf{S}_{\mathrm{MS}}(\tilde{\tau}^{(\mathrm{sca})},\tilde{\mathbf{P}},\tilde{\rho}) \\ &+ \sum_{i=1}^{N_{\mathrm{lay}}} \frac{\partial \mathbf{S}_{\mathrm{MS}}}{\partial \tau_{i}^{(\mathrm{sca})}} \left(\tau_{i}^{(\mathrm{sca})} - \tilde{\tau}_{i}^{(\mathrm{sca})}\right) \\ &+ \sum_{i=1}^{N_{\mathrm{lay}}} \sum_{s=1}^{6} \sum_{l=0}^{LM} \frac{\partial \mathbf{S}_{\mathrm{MS}}}{\partial \alpha_{i,l,s}} \left(\alpha_{i,l,s} - \tilde{\alpha}_{i,l,s}\right) \\ &+ \frac{\partial \mathbf{S}_{\mathrm{MS}}}{\partial \rho} \left(\rho - \tilde{\rho}\right). \end{split}$$
(11.18)

In Equation 11.18, the quantities with a tilde, e.g. $\tilde{\tau}^{(\text{sca})}$, refer to reference values representative of the bin, and wavelength-dependence is not explicitly written. $\tau_i^{(\text{sca})}$ is the optical depth in layer *i* due to scattering, so essentially a sum of extinction due to Rayleigh scattering and extinction due to aerosol scattering. The $\alpha_{i,l,s}$ are the phase matrix expansion coefficients (the vector equivalent of $\beta_{i,l}$) for layer *i* and expansion order *l*; the index *s* refers to one of the six independent entries of the phase matrix **P**. The last term corrects for the wavelength-dependent surface albedo. Partial derivatives are needed to evaluate Equation 11.18, which requires an RT model capable of returning these Jacobians in order for linear-*k* to still be a feasible method. Hasekamp et al. (2008) mention that the same principle is not applicable to calculate the multiple scattering Jacobians, as that would require second-order derivatives, which is not a standard feature of RT models. Instead, to obtain the Jacobians for every single spectral point, interpolation is used.

The linear-k method has been implemented into the UoL-FP algorithm from scratch without any third-party code. This specific implementation required significant modifications to the scheme as published by Hasekamp et al. (2008). Rather than using the partial derivatives with respect to scattering optical depth (which LIDORT does not produce natively), I was forced to modify the portion of the scheme which corrects for the varying scattering properties (Equation 11.18) and used a finite differencing method instead.

11.4 PCA-based Fast RT Method

The method based on empirical orthogonal functions (EOFs) or principal component analysis (PCA) was published first by Natraj et al. (2005), although Liu et al. (2006)

have independently developed a similar technique. During the research performed for this thesis, the PCA-based fast RT scheme was studied, implemented from scratch, and improved upon. In particular, the formalism has been extended in order to deal with spectrally varying scattering properties due to aerosols.

Similar to the LSI and linear-*k* methods, the PCA-based approach makes use of the large amount of information redundancy within the spectral range of an absorption band that is present in the optical properties. The representative profiles of the optical properties can be inferred from PCA. Within the small spectral range of a typical absorption band, the scattering properties and surface reflectance vary little compared to the dynamic range of gas absorption. This is even more so if the spectral points and the corresponding sets of optical parameters are previously divided into bins that are defined in gas optical depth space, where ideally, all points within a bin are optically similar. The computationally expensive MS calculations can be performed for representative profiles for a bin (rather than for every point in the bin), and the result for all points in that bin can then be derived from that rather effortlessly. In this regard, the PCA-based method is very similar to LSI, as both methods aim to correct a low-accuracy line-by-line calculation.

The following sections describe the steps needed to approximate a high-accuracy RT calculation for a band using the PCA-based approach. Before these steps are carried out, every spectral point in the band is assigned a bin, and each step has to be repeated for every bin.

As was the case for the linear-k method, no prior code was available that could be utilised to integrate the PCA-based method into the retrieval algorithm. I wrote two versions of the method - one being an "offline" implementation that I used to study the method, and then the final module that was then part of the UoL-FP algorithm. This is the first time that this method has been implemented into a full-physics retrieval scheme.

11.4.1 Preparation of Optical Properties

For every bin containing N_s spectral points, an $N_s \times N_o$ matrix **O** is constructed which contains a total of N_o layer-resolved, bulk or auxiliary optical properties. Auxiliary optical properties are wavelength-dependent quantities which, while not directly used in RT computations, are used to derive such quantities. An example would be the aerosol interpolation fraction c_i (Equation 4.9, Page 40). All properties may be transformed using a suitable function beforehand, which is denoted as f^F , and $f^{F}(\mathbf{O}) = \mathbf{F}$. This function f^{F} should have an inverse $f^{B} = (f^{F})^{-1}$ that transforms the quantity back into its original space, such that $f^B(\mathbf{F}) = \mathbf{O}$. Previous work (Natraj et al. 2005; Natraj et al. 2010; Spurr et al. 2013; Kopparla et al. 2016) chose f^F to be the natural logarithm in order to compress the range of gas optical depths, which results in less PCs needed to capture the same amount of variability. In general, f^F can be any bijective function $f^F: X \to Y$, as long as the optical properties are elements of X. Different functions could be applied to each type of optical property. Similar to Natraj et al. (2005), Natraj et al. (2010), and Kopparla et al. (2016), **O** is constructed using the total optical depth profiles $\tau_{i,j}$. However, instead of adding the single-scattering albedos $\omega_{i,j}$, the optical depth due to Rayleigh scattering $\tau_{i,i}^{(\text{Ray})}$ is used. The advantage of this choice is a better reconstruction of the single scatter albedos as well as aerosol properties, which will be explained in more detail in Section 11.4.4. To account for wavelength-dependent scattering properties of aerosols, the aerosol scattering coefficient $q_{i,a}^{(sca)}$ (see Equation 4.7) for each of the N_a aerosol types is included. Additionally, the aerosol interpolation coefficients c_i are added (see Equation 4.9). Finally, similar to Spurr et al. (2013), the last element are the Lambertian surface albedos ρ_i as a bulk property. For different surface models, the surface albedo should be replaced or extended by any number of other spectrally dependent parameters. For the ocean glint Cox-Munk kernel (Cox et al. 1954), this parameter would be the relative refractive index of water. The order of all of the mentioned quantities in the matrix **O** (the column position) is arbitrary as long as the same position of each quantity is used during the reconstruction.

Summarising, the matrix $\mathbf{F} = f^F(\mathbf{O})$ is written as in Equation 11.19, where the tilde over any property signifies that it has been transformed via the function f^F , i.e. $f^F(x) = \tilde{x}$. Thus, the total number of columns in \mathbf{F} is $N_o = 2N_l + N_a + 2$ (for this UoL-FP-specific implementation).

$$\mathbf{F} = \begin{pmatrix} \tilde{\tau}_{1,1} & \dots & \tilde{\tau}_{1,N_l} & \tilde{\tau}_{1,1}^{(\text{Ray})} & \dots & \tilde{\tau}_{1,N_l}^{(\text{Ray})} & \tilde{q}_{1,1}^{(\text{sca})} & \dots & \tilde{q}_{1,N_a}^{(\text{sca})} & \tilde{\rho}_1 & \tilde{c}_1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \tilde{\tau}_{N_s,1} & \dots & \tilde{\tau}_{N_s,N_l} & \tilde{\tau}_{N_{s,1}}^{(\text{Ray})} & \dots & \tilde{\tau}_{N_s,N_l}^{(\text{Ray})} & \tilde{q}_{N_s,1}^{(\text{sca})} & \dots & \tilde{q}_{N_s,N_a}^{(\text{sca})} & \tilde{\rho}_{N_s} & \tilde{c}_{N_s} \end{pmatrix}$$
(11.19)

Contrary to the other properties in **O**, c_i can be negative, in which case the transformation into logarithmic space would fail. For the work in this thesis, $f(x) = \ln(x)$ for all elements, however a constant value of 5 was added to every c_i to avoid the aforementioned issue, as all $c_i > -5$. This constant value is then subtracted again when calculating the binned optical properties.

11.4.2 Varying Scattering Properties and Generalisation of the Method

The specific preparation of the aerosol optical properties, described in Section 11.4.1 (also see Section 4.4, Page 37), is not an integral part of the PCA-based approach itself. Here, it was chosen to comply with the calculations within the UoL-FP retrieval scheme. For applications in which the spectral dependence of extinction and scattering properties of aerosols is ignored, the quantities $q^{(sca)}$, $q^{(ext)}$ and c can be omitted in Equation 11.19. In this case, there is no need to reconstruct the phase function expansion coefficients, since they are the same for each spectral point. Similarly, if the optical properties depend on additional quantities not mentioned here, then these quantities must be incorporated in **O** in the same way. If the coefficients are not parametrised or interpolated in a similar way as shown here, but rather calculated individually for every wavelength, then the PCA-based method would need to be reviewed, as the decomposition in the following section potentially proves difficult if N_{mom} (LM) is of the order of several hundred. On the other hand, for applications in which $N_{\rm mom}$ is considerably smaller, it is conceivable to directly ingest them into the optical property matrix **O**, thus allowing for a direct reconstruction of $\beta_{j,m}$ without any auxiliary parameters.

Neglecting the wavelength dependence of the scattering properties, if it is part of the forward model, will ultimately lead to an overall slope (with wavelength) in the residuals, as seen in Figure 11.3 (Page 154). Contrary to LSI and linear-k, no further RT computations are needed to correct for the varying scattering properties.

The method can potentially be generalised for other RT models. (V)LIDORT (Spurr 2008), TWOSTR (Spurr et al. 2011) and 2OS (Natraj et al. 2007) are models that ingest total atmospheric optical properties (τ, ω, β). Therefore, **O** is set up such that it contains all quantities to exactly calculate the inputs needed by the RT models, no information loss has occurred yet.

It was demonstrated in Efremenko et al. (2014) that the PCA approach can be generalised to other dimensionality reduction techniques, such as local linear embedding methods or discrete linear transforms. They showed for a single example that higher accuracies can potentially be achieved by employing e.g. local linear embedding rather than PCA. As such there is the potential of improving on the results shown in this chapter.

11.4.3 Calculation of the Empirical Orthogonal Functions

The principal components are calculated based on the mean-removed covariance matrix **C**, which is constructed by first subtracting the spectral mean of each optical

property from F, such that

$$\overline{F}_{i,l} = F_{i,l} - \frac{1}{N_s} \sum_{i'=1}^{N_s} F_{i',l}.$$
(11.20)

The covariance matrix \mathbf{C} is then given by

$$\mathbf{C} = \frac{1}{N_s - 1} \left(\overline{\mathbf{F}}^{\mathrm{T}} \overline{\mathbf{F}} \right).$$
(11.21)

Note the normalisation factor in Equation 11.21, which was not present in earlier publications. The eigenproblem for the $N_o \times N_o$ matrix **C** is then:

$$\mathbf{C} \, \mathbf{V}_k = \eta_k \mathbf{V}_k, \tag{11.22}$$

with eigenvalues η_k that correspond to eigenvectors \mathbf{V}_k . The eigenvectors \mathbf{V} , also called empirical orthogonal functions (EOFs), are scaled via the square root of their corresponding eigenvalues,

$$\mathbf{W}_k = \sqrt{\eta_k} \, \mathbf{V}_k, \tag{11.23}$$

and the PCs P are obtained by projecting F onto the EOFs:

$$\mathbf{P}_{k} = \frac{1}{\eta_{k}} \overline{\mathbf{F}} \, \mathbf{W}_{k}. \tag{11.24}$$

In practice, since **C** is real and symmetric, the LAPACK solver DSYEV or appropriate wrappers, such as NumPy's eigh (van der Walt et al. 2011) can be used to compute all \mathbf{V}_k and η_k . In terms of efficiency, the LAPACK routine is able to solve the eigenproblem for a 200×200 matrix in less than a second (Anderson et al. 2007); therefore, its impact on the overall computational effort can be neglected.

The η_k are subsequently ranked by decreasing values, such that \mathbf{V}_1 is the EOF that explains the largest amount of variability of **F**. Generally, the cumulative explained variance for the first three EOFs tends to be > 99%. As stated in Natraj et al. 2005, this implies that three to four EOFs are sufficient to reconstruct the optical properties for every spectral point in the bin to high accuracy.

Due to the symmetry of **C**, all Eigenvalues are positive and real. Eigensolvers such as the aforementioned DSYEV, however, might return some very small negative Eigenvalues for larger k. Since the magnitude of these eigenvalues is small and their contribution can be neglected, all eigenvalues and thus all corresponding scaled and unscaled eigenvectors and PCs can be set to zero if η_k is sufficiently small:

$$|\eta_k| < 10^{-15} \Rightarrow \mathbf{P}_{i,k} = \mathbf{W}_{l,k} = \mathbf{V}_{l,k} = \eta_k = 0, \ \forall \ i, l.$$
 (11.25)

The eigenvalues satisfying the above condition are usually of high order, such that they would not have been relevant at all for the reconstruction, which is often done using less than ~ 10 principal components.

To visualise as well as give physical meaning to the principal component analysis, the mean-removed, transformed optical state matrix $\overline{\mathbf{F}}$ for the O₂ A-band is visualised in Figure 11.5 (first bin, continuum-level) and Figure 11.6 (last bin, line core).



Figure 11.5: Representation of $\overline{\mathbf{F}}$ (excluding the surface albedo) for the first bin in the O₂ A-band. Each colour represents a different percentile based on the set of all spectral points within the bin.

These two figures show the location of the ith percentile of the various components of **F** apart from the surface albedo ρ , as in this particular example ρ was spectrally flat. **F** is the quantity on which the principal component analysis is actually performed, so in terms of the method, the deviations from the transformed mean optical profiles are more significant than the (transformed) mean profiles themselves. For the bin represented in Figure 11.5, the optical depths profiles are dominated by aerosol extinction, as is easily seen in the two peaks representing the high-altitude cirrus layer and the aerosol layer below. Especially in the O_2 A-band, continuum-level spectral points are spread out across the entire band, hence a large spectral variability is inherent in the optical profiles. The top right panel in Figure 11.5 reveals the origin of the spectral variation, which in this case is mostly due to the small aerosol type mixture. It has a much larger spectral variability in this band, compared to the large aerosol mixture or the cirrus mixture, which is spectrally almost flat. For the Rayleigh extinction profiles, the deviations are layer-independent. For the line-core dominated bin (Figure 11.6), the variability in optical depths is larger as can be seen in the top-left panel, despite this bin having only about 10% of the number of spectral points.



Figure 11.6: Same as Figure 11.5, but for the last bin $(5 < \tau \le \infty)$.



Figure 11.7: The first five scaled EOFs **W** for the first and last bins. For bin 22, the components corresponding to the Rayleigh optical depth profiles, aerosol scattering coefficient and fraction have been multiplied by 20 for visualisation purposes. ρ was omitted.

The scaled EOFs **W** (see Equation 11.23) represent mathematically the directions of the largest variability of $\overline{\mathbf{F}}$ in an N_o -dimensional space. The shapes of the profiles in Figure 11.7 are similar to those seen in Figures 11.5 and 11.6. Principal components are calculated by mapping the mean-removed and transformed optical property matrix onto the new set of coordinates, spanned by the EOFs. An example of the first five PCs is shown in Figures 11.8 and 11.9.



Figure 11.8: PCs in the vicinity of three weak oxygen lines in the O_2 A-band, from the first bin.



Figure 11.9: PCs at the position of a strong oxygen line, taken from bin number 21.

The first striking difference between Figure 11.8 and Figure 11.9 is the contribution of the first EOF, which for the weak lines is much smaller than for the strong one. For the strong absorption line, the first PC contributes a symmetric shape, whereas the subsequent PCs add asymmetric components to the line shape, with the fourth EOF strongly modifying the wings. This can be shown by performing the PCA

computation in reverse. F can be reconstructed through

$$\overline{F}'_{i,l} = \sum_{k=1}^{N_{\rm EOF}} P_{i,k} W_{k,l}, \qquad (11.26)$$

which then can be used to obtain the reconstructed optical property matrix

$$O'_{i,l} = f^B \left(\overline{F}'_{i,l} + \frac{1}{N_s} \sum_{i'=1}^{N_s} F_{i',l} \right).$$
(11.27)

The number of EOFs to reconstruct **O** is given by N_{EOF} . When all N_o EOFs are used, then the reconstruction is performed without any information loss, such that **O** = **O**'. Using the first $N_{\text{eof}} = 1...5$ to reconstruct **O**, the optical depths near the strong O₂ line from Figure 11.9 are shown as error percentages compared to the original values in Figure 11.10. What this comparison shows is how the first two PCs are not quite enough to reconstruct the shape of the absorption line to an acceptable degree. Note that the errors for cases $N_{\text{EOF}} = 1, 2$ were re-scaled.



Figure 11.10: Reconstruction of τ around a strong oxygen line using $N_{\text{EOF}} = 1...5$ PCs. For $N_{\text{EOF}} = 1$, the error (as a percentage of the original τ) was scaled by 0.1, and for $N_{\text{EOF}} = 2$ by 0.5 so that all five curves could be fit into the same graph.

11.4.4 Preparation of Binned Optical Properties

For every bin, a so-called mean optical state, $O^{(0)}$ is created by transforming the spectrally averaged **F** back into its original space:

$$O_l^{(0)} = f^B \left(\frac{1}{N_s} \sum_{i=1}^{N_s} F_{i,l} \right).$$
(11.28)

Constructing the total optical depths for the mean state is straight forward, since $\tau_l^{(0)} = O_l^{(0)}$ with $l = 1 \dots N_l$ (see Equation 11.19). The other mean properties, such as mean surface albedo, mean Rayleigh optical depth, and the mean auxiliary properties

 $c^{(0)}$ and $q_a^{(\text{sca},0)}$ are obtained analogously. The single-scattering albedos $\omega_j^{(0)}$, however, are a composite quantity in terms of the contents of **O**, and they have to be calculated from $\tau_j^{(0)}$, $\tau_j^{(\text{Ray},0)}$ and the mean aerosol scattering coefficient $q_a^{(\text{sca},0)}$ (see Equation 4.15, Page 41):

$$\omega_{j}^{(0)} = \frac{\tau_{j}^{(\text{Ray},0)} + \sum_{a=1}^{N_{a}} q_{a}^{(\text{sca},0)} \cdot \tau_{j,a}^{(\text{aer},755)}}{\tau_{j}^{(0)}}.$$
(11.29)

The final quantities needed for the binned calculations are the composite phase function expansion coefficients $\beta_{j,m}^{(0)}$. As explained in Section 11.4.1, the phase function expansion coefficients related to aerosols $\beta_{j,m}^{(aer)}$ are considered to be wavelength dependent and are calculated for two points near the edges of the spectral band. Through linear interpolation between the given points, the coefficients are obtained for every spectral point in the band. For the mean optical state, the mean interpolation factor $c^{(0)}$ is used to calculate the $\beta_{j,m,a}^{(aer,0)}$ via

$$\beta_{j,m,a}^{(\text{aer},0)} = \beta_{j,m,a}^{(\text{aer},\text{beg})} \cdot (1 - c^{(0)}) + \beta_{j,m,a}^{(\text{aer},\text{end})} \cdot c^{(0)}.$$
(11.30)

The mean composite phase function expansion coefficients can then be calculated as:

$$\beta_{j,m}^{(0)} = \frac{\tau_j^{(\text{Ray},0)} \cdot \beta_{m,j}^{(\text{Ray})} + \sum_{a=1}^{N_a} \beta_{j,m,a}^{(\text{aer},0)} \cdot q_a^{(\text{sca},0)} \cdot \tau_{a,j}^{(\text{aer},755)}}{\tau_j^{(\text{Ray},0)} + q_a^{(\text{sca},0)} \cdot \tau_{a,j}^{(\text{aer},755)}}.$$
(11.31)

With that, the total mean optical state is fully defined. The binned optical quantities for the perturbed states +k and -k are computed by perturbing the mean optical properties by the scaled EOFs *W*:

$$O_l^{(\pm k)} = f^B \left(\left[\frac{1}{N_s} \sum_{i=1}^{N_s} F_{i,l} \right] \pm W_{k,l} \right).$$
(11.32)

From $O_l^{(\pm k)}$, the perturbed properties, $\tau_j^{(\pm k)}$, $\tau_j^{(\text{Ray},\pm k)}$, $q_a^{(\text{sca},\pm k)}$, $\rho^{(\pm k)}$ and $c^{(\pm k)}$ are obtained as before, and the composite quantities $\omega_j^{(\pm k)}$ and $\beta_{j,m}^{(\pm k)}$ are calculated analogously to Equations 11.29 to 11.31.

11.4.5 Binned Calculations and Radiance Reconstruction

For each bin, $2N_{\text{EOF}} + 1$ low-stream and high-stream (or first-order and second-order) calculations are performed using LIDORT, 2Stream and 2OS, and the optical inputs $\mathbf{O}^{(0)}$ and $\mathbf{O}^{(\pm k)}$ that were prepared in Section 11.4.1. Further optimisation can be achieved by choosing the number of EOFs individually for each bin, which could reduce the total number of binned calculations.

Using the notation in Spurr et al. (2013), the differences between high- and low-stream binned calculations in logarithmic space are defined as

$$J_{I}^{(0)} = \ln \left[\frac{I_{\rm LD}(\mathbf{O}^{(0)}) + I_{\rm FO}(\mathbf{O}^{(0)})}{I_{\rm 2S}(\mathbf{O}^{(0)}) + I_{\rm FO}(\mathbf{O}^{(0)})} \right]$$
(11.33)

and

$$J_{I}^{(\pm k)} = \ln \left[\frac{I_{\rm LD}(\mathbf{O}^{(\pm k)}) + I_{\rm FO}(\mathbf{O}^{(\pm k)})}{I_{\rm 2S}(\mathbf{O}^{(\pm k)}) + I_{\rm FO}(\mathbf{O}^{(\pm k)})} \right]$$
(11.34)

for the mean and perturbed optical states respectively. The subscripts in $J_I^{(k)}$ signify that this formulation of the correction term is only valid for the intensity component of the Stokes vector, as (modelled) intensities are always positive and hence taking the logarithm will always be possible. For other Stokes components $S \neq I$, which can be both positive and negative, a simple difference between second- and first-order calculations is performed:

$$J_{S}^{(0)} = S_{O2}(\mathbf{O}^{(0)}) - S_{O1}(\mathbf{O}^{(0)}), \qquad (11.35)$$

and

$$J_{S}^{(\pm k)} = S_{O2}(\mathbf{O}^{(\pm k)}) - S_{O1}(\mathbf{O}^{(\pm k)}).$$
(11.36)

Still in logarithmic space for the intensities, the second order central-difference expansion is then applied, where the PCs transform the result back into the spectral space:

$$C_{I,i} = \exp(J_{I,i}),$$

 $C_{S,i} = J_{S,i}.$
(11.37)

with

$$J_{S,i} = J_S^{(0)} + \sum_{k=1}^{N_{\rm EOF}} \frac{J_S^{(+k)} - J_S^{(-k)}}{2} P_{k,i} + \sum_{k=1}^{N_{\rm EOF}} \frac{J_S^{(+k)} - 2J_S^{(0)} + J_S^{(-k)}}{2} P_{k,i}^2.$$
(11.38)

The correction factors C_i are then used to reconstruct the approximate radiance for each spectral point:

$$I_{\text{approx},i} = \left[I_{\text{FO},i} + I_{2S,i}\right] \cdot C_{I,i},$$

$$S_{\text{approx},i} = S_{\text{FO},i} + C_{S,i}$$
(11.39)

Both atmospheric and surface weighting functions are calculated exactly like Stokes components $S \neq I$.

11.5 Detailed Analysis of the PCA-based Method

Before the PCA-based method was actually implemented for use in XCO_2 retrievals, a comprehensive study on the effects of the approximation error was conducted using

GOSAT instrument specifications. To realistically characterise retrieval errors and the spectral residuals of the PCA-based method, a set of simulations was performed that represent the global distribution of expected scenarios covering a wide range of geophysical parameters. Roughly 30 000 model atmospheres have been used to compute the TOA radiances which correspond to locations of GOSAT soundings from two seasons: May, June and July 2011 (summer), and November, December and January 2011/2012 (winter). The GOSAT soundings have been chosen for nearuniform global coverage (cut-off at 60° southern latitude to exclude Antarctica), such that there are at most 5 clear-sky locations in every 2.0 by 2.0 degree grid box. The cloud screening was carried out prior to the simulations via an O_2 A-band fit, where cloudy scenes were identified when the apparent surface pressure deviated more than 30 hPa from the value obtained from European Centre for Medium Range Weather Forecasts (ECMWF) ERA-Interim data. For each season, simulations for ocean-glint scenes have been performed as well, where the wind speed parameter is also taken from ECMWF ERA-Interim. The two seasons are in contrast to each other mostly in terms of solar zenith angles and the associated signal level.

The model atmospheres for each individual sounding are based on Boesch et al. (2013). Temperature, humidity and gas profiles are extracted from ECMWF ERA-Interim and the model CO_2 from the Copernicus Atmosphere Monitoring Service (CAMS) respectively; the spectrally invariant surface albedos (over land) per band are estimated from the radiances in the measured GOSAT spectra. Aerosol profiles are calculated on a per-scene basis from CAMS and consist of five different tropospheric types (sea salt, dust, organic matter, black carbon and sulphate), which are partly further differentiated by size bins and dependence on relative humidity. Every scene contains a thin cirrus cloud aerosol mixture with a constant total optical depth of 0.005. The Gaussian height profile is parametrised by a latitude-dependent mean height and width according to Eguchi et al. (2007).

Due to the global distribution of the sounding locations (see Figure 11.11), the simulations cover a wide range of scenes with different surface types, varying aerosol loadings and solar zenith angles.

The linear sensitivity analysis framework described in Rodgers (2000) is applied here, in which the retrieval error can be estimated from the forward model error, which equals to the spectral error of the PCA-based fast RT method. The error of the state vector $\Delta \mathbf{x}^{EOF}$ is then derived from the gain matrix **G**, which maps spectral features into state vector space:

$$\mathbf{G} = \mathbf{S}_{\mathbf{a}} \mathbf{K}^{\mathrm{T}} (\mathbf{K} \mathbf{S}_{\mathbf{a}} \mathbf{K}^{\mathrm{T}} + \mathbf{S}_{\epsilon})^{-1}, \qquad (11.40)$$

$$\Delta \mathbf{x}^{\text{EOF}} = \mathbf{G} \Delta \mathbf{f} = \mathbf{G} \left[\mathbf{f}^{\text{true}}(\mathbf{x}, \mathbf{b}) - \mathbf{f}^{\text{approx}}(\mathbf{x}, \mathbf{b}) \right].$$
(11.41)



Figure 11.11: Locations for which forward model simulations are performed, and corresponding total aerosol optical depths. For both seasons, the coverage is almost global, where the shift towards high southern latitudes is due to the changing SZA for the winter season. Note the two different latitudinal bands covered by the glint observations over the oceans in the two seasons. Apart from boreal forests and Greenland, most landmasses are covered in both seasons.

 \mathbf{S}_{a} is the prior covariance matrix, and the noise covariance \mathbf{S}_{ϵ} is calculated from the standard deviation of the GOSAT out-of-band signal. The weighting functions $\mathbf{K} = \partial \mathbf{f}^{true} / \partial \mathbf{x}$ are obtained from the high-accuracy simulation \mathbf{f}^{true} , which is a lineby-line simulation using 16 quadrature streams. \mathbf{f}^{approx} is the simulation using the PCA-based approach. Both simulations \mathbf{f}^{true} and \mathbf{f}^{approx} are evaluated using the same state vector \mathbf{x} and auxiliary parameters \mathbf{b} so that the only difference between the two runs is the RT portion of the forward model. \mathbf{b} contains non-statevector quantities that are still needed for forward model computations, such as the volume mixing ratios of non-retrieved gases (O₂, CH₄, H₂O) and various instrument parameters.

The parts of the state vector error that correspond to the CO₂ profile ($\Delta \mathbf{x}^{\text{EOF}(\text{CO}_2)}$) are converted to a column-averaged value using the pressure weighting function **h** (O'Dell et al. 2012):

$$\Delta XCO_2 = \mathbf{h}^{\mathrm{T}} \Delta \mathbf{x}^{\mathrm{EOF}(\mathrm{CO}_2)}.$$
 (11.42)

11.5.1 Spectral Residuals

Spectral residuals are assessed using a quantity that gives the relative error of the radiances on a per-wavelength basis with respect to the full line-by-line calculation $I_{\text{lbl},i}$ for each wavelength index *i* as

$$\Delta I_{\text{rel},i} = \frac{I_{\text{approx},i} - I_{\text{lbl},i}}{I_{\text{lbl},i}}.$$
(11.43)

Finally, residuals for the entire ensemble of simulations are investigated by assessing the histograms of the IQRs in Figure 11.12. The histograms are generally consistent

with the intuitive expectation; increasing the number of EOFs typically reduces the spectral residuals. For simulations that employ the Lambertian surface model (land), both the O_2 A-Band and the strong CO_2 band show an improvement when increasing the number of EOFs from 2 to 3, whereas the improvements for the weak CO_2 band are only marginal when considering the entire ensemble. For the Cox-Munk surface model, on the other hand, the major improvement seems to be when going from 1 to 2 EOFs.



Figure 11.12: Histograms of the spectral residuals. The abscissas represent the interquartile ranges of the absolute values of the relative errors $|I_{\text{rel},i}|$ (see Equation 11.43). The inset text boxes in each sub-plot show the median and the interquartile ratios for each distribution (number of EOFs).

Ocean glint scenes exhibit systematically higher residuals for the O_2 A-band and the strong CO_2 band, and lower residuals for the weak CO_2 band, when compared to land scenes. This result is counter-intuitive, since glint scenes are more dominated by the SS contributions, and thus have smaller MS contributions compared to the SS magnitude. With a small MS contribution, the reconstruction accuracy will depend linearly on the ratio of MS to SS radiances. However, when the MS contributions are higher (in our case up to ~120 % for land scenes, and ~50 % for ocean scenes), the linear regime is no longer valid. Here, the residuals depend on the observation geometry, surface model, and aerosol scattering properties in a non-trivial fashion.

To explore the spatial structure of the error distribution in Figure 11.12, $IQR(|I_{rel,i}|)$

is displayed on global maps in Figure 11.13. Due to the varying scales of the residuals for each map, the respective data is scaled by a factor that is indicated in every sub-plot. The maps reveal apparent drivers for the residuals. For glint simulations, the largest residuals are seen in regions associated with heavy aerosol loadings, such as the Mid-Atlantic ocean with its desert dust outflow, and larger viewing and solar zenith angles. Over land, the residuals are larger with increasing solar zenith angles as well as increasing aerosol optical depths.



Figure 11.13: Global map of residual magnitudes when using 1 EOF. Bins for ocean simulations have been chosen to be large enough to allow for continuous patches despite the spacing of orbits (see Figure 11.11), which makes a visual assessment easier. The factors shown in the top-right corner of each map indicate the factors, with which the respective ensembles have been multiplied (L for land, G for glint) such that the values can be plotted using the same colormap. Regions over land with larger residuals are easily identified, being dusty deserts (Sahara, Arabian peninsula), regions with large amounts of sea salt (high southern latitudes) and aerosol outflow regions (mid-Atlantic, western North-Pacific) as well as high latitudes with higher solar zenith angles.

11.5.2 XCO₂ Errors

Section 11.5.1 highlights certain regions and scenarios for which the PCA-based approach produces larger residuals. These regions, however, do not directly translate into the regions with the largest XCO_2 errors, as the signal-to-noise ratio and the gain matrix affects those errors according to Equation 11.40.

Figure 11.14 shows the calculated errors for the two seasons distributed on global maps as well as separate histograms for land and ocean simulations. There is a notable difference, between the summer and winter seasons, in XCO_2 error behaviour for scenes over the Sahara, the Arabian peninsula and over parts of India; these are

all regions with high signal-to-noise ratio measurements (or equivalently, high albedo in the strong CO_2 band). For the summer season, the XCO_2 error is larger over these areas compared to others when 1, 2 or 3 EOFs are used. In contrast, for the winter season, this is true only when 1 EOF is used.



Figure 11.14: Estimated XCO_2 errors displayed as both maps and histograms. For the case of using only 1 EOF, the errors on the map were scaled by a factor of 0.1 such that the same colorbar could be used. Regions with higher surface reflectance and thus higher SNR clearly result in higher XCO_2 errors (deserts) over land. Over ocean, errors are increased in the mid-Atlantic due to larger AODs (compare Figure 11.11) and the larger viewing angles at the edges of the glint observation bands.

This difference can be seen in the spectral residuals (Figure 11.13) and is driven mainly by the changing solar zenith angle, which is roughly between 10° and 20° in the summer and between 30° and 60° in the winter season.

The effect of changing the number of EOFs on the dependence of the residual errors with respect to optical properties is also seen in Figure 11.14. For summer scenes over the Sahara, the estimated XCO_2 error is largely independent of the total AOD

using only 1 EOF, but shows a rather clear dependence when using 2 or 3 EOFs. The exact opposite behaviour is seen for the winter season. SNR-dependence of the estimated error is similar to the AOD-dependence: for summer scenes, using more than 1 EOF leads to a clear dependence, whereas for winter scenes, this dependence is much weaker.

The generally larger spectral error for glint observations (see Figure 11.12), translates into a larger mean XCO_2 error. Land scenes in the summer season show a higher error over the Sahara and Arabian peninsula, coinciding with those scenes with a high signal-to-noise ratio. There is a change in sign for the overall error for land observations, but the magnitude of the errors is already small with an IQR of the distributions of around 0.06 ppm and lower. Glint-type soundings using the Cox-Munk surface model show higher scatter for each season, when only 1 EOF is used, and is comparable to land scenes when using 2 or 3 EOFs. To summarise the dependence of the XCO_2 error on the number of EOFs, the percentage of scenes with errors larger than 0.1 ppm are stated in Table 11.1.

	Summer		Winter	
# EOFs	Land	Glint	Land	Glint
1	67.09%	82.10 %	78.85%	76.84%
2	5.59%	1.36~%	5.76%	0.29%
3	5.63%	0.75%	2.18%	0.47%
5	0.39%	0.78%	0.21%	0.40%

Table 11.1: Percentage of scenes with XCO_2 errors larger than ± 0.1 ppm.

11.5.3 Influence of Cirrus Optical Depth

To assess the errors induced by different cirrus optical depths, the XCO_2 errors of a small subset of $N\sim500$, globally distributed scenes from the winter/land season set was additionally simulated for four different total optical depths of cirrus clouds: 0.01, 0.1, 0.25 and 0.5. Box plots illustrating the results are shown in Figure 11.15.

Since XCO_2 retrievals tend to be filtered very restrictively using the total retrieved cirrus OD (e.g. a filter threshold of < 0.05 as stated in Cogan et al. (2012)), the cases of 0.01 and 0.1 are most relevant. For the latter case, the standard deviation (as well as the IQR) of the errors in this subset increased roughly by a factor of 2 to 3 compared to the original scenario of 0.005. These IQRs do not exceed 0.7 ppm using 1 EOF, and decrease to 0.05 ppm using 3 EOFs. Comparing these numbers to the those in Figure 11.14 one can see that the IQRs of the respective distributions roughly doubled/tripled when the total cirrus OD was increased to 0.1.



Figure 11.15: XCO_2 errors at various cirrus cloud optical depths for a small subset of the winter season scenes. There is a dramatic reduction of the error magnitude going from 1 EOF to 2 EOFs, especially for cases with cirrus optical depth larger than 0.1. The tick labels for the 1 EOF case are seen on the left hand side of the three plots, whereas the tick labels for the cases of 2 and 3 EOFs are seen on the right hand side. Whiskers in this plot have been suppressed.

The conclusion here is that high-altitude aerosols, such as cirrus clouds, have a similar effect on the reconstruction accuracy, and ultimately the retrieval error, as aerosols in the lower troposphere in the glint observation mode. The PCA-based approach performs well for cirrus optical depths as large as 0.5 when 3 EOFs are used.

11.5.4 Investigating Different Instrument Configurations

The XCO₂ error from the PCA-based method will change for different instruments. To evaluate how the PCA-based approach performs for other instruments, linear error analysis is performed for a subset ($N \sim 5500$) of scenes using different instrument models. The assumed instrument models are described by their spectral resolution, i.e. the full width at half maximum (FWHM) of the Gaussian instrument line shape (ILS) functions, the noise model, and the dispersion relation as derived from the sampling per FWHM. A potentially different field of view or footprint size is not taken into account, and neither is the sensitivity towards the state of polarisation. The spectral windows and the viewing geometries also remain unchanged to allow for a direct comparison between the simulations with the only difference being the instrument models. The scenes themselves still correspond to specific GOSAT sounding locations found in both the summer and winter season sets.

The subset of scenes was chosen to contain predominantly four different surface types: tree cover, bare areas, sparse vegetation and water. A scene falls into one of these categories if at least 95% of the land cover class grid boxes enclosed by the GOSAT footprint belong to one of the listed indices in Table 11.2.

Three additional instruments are considered, all of which are grating-spectrometer types, as opposed to GOSAT being a Fourier-transform spectrometer. The first one

Surface type	Land cover class indices		
Tree cover	50, 60, 61, 62, 70, 71, 72, 80, 81, 82, 90		
Bare area	200, 201, 202		
Sparse vegetation	120, 121, 122, 130, 140, 150		
Water	220		

Table 11.2: Land cover class indices according to the ESA-CCI Land Cover map (epoch 2008-2012) associated with the four surface types.

is an OCO-2 like (Crisp et al. 2017) instrument that features high SNR and high resolution across all three bands. The second instrument is to resemble the Earth Explorer 8 candidate mission CarbonSat (Buchwitz et al. 2010) which has lower resolution than the first instrument, especially in band 2 and 3. Finally, the last instrument resembles Sentinel 5-like specifications (Ingmann et al. 2012), with a much lower resolution in the O_2 A-Band, but higher resolving power in the strong CO_2 band at 2.06 µm. The 2.06 µm band is not present in the design specifications of Sentinel 5; the instrument characteristics for that band from earlier studies was assumed (European Space Agency 2011b, 2011a). Table 11.3 lists the spectral characteristics for the various instruments. Various noise models are utilised to

Instrument	ILS FWHM	Sampling Ratio
GOSAT	$0.35, 0.25, 0.24 \ [\mathrm{cm}^{-1}]$	1.4
A (OCO-2 like)	0.042, 0.076, 0.097 [nm]	2.5
B (CarbonSat like)	0.1, 0.3, 0.55 [nm]	3
C (Sentinel 5 like)	0.4, 0.25, 0.125 [nm]	3

Table 11.3: Instrument model characteristics, the FWHMs of the Gaussian instrument line shape functions are stated for all three bands (O_2 A-band, weak CO_2 and strong CO_2). Sampling ratio is the number of spectral points per FWHM.

simulate different instruments in order to assess the impact of noise and spectral resolution on the reconstruction accuracy. The noise model for instrument A (OCO-2 like) is a simplified version of the one stated in Eldering et al. (2015):

$$\mathrm{SNR}_A = \sqrt{\frac{L^2}{aC_{\mathrm{back}}^2 + bC_{\mathrm{ph}}^2 L}},\tag{11.44}$$

with *L* being the radiance value in units of ph s⁻¹ m⁻² sr⁻¹ μ m⁻¹. While in Eldering et al. (2015), the coefficients C_{back} and C_{ph} are detector-pixel dependent, here they were chosen to be constant values for each band. Similarly, for the CarbonSat like instrument B, the following noise model was used (Buchwitz et al. 2010),

$$SNR_B = \sqrt{3} \frac{SNR_{ref} \cdot \frac{L}{L_{ref}}}{\sqrt{2 + \frac{L}{L_{ref}}}}.$$
(11.45)

Finally, instrument C was modelled after Sentinel-5 like specifications, with the noise model

$$\text{SNR}_C = \text{SNR}_{\text{ref}} \sqrt{\frac{L}{L_{\text{ref}}}}.$$
 (11.46)

The coefficients used in the above noise models are summarised in Table 11.4.

Coefficient	O_2 A-Band	Weak CO ₂	Strong CO ₂			
Instrument A (OCO-2 like)						
а	$4.9\cdot 10^{37}$ $6.0025\cdot 10^{36}$ $1.5625\cdot 10^{36}$					
b	$7.0\cdot10^{18}$	$2.45\cdot10^{18}$	$1.25\cdot10^{18}$			
C_{back}	0.00497	0.00671	0.0149			
$C_{ m ph}$	0.00961	0.00706	0.008			
Instrument B (CarbonSat like)						
SNR_{ref}	150	160	130			
$L_{ m ref}$	$4.2\cdot10^{19}$	$1.5\cdot 10^{19}$	$3.8\cdot10^{18}$			
Instrument C (Sentinel 5 like)						
SNR_{ref}	500	300	100			
$L_{ m ref}$	$4.49\cdot 10^{19}$	$1.15\cdot10^{19}$	$5.0\cdot10^{18}$			

Table 11.4: Instrument model coefficients. Radiances (L, L_{ref}) and coefficient b are in ph s⁻¹ m⁻² sr⁻¹ µm⁻¹, a in $(ph s^{-1} m^{-2} sr^{-1} µm^{-1})^2$, and C_{back} , C_{ph} and SNR_{ref} are unitless.

The comparison of the estimated XCO_2 errors for the four instrument types is visualised in Figure 11.16, where the scenes have been aggregated according to AOD bins. For these simulations 3 EOFs were used for the radiance reconstruction.

Instrument C exhibits systematically larger errors than instrument B for all surface types apart from ocean, indicating that the spectral resolution in the strong CO_2 band is of higher relevance to the retrieval error than the noise levels. When only 1 EOF is used (not shown), the dependence on the AOD for all surface types and instrument models is much higher.

From this analysis, it can be concluded that if 3 EOFs are used, the PCA-based approach is able to reconstruct the radiances to an accuracy at which the forward model errors result in less than ± 0.2 ppm error in the retrieved XCO₂ for a variety of instrument models. Using only 1 EOF will inadvertently cause biases in high-AOD scenes on the order of 1 ppm.

11.6 Performance Aspects of the Fast RT Methods

The three main fast RT methods presented here vary both in terms of the methodology, complexity, as well as computational efficiency and finally accuracy. To understand



Figure 11.16: XCO_2 errors (using 3 EOFs) of scenes grouped into aerosol optical depth bins (the dashed, grey lines indicate the bin boundaries) for the four instruments listed in Table 11.3 (left to right, GOSAT: grey, A: blue, B: red, C: green) for different surfaces.

the potential impact of choosing one method over the other, it is helpful to compare them side by side and highlight the differences. It must be noted that a legitimate comparison between the methods is only possible if all methods are implemented consistently in the same retrieval framework. Naively comparing results across publications (Hasekamp et al. 2008; O'Dell 2010; **Somkuti**, Boesch, Natraj, et al. 2017a) is a flawed approach as they all use different retrieval setups.

The PCA-based method and LSI are similar in their general approach - both methods aim to correct a low-accuracy line-by-line calculation. The correction is done by binning spectral points according to their optical depth, and performing both lowand high-accuracy binned calculations for this small number of representative optical profiles. Using the results of the binned calculations, a correction term is derived for every spectral bin and subsequently applied to obtain an approximation to the full high-accuracy calculation. Both methods also take into account the vertical structure of the gas absorption, although in different ways. LSI collapses the vertical profile information in a single value ξ , whereas in the PCA-based method the variability in the vertical profiles is directly embedded into the binned calculations and the line-by-line principal components.

Linear-k takes a slightly different approach by separating the SS and MS components

from the start. Since the SS contribution can be calculated very quickly for a large number of spectral points once the optical properties and the phase matrix are known, the focus is to approximate the MS calculations through representative, binned ones. The vertical profiles are taken into account directly by correcting the binned MS calculations using the (analytic) derivative $\partial \mathbf{S}/\partial \mathbf{n}$. The linear-*k* method does not rely on line-by-line MS calculations, making it the fastest of the three schemes.

The computational efficiency of the three methods is somewhat difficult to measure, as the performance will depend on the implementation as much as on the method itself. While it is possible to simply measure the elapsed CPU time for a forward model simulation or an entire retrieval, this number might be more representative of the implementation into the retrieval algorithm than the fast RT scheme. This is definitely the case for the UoL-FP algorithm. It is structured such that either SS, MS, or a combination of both radiance calculations are carried out in a monochromatic loop. This allows to exchange the underlying RT code(s) easily, without having to modify any parts the main loop. Unfortunately this creates a very large overhead, for example when only SS calculations are required, which could be performed in one single, vectorised array operation. Comparing the elapsed CPU time for retrievals using the UoL-FP algorithm would therefore penalise the linear-k method, as an alternative SS implementation would reduce the computation times even further.

Rather than stating a measured, but biased value of the computation time of the various methods, the computational efficiency can assessed by the total number of SS and MS calculations, and adjusting for whether Jacobians are needed or not. O'Dell (2010) has discussed this, however did not provide any detailed numbers or estimates for the speedup apart from a wide range of 1 to 2 orders of magnitude. The speedup & is generally defined as the ratio between the durations of the approximated calculations and the full line-by-line calculations

$$\mathcal{S} = \frac{t(\text{line} - \text{by} - \text{line})}{t(\text{approximation})}.$$
 (11.47)

The duration for line-by-line calculations are easily written as the sum of MS and SS durations multiplied by the number of spectral points N_s .

$$t_{\rm rad}(\text{line-by-line}) = N_s \left(t_{\rm rad}^{\rm SS} + t_{\rm rad,high}^{\rm MS} \right),$$

$$t_{\rm jac}(\text{line-by-line}) = N_s \left(t_{\rm jac}^{\rm SS} + t_{\rm jac,high}^{\rm MS} \right).$$
(11.48)

In Equation 11.48, the *t* represent the CPU time needed for a single call to an RT model, calculating either the SS or MS contribution to the total TOA radiance. The subscripts *rad* or *jac* signify whether the RT model is assigned to calculate radiances only, or to also calculate analytic weighting functions. For LSI, the duration due to

RT calls is

$$t(\text{LSI})_{\text{rad}} = N_s \left(t_{\text{rad}}^{\text{SS}} + t_{\text{rad,low}}^{\text{MS}} \right) + \left(2N_{\text{bin}} - N_{\text{single}} + 1 \right) \cdot \left(\underbrace{t_{\text{rad}}^{\text{SS}} + t_{\text{rad,low}}^{\text{MS}}}_{\text{low accuracy binned}} + \underbrace{t_{\text{rad}}^{\text{SS}} + t_{\text{rad,high}}^{\text{MS}}}_{\text{high accuracy binned}} \right), \quad (11.49)$$

and analogously for when calculating weighting functions. N_{bin} represents the number of bins and N_{single} is the number of bins for which only one $\sqrt{\xi}$ bin is used, usually for very low and very large gas optical depths. Similarly, the duration using the PCA-based method can be written as

$$t(\text{PCA})_{\text{rad}} = N_s \left(t_{\text{rad}}^{\text{SS}} + t_{\text{rad,low}}^{\text{MS}} \right) + N_{\text{bin}} \left(2N_{\text{EOF}} + 1 \right) \cdot \left(\underbrace{t_{\text{rad}}^{\text{SS}} + t_{\text{rad,low}}^{\text{MS}}}_{\text{low accuracy binned}} + \underbrace{t_{\text{rad,high}}^{\text{SS}} + t_{\text{rad,high}}^{\text{SS}}}_{\text{high accuracy binned}} \right), \quad (11.50)$$

assuming that the same number of EOFs are used for every bin. The final method is linear-k, with a total computation time of

$$t(\text{linear-}k)_{\text{rad}} = \left(N_s \ t_{\text{rad}}^{\text{SS}}\right) + \left(N_{\text{bin}} \ t_{\text{jac}^1,\text{high}}^{\text{MS}}\right),$$

$$t(\text{linear-}k)_{\text{jac}} = \left(N_s \ t_{\text{rad}}^{\text{SS}}\right) + \left(N_{\text{bin}} \ t_{\text{jac}^2,\text{high}}^{\text{MS}}\right).$$
 (11.51)

For linear-k, the MS computations for the binned calculations always require Jacobians to be calculated as well. For radiance only computations, either one (Equation 11.17) or more (Equation 11.18) atmospheric and surface weighting functions need to be calculated. When Jacobians are also required, the same weighting functions need to be calculated potentially *in addition* to those required by the retrieval algorithm.

To calculate the speedup for each method (and for each band), values for the per-point RT calculation durations have to be simply entered into the relevant equations above, along with the numbers of bins, EOFs and spectral points per band. As mentioned before, the SS contributions alone can be very quickly calculated in a fully vectorised array operation, as laid out in the ATBD of the SRON CCI-GHG full-physics XCO₂ data product (Hasekamp et al. 2016). It is therefore reasonable to assume that in an optimised implementation $t^{SS} \ll t^{MS}$ in either Jacobian or radiance mode. After neglecting the single-scatter contributions, a further simplification can be made, which is to relate a high-accuracy to a low-accuracy computation duration through a factor $\eta = t_{low}^{MS}/t_{high}^{MS}$. This factor will be different for radiance-only and Jacobian RT model calls, and also depend on the number of Jacobians to be calculated, as well as the number of layers in the model atmosphere. The speedup value for each of the

three fast RT methods using these assumptions can thus be written as

$$S(\text{LSI}) = \frac{N_s}{\eta N_s + (2N_{\text{bin}} - N_{\text{single}} + 1)(\eta + 1)},$$
 (11.52)

$$S(PCA) = \frac{N_s}{\eta N_s + N_{\rm bin}(2N_{\rm EOF} + 1)(\eta + 1)},$$
(11.53)

$$S(\text{linear-}k) = \frac{N_s}{N_{\text{bin}}\eta^{\dagger}}.$$
(11.54)

It is important to note that η^{\dagger} is different from η , since the linear-*k* method requires Jacobians to be calculated even for radiance-only computations. η^{\dagger} is closer to 1, considering that only a few more Jacobians are added to the calculation. The speedup values for an exemplary case are shown in the following figures, where $N_s = 30\,000$ which is roughly equivalent to an O₂ A-band calculation on a 0.01 cm⁻¹ grid. Obviously, η is highly dependent on the number of quadrature streams for the



Figure 11.17: Speedup *S* as a function of η and the number of bins N_{bin} . For the PCA-based method, the number of EOFs was fixed at 3 for every bin.

high-accuracy calculations, as well as the number of required atmospheric weighting functions. Depending on the way of measuring η , the actual value of S can significantly vary, however a few general statements can be made (see Figure 11.17): both LSI and EOF result in speedups of one to two orders of magnitude, with LSI being faster for the same number of bins in $\tau^{(gas)}$ space. Linear-k, due to the omission of low-accuracy MS calculations, offers a speedup between two and three orders of magnitude. The speedups presented in Figure 11.17 are extreme cases in which the SS calculations do not play a role and are not quite representative of the speedup of an entire retrieval.

Chapter Twelve

CO₂ Retrievals from OCO-2



O retrieve XCO₂ from OCO-2 measurements, the UoL-FP algorithm was adapted to include instrument specific extensions to the forward model, such as the footprint-dependent residual fitting technique. Using the PCAbased radiative transfer acceleration technique discussed in Section 11.4, the OCO-2

retrievals were assessed against TCCON to validate the algorithm and the retrieval setup.

An investigation of XCO_2 errors due to the RT approximation technique, described in Section 11.5, allowed the estimation of that quantity without having to actually compute retrievals. The linear error analysis framework, however, assumes a fully linear retrieval problem, which for XCO₂ retrievals does not always hold.

In this chapter, the three methods described in the previous one were used for retrievals using real OCO-2 measurements, and the results were compared. This, again, is the first time three contemporary fast RT methods were implemented in a consistent framework. This facilitates the study of the effect of the fast RT method on the retrieved CO_2 concentrations when every other aspect of the retrieval is kept the same.

Retrieval Setup and Implementation of Fast RT Methods 12.1

The used retrieval setup is an adaption from the UoL-FP retrieval scheme used in the GHG-CCI initiative (Buchwitz, Reuter, Schneising, Hewson, et al. 2017). Meteorological data (surface pressure, humidity and air temperature profiles), already sampled at the corresponding OCO-2 sounding locations, was obtained from "GES DISC" (2018). SIF priors for land measurements are taken from the OCO-2 IDP v8 data product, available at the same data portal. Prior XCO₂ profiles are informed from the MACC 16r1 model. The UoL-FP aerosol scheme is employed with a minor modification: a fourth aerosol mixture, based on a sulphate-type aerosol is added as an optically thin ($\tau = 0.005$) layer into the stratosphere at 20 km altitude and $3 \,\mathrm{km}$ width. This is to mitigate a known issue, where ocean glint retrievals show

an unphysical enhancement at high southern latitudes. Stratospheric sulphate-type aerosols, also known as the Junge layer, are mostly due to volcanic eruptions.

The retrieval utilises a purely Lambertian, non-polarising BRDF for both land and ocean scenes. Initial guesses (= a priori value) for the surface albedo are estimated from the measured radiances:

$$\rho_{\rm a} = \frac{\pi \overline{I}_{\rm meas}}{\mu_0 I_0},\tag{12.1}$$

where $\overline{I}_{\text{meas}}$ is the band-averaged continuum-level intensity, measured by the OCO-2 instrument, and μ_0 is the cosine solar zenith angle. I_0 is the intensity component of the solar irradiance after application of the instrument Müller matrix

$$\mathbf{S}_0 = \mathbf{M}\mathbf{S}_0',\tag{12.2}$$

with \mathbf{S}'_0 being the Stokes vector of the unpolarised solar irradiance $\mathbf{S}'_0 = (I'_0, 0, 0, 0)^{\mathrm{T}}$, the only relevant entry in \mathbf{M} , m_{11} is usually 0.5 for the OCO-2 instrument. For the three bands, I'_0 was set to $4.7 \cdot 10^{21}$, $2 \cdot 10^{21}$, and $1.3 \cdot 10^{21}$ ph s⁻¹ m⁻² sr⁻¹ µm⁻¹, respectively.

Residual fitting, with residual waveforms having been derived through principal component analysis from a previous run, is employed by fitting three waveforms per band¹. They are weighted by a coefficient r_k and added to the ILS-convolved radiances $I^*_{\text{conv}}(\lambda_i)$,

$$I_{\text{conv}}(\lambda_i) = I_{\text{conv}}^*(\lambda_i) + \sum_{k=1}^N r_k R_k(\lambda_i).$$
(12.3)

These waveforms correct mostly for imperfections in the spectroscopic data, and it is assumed that the choice of fast RT method does not influence the derivation of the residual waveforms, or their effectiveness during the retrieval. Without applying the waveforms, the final reduced χ^2 values would not be close to 1, but rather five to ten times larger. Waveforms are derived and applied for each OCO-2 footprint individually, but not for the different measurement modes or surface types. The coefficients r_k are part of the state vector, and in this setup, three waveforms are used for every band.

The retrieval setup differs between ocean and land scenes. For measurements over land, the state vector includes the SIF radiance at 755 nm as well as the ratio of the SIF radiances at 755 nm and 772 nm, therefore imposing a slope. Tests have shown that better convergence is obtained if the initial value of the Levenberg-Marquardt γ -parameter is changed. For ocean scenes a value of $\gamma = 5000$ is used, whereas

¹The residual waveforms are often and confusingly called EOFs. In this thesis, the notion of the EOF is reserved for the eigenvectors of the covariance matrix (see Equation 11.22), whereas the principal components are the original data matrices projected onto the EOFs. The waveforms $R_k(\lambda_i)$ should therefore be called PCs, rather than EOFs.

for land scenes $\gamma = 500$. The main effect from increasing the initial value of γ (compared to a GOSAT setup of $\gamma = 10$) is to reduce the step size of the state vector changes between iterations, which resulted in many divergent steps due to the aerosol extinction profiles exhibiting very large updates. For land retrievals, the convergence criterion using the convergence scale factor (see Equation 5.15, Page 52) was loosened by using a value of 10 rather than 3.

The code for the LSI method was provided by C. W. O'Dell and was then integrated into the UoL-FP retrieval algorithm. The linear-k method was implemented independently, however two minor modifications to the original publication by Hasekamp et al. (2008) had to be made. First, rather than binning in $\tau^{(abs)}$ space, a binning in $\tau^{(\text{gas})}$ was used (as is done in the other two methods). Secondly, the partial derivatives with respect to scattering optical depth, as well as phase function expansion coefficients (see Equation 11.18) did not work as expected with the RT models present in the UoL-FP algorithm. To achieve a similar correction for the spectrally varying scattering and surface properties, the analytic derivative was essentially replaced by a finite difference approach: for every binned calculation with the scattering and surface properties taken from the centre of the band, a second binned calculation was performed, but the scattering and surface properties were taken from one of the band edges. While this effectively doubles the number of binned calculations, this approach achieves the same effect as the use of the analytic derivatives. The essential portion of the linear-k method, which is the correction for the vertical profiles, remains the same. Using multiple grids for different absorbers has also been omitted, as the net gain in accuracy improvement are small, but would make the implementation more complicated². The PCA-based method was implemented as described in Section 11.4.

The binning strategy for each method was taken from the respective publications. For LSI, the bin boundaries from Table 3 of O'Dell (2010) are used, which prescribes 18, 8, and 14 bins for the three bands. Linear-k bins were chosen similarly to the description in Hasekamp et al. (2016), where the bins are logarithmically spaced in $\tau^{(\text{gas})}$ space,

$$\tau_k^{(\text{gas})} = \exp\left\{\log\left(\tau_{\min}^{(\text{gas})}\right) + (k-1)\left[\log\left(\tau_{\max}^{(\text{gas})}\right) - \log\left(\tau_{\min}^{(\text{gas})}\right)\right]/N\right\}$$
(12.4)

Here, the largest total-column gas optical depth $\tau_{\max}^{(\text{gas})}$ is not limited to 15, as was described in the ATBD, but left unchanged. This way, the bin boundaries are constructed through the actual range of gas optical depths within the band. Should

²Private communication with O. Hasekamp.

there be spectral points with $\tau^{(\text{gas})} = 0$, these points are discarded for the purposes of the bin boundary calculation, and are then assigned to the first bin. Tests have revealed that this approach to generating the bin boundaries results in smaller residuals, especially in the O₂ A-band. For each band, 15 bins are generated. For the PCA-based method, the binning strategy from Kopparla et al. (2017) was utilised which results in 10 bins in gas optical depth space, where each bin is sub-divided into two further bins along the median of ξ (see Equation 11.9), thus resulting in 20 bins in total.

For the calculation of the Jacobians, the interpolation and approximation is actually performed on the atmospheric and surface weighting functions for all three methods.

12.2 OCO-2 Retrievals Using the PCA-based Method

The UoL-FP retrieval setup was historically tailored to process the GOSAT measurement record. To confirm that the changes in the retrieval scheme and processing pipeline work as intended for OCO-2, a set of scenes was retrieved and validated against the Total Column Carbon Observing Network (TCCON). TCCON is a global network of ground sites equipped with Fourier-transform spectrometers that perform regular measurements. The instruments are pointed directly towards the sun, thus guaranteeing a high signal-to-noise ratio. Wunch, Toon, et al. (2011) describe the retrieval methodology in detail. In the development of globally-spanning satellite records of atmospheric CO_2 , TCCON has proven to be an invaluable validation source, as it allowed deriving quality filter criteria and explore biases in the XCO₂ retrievals. Figure 12.1 shows the location of TCCON sites that were used to validate the OCO-2 retrieval setup.



Figure 12.1: Location of the 13 TCCON sites that were used for the validation.

Matching a satellite overpass to a TCCON measurement is a difficult task in itself, as one must consider that given the meteorological conditions, a satellite measurement at a distance away from the ground station potentially samples different air masses. Wunch, Wennberg, Toon, et al. (2011) have proposed a method that uses the air temperature at a pressure level of 700 hPa, T_{700} , to determine collocation between any satellite measurement and a coincident TCCON measurement. If T_{700} at the satellite measurement location is within ΔT of T_{700} at the TCCON location, one can consider the satellite sampling the same air parcel as the TCCON station. Thankfully, OCO-2/TCCON pairings acquired via this method were provided by Brendan Fisher from the Jet Propulsion Laboratory.

The supplied validation set using version 8 L1B radiances contains $N\sim 245\,000$ soundings from September 2014 to December 2016, in nadir and glint measurement modes for both land and ocean surfaces. A basic filter was applied to retain only converged retrievals for which the fit quality in each band fulfils $0.75 < \chi^2 < 2$, and less than three divergent steps were performed. This leaves about $N\sim 200\,000$ (~80%) retrievals, of which $N\sim 175\,000$ had appropriate TCCON overpasses. The retrievals were computed using the PCA-based fast RT method using 3 EOFs per bin.

A known feature of OCO-2 retrievals related to the instrument is the footprintdependent bias (Wunch et al. 2017). As each footprint can be treated as a separate instrument (see Section 3.2, Page 23), and the footprints are spatially adjacent, a bias can be inferred. Assuming that XCO_2 is approximately constant across a frame (all eight footprints), differences to the median value can be attributed to instrument or retrieval biases. For the determination of the footprint bias, only frames with at least six valid retrievals were taken into account. The derived footprint biases are shown in Figure 12.2.



Figure 12.2: Footprint biases for each measurement mode separately.

While there are slight differences in the different measurement modes and surface types, footprints number six and eight are large outliers for all modes, which is similar to the biases reported in Osterman et al. (2017). Given that the biases are found to be close to those published by Osterman et al. (2017), the biases are indeed

likely to be related to the instrument or the instrument calibration, rather than the algorithm. The median value of the footprint-dependent differences to the frame medians (orange lines in Figure 12.2) is then subtracted from every retrieved XCO_2 value according to the footprint number.

The comparison of the retrieved and footprint bias-removed XCO_2 against the TCCON-collocated measurements (for all stations) is shown in Figure 12.3 and Figure 12.4.



Figure 12.3: Comparison of land scenes (nadir & glint) against TCCON measurements, which were collocated by Brendan Fisher (JPL). The left panel shows the point cloud coloured by the mode of the footprints within the bin. An apparent footprint bias would be seen as differently coloured and distinct regions in the correlation plot. Footprint number eight is slightly overrepresented in the sample, hence why they are also represented as the dominant colour in this plot.

It has to be noted that the processed subset of OCO-2 soundings is pre-selected and contains mostly "well-behaved" scenes, i.e. feature low aerosol loading. Hence the statistics shown in those two figures are very favourable with a mean bias of \sim 1 ppm. Land scenes show a higher overall scatter with 1.6 ppm compared to the ocean measurements with 1.1 ppm. The left panels in the figures show the distribution of footprints. For every cell in the two-dimensional histograms, the colour is determined by the most prevalent footprint number. The somewhat random-seeming distribution of the colours associated with the footprints is a good indication that there are no residual footprint-dependent features left. In Figure 12.3, footprint number eight seems to be predominant in this TCCON comparison, however this is a result of a non-uniform distribution of footprints in the data set (Figure 12.5).



Figure 12.4: Same as Figure 12.4, but for ocean scenes.



Figure 12.5: Distribution of footprints in the subset after minimal filtering for convergence and fit quality.

Retrieval biases were identified empirically by visually investigating diagnostic graphs like Figure 12.6, in which a linear model is fitted to the UoL-FP/TCCON differences as a function of a retrieval-related parameter.

For land retrievals, biases with respect to the following parameters were identified: the difference between retrieved and prior surface pressure (Δp_{surf}) , the ratio of retrieved albedos between the O₂ A-band and the strong CO₂ band, as well as the ratio between the surface- and 15th-level of the retrieved CO₂ profile (CO₂^{grad}). For ocean scenes, bias correction parameters are the XCO₂ uncertainty as well as Δp_{surf} .

The bias correction procedure was performed for land (nadir & glint) and ocean glint measurements separately, and the derived coefficients are tabulated in Table 12.1.



Figure 12.6: This figure demonstrates а systematic surface pressure bias for land explains which retrievals, almost 23% of the variance the UoL-FP/TCCON of differences. The bias is roughly -0.35 ppm hPa⁻¹. Inspired by C.W. O'Dell's diagnostic plots.

Table	12.1:	Bias	correction	coefficients	5.

The bias-corrected XCO_2 values are obtained by subtracting the bias term:

$$\operatorname{XCO}_{2}^{\operatorname{corr}} = \operatorname{XCO}_{2} - \left(\sum_{i=0}^{N_{\operatorname{coef}}} b_{i} p_{i}\right), \qquad (12.5)$$

where b_i are the bias correction coefficients from Table 12.1, and the p_i are the corresponding parameters of the particular sounding (Δp_{surf} , etc.). After bias correction, the values are compared against the matched TCCON measurements, and summarised in Table 12.2. The overall bias is by definition 0 ppm, and the overall scatter (standard deviation) is 1.2 ppm. The values are comparable with Wunch et al. (2017) and thus show that the retrieval set-up produces results that are comparable with the official OCO-2 data product. Admittedly, the pre-selection and filtering of soundings is an arduous and complicated task, which was already done for the set of measurements provided by Brendan Fisher.

12.3 Intercomparison of Fast RT Methods

To visualise the radiance residuals for all three fast RT methods, a set of typical retrieval scenarios was constructed for which full line-by-line calculations as well as approximations were computed. The surface was chosen to represent a typical grass-like biome with albedos 0.47, 0.29 and 0.07 for the three bands. The viewing

Site	Δ [ppm]	σ [ppm]	N
Anmeyondo	-0.38	1.53	721
Armstrong	0.40	1.54	17512
Ascension	-0.13	0.93	8077
Bialystok	-0.02	1.11	3182
Bremen	1.12	1.39	467
Caltech	-0.09	1.41	1253
Darwin	0.22	0.91	22383
Garmisch	0.36	1.89	3101
Izana	-0.70	1.18	795
Karlsruhe	0.93	1.31	4571
Lamont	-0.14	1.10	63783
Lauder	-0.17	0.72	7531
Manaus	-0.38	1.47	5
Orleans	0.61	1.28	3635
Paris	-0.06	1.42	3594
Park Falls	-0.24	1.16	10736
Reunion	-0.08	0.81	8269
Rikubetsu	0.55	1.04	1326
Saga	-1.03	1.35	2878
Sodankyla	0.17	1.04	608
Tsukuba	-0.16	1.95	1276
Wollongong	-0.25	1.12	8366
Station-to-Station	0.02	0.30	-
All	0.00	1.20	174069
All (land)	-0.02	1.28	131538
All (ocean)	0.07	0.90	42531

Table 12.2: Statistics (bias Δ and scatter σ) of the UoL-FP/TCCON matching and comparison, after the bias correction procedure. Station-wise biases are calculated as the mean and standard deviation of the per-station means. Information on the sites themselves can be found at the official TCCON website ("TCCON - Tccon-Wiki" 2018).

geometry was close to nadir with a 1° viewing zenith angle, and the solar zenith angle was varied at either 15° or 65°. For each of the two solar geometries, four different aerosol scenarios were chosen which differ in optical depth only. The aerosol mixtures remained the same: a small-type carbonaceous, continental black carbon type, as well as a large-type carbonaceous dusty mixture, corresponding to types 5c and 4c from R. Kahn et al. (2001), respectively. Height profiles of both mixtures are parametrised as Gaussians with respect to altitude, the small one is centred at 2 km with a width (standard deviation) of 1.5 km, whereas the large type is centred at 1.5 km with a width of 1 km. Additionally, each scenario also features a cirrus cloud aerosol type at 10 km altitude with 0.5 km width. The residuals for the most challenging case (AOD: 0.5^3) for both aerosol mixtures, cirrus AOD: 0.15, solar zenith angle: 65°), are presented in Figures 12.7 to 12.9. In those figures, the instrument-independent high-resolution residuals are assessed through the relative values of the Stokes components (see Equation 11.43). The radiance differences after multiplication with the high-resolution solar spectrum, relevant entries of an instrument Müller matrix and convolution with an OCO-2 like instrument line shape (Table 11.3) are also shown in the last panel. The Müller matrix coefficients were chosen to be $m_{11} = 0.5$, $m_{12} = 0.24$, $m_{13} = -0.44$, $m_{14} = 0$. For demonstration purposes, five EOFs were chosen for the PCA-based method. The scenarios themselves represent rather extreme cases in terms of full-physics XCO₂ retrievals, as most soundings with aerosol loadings ≥ 0.3 end up being flagged during quality filtering procedures. The same holds for large solar zenith angles. Residuals



Figure 12.7: Residuals of the O₂ A-band for a challenging scenario with high aerosol loadings and large solar zenith angle. The top three panels show the relative point-by-point residuals ΔI_{rel} (as percentage), and Stokes fractions Δq and Δu . They are based on the full SS+MS Stokes vectors and are thus instrument-independent. The last panel shows the total intensity after **S** is multiplied with the Müller matrix components (m_{11} , m_{12} , m_{13}) and subsequently convolved with a Gaussian instrument line shape function.

are generally accentuated in line cores for most bands and methods, mainly due to

³Aerosol optical depth at 750 nm.



Figure 12.8: Same as Figure 12.7, but for the weak CO_2 band. Surprisingly, the linear-*k* method exhibits the smallest residuals, however this can be explained through the dynamic binning strategy that the implementation of the other two methods lack.

the way how $I_{\rm rel}$ and the residual Stokes fractions Δq and Δu are calculated. For all three methods and bands, the residuals appear mostly flat, despite the rather heavy aerosol loading, showing that the various approaches to deal with spectrally varying scattering parameters work well. A small residual slope for the Stokes fractions qand u is seen for the LSI method in the O₂ A-band. Looking at the residuals of the convolved radiances, it becomes clear that the shape of the residuals is significantly different for the various fast RT methods.

Using linear error analysis, these spectral residuals can be translated into XCO₂ errors, as was done in Section 11.5. Here, the gain matrix was calculated using the Jacobians obtained by the PCA-based method, which provides the most accurate reconstruction in that regard. Rather than reporting only one number, however, the contributions of the bands to the total error are listed in the following tables. Along with the XCO₂ errors, the interquartile ranges of $|\Delta I_{\rm rel}|$ are listed as well. Table 12.6 is a summary of the XCO₂ errors that includes also a sum of all partial contributions for each band.

While it is difficult to draw universal conclusions from this very small number of


Figure 12.9: Same as Figure 12.7, but for the strong CO_2 band.

scenarios, Tables 12.3 to 12.5 suggest the following. Reconstruction accuracies tend to be best for the weak CO_2 band, which is not surprising given the smaller range of gas optical depths occurring in that band. It is surprising, though, that for that particular band, the linear-k method performs so well and is easily outclassing even the PCA-based method using five EOFs. This could be attributed to the binning strategy, which is static for both PCA-based and LSI methods. In terms of residuals, both the O_2 A-band and the strong CO_2 band show similar relative magnitudes, however the XCO₂ errors are larger for the strong CO_2 band. Radiance errors in the O_2 A-band cause XCO_2 errors due to errors in the retrieved surface pressure, whereas in the CO_2 bands, the mismatch of spectral lines is more directly responsible for a difference in CO_2 concentrations.

The general behaviour remains as expected, with increasing aerosol loadings as well as increasing solar zenith angle, the residual magnitudes increase as well. Across most scenarios and bands, the PCA-based method performs best in terms of both residuals and XCO_2 errors, where as LSI and linear-*k* are similar. For scenarios 5 and 8, the XCO_2 errors for PCA-based method stem mostly from the weak CO_2 band, and these two cases are the only ones in which the linear-*k* method outperforms the PCA-based approach. O'Dell (2010) does not state XCO_2 errors as a result of the approximation error, and the numbers reported in Hasekamp et al. (2008) are in line with the values listed in Table 12.6.

	Aerosol	0.05	/ 0.05	0.50	/ 0.05	0.05	/ 0.50	0.50	/ 0.50
	Cirrus	0.	01	0.	01	0.	01	0	.15
	SZA	15°	65°	15°	65°	15°	65°	15°	65°
	Scenario #	1	2	3	4	5	6	7	8
	linear-k	0.79	1.30	0.95	1.69	2.17	2.56	2.84	4.09
$\Delta I_{ m rel}$	LSI	1.51	3.67	1.99	5.56	3.06	8.21	2.35	5.44
	PCA	0.04	0.08	0.12	0.25	0.13	0.28	0.10	0.24
	linear-k	0.09	2.53	0.23	5.57	0.21	4.59	0.60	7.46
Δq	LSI	0.23	3.39	0.26	6.42	0.40	5.94	0.29	10.53
	PCA	0.01	0.11	0.01	0.33	0.02	0.21	0.03	0.28
	linear-k	0.04	0.93	0.09	2.04	0.08	1.68	0.23	2.74
Δu	LSI	0.09	1.24	0.10	2.36	0.16	2.17	0.12	3.86
	PCA	0.00	0.04	0.00	0.12	0.01	0.08	0.01	0.10
	linear-k	2.41	2.88	5.35	4.13	8.62	8.28	9.50	10.52
$\Delta I_{\rm rel}$ (conv.)	LSI	1.96	5.76	2.77	9.40	4.37	13.38	4.13	10.31
	PCA	0.08	0.24	0.22	0.55	0.29	0.70	0.23	0.42
ΔXCO_2	linear-k	-0.04	0.02	0.07	-0.12	-0.49	0.33	1.22	-0.68
	LSI	-0.01	-0.51	-0.10	-0.65	-0.70	1.12	0.33	-0.56
	PCA	0.00	-0.01	0.00	0.02	0.01	0.05	0.00	-0.04

O₂ A-band

Table 12.3: The first three lines in the table header define the scenario. Aerosol optical depths for small and large-types, cirrus cloud optical depth, and solar zenith angle. For each of the eight scenarios, the interquartile range of the absolute value of the residuals are listed. $\Delta I_{\rm rel}$ and $\Delta I_{\rm rel}$ (conv.) are in units of 0.01%, Δq and Δu are in units of $1 \cdot 10^{-5}$. XCO₂ errors are in ppm.

Performing linear error analysis on eight scenarios already reveals that for the globally most common aerosol configuration (optical depth ≤ 0.1), the estimated errors are very small and thus will most likely not have any significant effect on the final XCO₂ or any derived quantities (i.e. fluxes). Especially considering that bias correction procedures affect the retrieved concentrations to a larger extent.

All three discussed fast RT methods were fully implemented into the UoL-FP retrieval scheme. This allows running full end-to-end retrievals with one consistent algorithm where only the fast RT method is changed between runs. Such an analysis has not been published for XCO_2 retrievals from the SWIR so far.

The selection of soundings to retrieve was based on a compromise between global coverage, distribution of measurement modes (land nadir, land glint, and ocean glint), as well as data volume (number of orbits). At the end, soundings (extracted from the

Aerosol		0.05 / 0.05		0.50 / 0.05		0.05 / 0.50		0.50 / 0.50	
	Cirrus	0.	01	0.	01	0.	01	0	.15
	SZA	15°	65°	15°	65°	15°	65°	15°	65°
	Scenario #	1	2	3	4	5	6	7	8
	linear-k	0.03	0.05	0.10	0.13	0.07	0.09	0.06	0.14
$\Delta I_{ m rel}$	LSI	0.32	0.95	0.64	2.39	0.69	3.35	0.99	4.56
	PCA	0.10	0.25	0.48	1.50	0.83	2.74	0.95	2.18
	linear-k	0.00	0.01	0.01	0.04	0.01	0.03	0.02	0.05
Δq	LSI	0.02	0.25	0.14	2.01	0.64	0.81	1.42	1.40
	PCA	0.01	0.13	0.04	1.26	0.09	1.13	0.12	1.54
	linear-k	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.02
Δu	LSI	0.01	0.09	0.05	0.74	0.24	0.30	0.54	0.51
	PCA	0.00	0.05	0.02	0.46	0.03	0.41	0.05	0.56
	linear-k	0.05	0.08	0.11	0.16	0.10	0.15	0.26	0.49
$\Delta I_{\rm rel}$ (conv.)	LSI	0.28	0.79	0.50	1.69	0.51	3.34	0.75	3.79
	PCA	0.10	0.28	0.59	1.86	0.80	2.70	0.79	1.79
ΔXCO_2	linear-k	0.02	0.07	-0.02	0.02	-0.03	-0.03	0.12	-0.37
	LSI	-0.04	0.09	-0.03	-0.33	-0.24	0.21	0.31	0.17
	PCA	0.00	-0.01	0.02	0.08	-0.02	0.54	0.09	0.62

Weak CO₂

Table 12.4: Same as Table 12.3, but for the weak CO_2 band.

Strong	CO_2
--------	-----------------

	Aerosol	0.05 / 0.05		0.50 / 0.05		0.05 / 0.50		0.50 / 0.50	
	Cirrus	0.	01	0.	01	0.	01	0.	15
	SZA	15°	65°	15°	65°	15°	65°	15°	65°
	Scenario #	1	2	3	4	5	6	7	8
	linear-k	0.34	0.66	1.48	2.69	1.68	2.98	1.74	3.97
$\Delta I_{ m rel}$	LSI	0.21	0.60	0.63	3.07	0.56	4.89	1.31	4.74
	PCA	0.07	0.14	0.33	1.37	0.35	2.49	0.71	2.56
	linear-k	0.04	0.60	0.21	6.91	0.24	6.96	0.56	12.42
Δq	LSI	0.11	1.96	1.81	20.08	2.21	13.41	3.36	20.22
_	PCA	0.07	0.88	1.07	12.97	1.33	10.96	2.25	15.72
	linear-k	0.02	0.22	0.08	2.54	0.09	2.55	0.22	4.59
Δu	LSI	0.04	0.72	0.69	7.37	0.84	4.93	1.27	7.42
	PCA	0.03	0.32	0.41	4.76	0.50	4.02	0.85	5.77
	linear-k	0.37	1.09	1.64	4.47	1.88	5.29	4.33	17.55
$\Delta I_{\rm rel}$ (conv.)	LSI	0.25	1.99	0.90	6.85	1.23	10.35	4.10	14.41
	PCA	0.10	0.33	0.38	1.87	0.58	5.03	1.39	4.39
	linear-k	-0.02	-0.07	-0.07	-0.27	-0.14	-0.16	-0.44	0.64
ΔXCO_2	LSI	0.00	0.01	-0.03	-0.08	-0.04	-0.19	0.06	-0.39
	PCA	0.00	-0.01	0.00	-0.06	-0.01	-0.03	0.09	0.09

Table 12.5: Same as Table 12.3, but for the strong CO_2 band.

	Aerosol	0.05 / 0.05		0.50 / 0.05		0.05 / 0.50		0.50 / 0.50		
	Cirrus	0.	01	0.	0.01		0.01		0.15	
	SZA	15°	65°	15°	65°	15°	65°	15°	65°	
	Scenario #	1	2	3	4	5	6	7	8	
	linear-k	-0.04	0.02	0.07	-0.12	-0.49	0.33	1.22	-0.68	
O_2 A-Band	LSI	-0.01	-0.51	-0.10	-0.65	-0.70	1.12	0.33	-0.56	
	PCA	0.00	-0.01	0.00	0.02	0.01	0.05	0.00	-0.04	
	linear-k	0.02	0.07	-0.02	0.02	-0.03	-0.03	0.12	-0.37	
Weak CO ₂	LSI	-0.04	0.09	-0.03	-0.33	-0.24	0.21	0.31	0.17	
	PCA	0.00	-0.01	0.02	0.08	-0.02	0.54	0.09	0.62	
	linear-k	-0.02	-0.07	-0.07	-0.27	-0.14	-0.16	-0.44	0.64	
Strong CO ₂	LSI	0.00	0.01	-0.03	-0.08	-0.04	-0.19	0.06	-0.39	
Ũ	PCA	0.00	-0.01	0.00	-0.06	-0.01	-0.03	0.09	0.09	
Total	linear-k	-0.04	0.02	-0.02	-0.37	-0.66	0.14	0.89	-0.41	
	LSI	-0.06	-0.41	-0.16	-1.06	-0.98	1.14	0.70	-0.78	
	PCA	-0.01	-0.02	0.02	0.04	-0.03	0.57	0.18	0.67	

XCO₂ Errors [ppm]

Table 12.6: Summary of the XCO_2 errors per band from Tables 12.3 to 12.5, including the total error as the sum of all contributions.

OCO-2 L2 Lite data, version 8r) from every second day of May 2016 were considered. From these, only every 10th OCO-2 frame is kept. In this subset, the number of ocean glint measurements is still large compared to land ones, so as an additional filter, only every 4th ocean glint sounding is kept. The warn-level classification present in the OCO-2 L2 Lite data is not considered, as the maximal range of geophysical parameters is desired. This produces a sounding list of roughly 130 000 soundings from 222 different orbits, where the three different modes (land glint, land nadir and ocean glint) are approximately equally represented (~50 000, ~43 000 and ~36 000). The resulting selection is therefore not fully representative of the actual OCO-2 sampling, however the aim here was not to produce a representative set of OCO-2 measurements, but rather to cover as much of the geophysical parameter space as possible, while keeping the number of soundings small. Figure 12.10 shows the global distribution of the sounding locations.



Figure 12.10: Distribution of selected OCO-2 soundings from May 2016, gridded to a rough 2° by 2° grid. The set covers most landmasses and oceans apart from Antarctica and the southern oceans, with some noticeable gaps near the Indian ocean, Patagonia, tropical Africa and patches in Eastern Asia and Central Northern America.

To analyse the results, the retrievals are sub-setted so that only soundings which were successfully completed by the algorithm, regardless of their outcome or quality. For the LSI and linear-k methods, the completion rate⁴ was over 99%, the retrievals utilising the PCA-based method however exhibited a lower rate with ~97.5%. The difference in completion rate is due to the eigenproblem solver DSYEV returning an error during the computation of the eigenvectors and eigenvalues (see Equation 11.22, Page 161) in some cases.

Before the retrieval results are assessed, the computation performance of the methods can be compared. The total execution time of each retrieval is measured via the UNIX date command. This way of measuring the runtime does not fully correspond to CPU time, as any delay by I/O processes are included in the total runtime. In Figures 12.11 and 12.12 the runtimes are visualised per iteration (runtime divided by the sum of iterations), and show the expected trend - the two runs using the PCA-based method are the slowest, and linear-k is the fastest. Since all retrievals are performed using the same set-up (within the respective land/ocean group), the width of each distribution represents the variability of the runtime due to various influences, such as hard-drive access times or current CPU load on the computation node. Table 12.7 lists the per-iteration runtimes as well as the total integrated runtime for the subset of all retrievals that converged in all of the four runs.



Figure 12.11: Retrieval runtimes per iteration in minutes, showing that linear-k is, on average, 18% shorter than LSI, which is about 25% shorter than the PCA-based method using one EOF. Using five EOFs increases the runtime by roughly 65% compared to using only one EOF.

Apart from the number of utilised CPU hours, the retrievals themselves show significant differences, as a quick look at the number of iterations and number of divergent steps reveals the influence of the fast RT method. About 25 % of ocean retrievals, and 1 % of land retrievals exhibit a different number of iterations between the four runs. In terms of convergence rates, linear-k has the lowest value with 81 % (88 % for land),

⁴Fraction of successful executions of the retrieval algorithm compared to the total number of executions.



Figure 12.12: Same as Figure 12.11, but for land soundings. The overall picture is the same, however the distributions feature more prominent tails.

	Oc	ean	Land			
	per iteration	total runtime	per iteration	total runtime		
PCA(5)	4.09 ± 0.75	8588	3.49 ± 0.92	17508		
PCA(1)	2.47 ± 0.62	5184	2.42 ± 0.60	11965		
LSI	1.86 ± 0.33	3933	2.19 ± 0.61	10998		
linear-k	1.54 ± 0.38	3561	1.73 ± 0.61	9464		

Table 12.7: Runtimes per iteration (in minutes) as well as total runtime in hours for roughly 31 000 (ocean) and 85 000 (land) retrievals. A full line-by-line computation takes about 45 min per iteration.

whereas the other three runs have a 88 % (92 %) rate. Looking at the distribution of iterations and divergent steps (Figure 12.13) shows that more retrievals using the linear-k method encountered a fourth divergent step, at which the retrieval is aborted and counted as non-convergent.



Figure 12.13: Distributions of numbers of iterations (left) and numbers of divergent steps (right) for ocean soundings. The retrievals were set up to abort after more than 3 divergent steps.



Figure 12.14: Same as Figure 12.13 but for land soundings.

The fit quality, characterised by χ^2 , is similar for all three methods for the weak and strong CO₂ bands, but diverges slightly for the O₂ A-band. Here, only those retrievals are considered for which the respective method yielded a convergent retrieval. Measurements near the South Atlantic Anomaly have been excluded as well. For the O₂ A-band, the LSI method exhibits on average slightly lower χ^2 values, and linear-k retrievals are on average higher. Box-plot representations of the χ^2 distributions are shown in Figures 12.15 and 12.16, and highlight that the fit quality of the ensemble of converged retrievals only significantly changes between runs for the O₂ A-band.



Figure 12.15: Box-plots illustrating the χ^2 values for the three bands and the four different runs (ocean retrievals).

The increased χ^2 for land retrievals can be explained through the equally increased SNR. While one would intuitively expect the SNR for ocean glint measurements to be higher, the actual measurements show the opposite. This is the result of the polarisation anomaly mitigation strategy (Crisp et al. 2017), in which the OCO-2 spacecraft is rotated by 30° (yaw angle) for glint observations. The mitigation



Figure 12.16: Same as Figure 12.15 but for land retrievals.

strategy was a response to a design mistake due to which the OCO-2 instrument's slit alignment was mistakenly constructed to attenuate the ocean glint signal. While this modified yaw angle has very little effect for mostly unpolarising land surfaces, the SNRs for ocean glint measurements are recovered to values between 200 and 600.

With increasing SNR, the impact of the residual waveforms to compensate for systematic errors in the forward model becomes more important. Ideally, there should be no relationship between the SNR and the fit quality, as the reduced χ^2 statistic takes into account the expected noise level. Looking at 2D histograms between these two quantities in Figure 12.17 reveals that there is a relationship for both land and ocean retrievals. For land retrievals, however, the relationship is more striking as the range of SNRs is greater.

Summarising the statistics regarding convergence, the linear-k runs show a higher number of non-converging retrievals as well as a noticeably wider distribution of χ^2 values for the O₂ A-band, for both land and ocean scenes.

Further analysis of the retrieved quantities is performed on a subset of data, where only those retrievals are included, for which the convergence criteria was met for all four runs. This leaves the final subset with \sim 32 000 land nadir, \sim 36 000 land glint, and \sim 25 000 ocean glint retrievals.

A simple way of assessing the differences of the runs globally is to calculate averages for latitudinal bins. Using 4° bins, means for XCO_2 are calculated for every run independently, and the latitudinal means are then subtracted from the bin median across the four values. Figure 12.18 shows these differences and highlights that the run using linear-k is consistently an outlier for all measurement geometries. For both land and ocean glint modes, the linear-k retrievals diverge at northern latitudes past 20° . For the land nadir observation mode, the retrievals show the largest differences at high northern latitudes, the tropics and southern mid-latitudes. In Figure 12.18,



Figure 12.17: Fit quality (χ^2) versus signal-to-noise ratio for the PCA(5) run.

the two runs using the PCA-based method are almost indistinguishable, whereas the land retrievals utilising the LSI show differences in regions similar to the linear-k runs, albeit of smaller magnitude. The differences in the latitudinal means confirm the results from the linear error analysis (Table 12.6), which predicted errors up to 1 ppm, with different signs.

The differences between the four runs are visualised in Figure 12.19. This matrix of maps is to be read in the following way: in the outer rows and columns, maps of the retrieved XCO₂ are displayed - land scenes at the top row and rightmost column, ocean scenes at the leftmost column and bottom row. XCO₂ values are plotted using a sequential colour map, darker colours representing higher concentrations. These absolute concentration maps share the same colour mapping, the first map in the top row as well as leftmost column feature an indicator which state the limits of the colour map to give an indication of the value range. The inner maps of the matrix show the pairwise differences between the four runs, where the (i, j)-th (row, column) map of the inner matrix is calculated as $XCO_2(i) - XCO_2(j)$. For example, the map of the first row / last column (1,3) of the inner matrix is the difference between the PCA(5) run and the linear-k run. The differences are calculated on a per-sounding basis before the data is then gridded (4° grid cell size) and plotted on the map. For the pairwise differences, a divergent red-blue colour map is used with red representing



Figure 12.18: Latitudinally binned XCO_2 displayed as differences to the median of all four values of that bin (4° spacing). The shaded areas represent the standard error of the latitudinal bin.

positive values, blue representing negative values, and white representing the central data bins $-\Delta x < x < \Delta x$.

Looking at the ocean glint measurement mode (lower triangle of the map matrix), the differences between the two PCA runs and the LSI run behave very similarly, apart from the magnitude. These three runs show pairwise differences in the southern tropics and southern sub-tropics. A similar band is seen at higher northern latitudes. In addition, the PCA(1) and PCA(5) runs show differences in the Atlantic ocean at the dust outflow regions, hinting at a dependence on aerosol optical depth. The linear-k run behaves differently, as the distinct band south of the equator is not present, however differences at the dust outflow regions suggest a dependence on aerosol loadings as well. The two runs using the PCA-based method show a surprisingly low overall difference - for ocean scenes, the 95% of retrievals exhibit a difference less than 0.033 ppm.

For land scenes, the magnitude of differences is roughly two to three times larger than for the ocean glint measurements. Intuitively, this is expected as the SNR for land-type measurements is, on average, higher than for ocean glint observations.

The maps in Figure 12.19 show the spatial patterns of the differences, however also



Figure 12.19: Matrix of pairwise differences between the different runs, separated into land (upper right) and ocean (lower left) retrievals. Each map in the inner rectangle (blue colour-scale) represents the difference between (row) and (column), and the inset text box shows the range of XCO_2 difference for each map.



Figure 12.20: Differences between the respective retrieval runs to the median (like in Figure 12.19) as functions of the prior aerosol optical depths (for small and large type aerosol mixtures) and the geometric airmass factor. The shaded areas represent the standard error of the ΔXCO_2 values in the respective AOD or AMF bin.

obfuscate the geophysical parameters that cause them. The two main parameters associated with RT approximation errors are the viewing geometry and aerosol loading. Plots of the differences as functions of these parameters are shown in Figure 12.20 for the three observation modes. The most distinct bias for the linear-k method appears for ocean glint retrievals as a function of the small type aerosol optical depth. The dependence for land measurements is not as clear, as the differences become smaller for larger AOD. For the large type aerosol mixture, the bias is similarly clear for ocean glint measurements - the retrieved XCO₂ diverges with larger AOD for the PCA-based and LSI methods. The AOD-dependence for land scenes is not as straightforward, as for example the LSI runs diverge from the PCA runs at lower AODs. For land nadir retrievals, the largest linear-k differences are seen at larger airmass factors.

To investigate the cause of the large departures of the retrieved XCO_2 for the linear-k retrievals, the retrieved aerosol optical depths for the small-type, large-type and cirrus mixtures are plotted as functions of prior AODs (and latitude for the retrieved cirrus)



Figure 12.21: Similar to Figure 12.20 (and same colour coding), but with retrieved AODs rather than the retrieved XCO₂. For the last row (cirrus AODs), the retrievals were binned according to their latitude, as the prior is the same for every scene.

in Figure 12.21. It becomes immediately clear that the XCO_2 differences inherent in the linear-k retrievals are linked to the retrieved aerosols. Retrieved aerosol optical depths for linear-k retrievals substantially differ from the other 3 runs for most prior scenarios. Land glint and land nadir behave very similarly in that regard, whereas the ocean glint retrievals are distinct, especially for the retrieved cirrus AOD.

A possible explanation as to why the retrievals using linear-k method show such different results points to the way that the method handles atmospheric weighting functions (or Jacobians). As established in Section 11.3, linear-k approximates the MS contributions and weighting functions, and adds them to the exactly calculated SS contributions of the Stokes vectors. The method itself differs from LSI and the PCA-based method (see also Section 11.6) in that the other two methods use a low-accuracy MS calculation as a basis for the correction procedure. Linear-k lacks these low-accuracy calculations and relies only on linear interpolation to obtain the MS contributions to the weighting functions at every spectral point. For most weighting functions, this appears to be a feasible procedure, as long as the MS contributions to the total weighting function are not dominant. Since the UoL-FP algorithm uses

full aerosol extinction profiles (for each mixture), the Jacobians for these state vector elements are mostly MS dominated. The reconstruction accuracy for these Jacobians is thus much lower when compared to calculations using LSI and the PCA-based method. After the first iteration, the aerosol Jacobians drive the state vector into a slightly different direction and thus result in a different final state vector.

12.4 Conclusions

The OCO-2 retrievals in Section 12.2 show that the UoL-FP algorithm was successfully adapted to the OCO-2 instrument. When validated against the TCCON ground network, the retrieved XCO₂ exhibit standard deviations of $\sigma_{\text{land}} = 1.28 \text{ ppm}$ and $\sigma_{\text{ocean}} = 0.90 \text{ ppm}$, which compare favourably against published values in Wunch et al. (2017) ($\sigma_{\text{ocean}} = 1.42 \text{ ppm}$, $\sigma_{\text{land-glint}} = 0.90 \text{ ppm}$, $\sigma_{\text{land-nadir}} = 1.50 \text{ ppm}$); keeping in mind that Wunch et al. (2017) used daily aggregates whereas individual TCCON/OCO-2 collocations were used in Section 12.2.

Implementing three different fast RT schemes into the same full-physics algorithm has revealed interesting results in the final, retrieved XCO₂. Considering only the ocean retrievals, more than 98% of the converged retrievals are less than 1 ppm apart, however this percentage drops when considering land scenes (79%). Both magnitude and spatial pattern (see Figure 12.19) of the differences imply that the choice of fast RT method does have a significant and systematic impact on the retrieved atmospheric CO₂ concentrations. As pointed out by Chevallier et al. (2007), even sub-ppm regional biases can significantly affect fluxes obtained from the inversion of satellite data.

The analysis in Section 12.3 does not factor in any aspect regarding bias correction procedures (Wunch et al. 2017), which are done for most scientific data releases. It is expected that any correction to some ground truth, be it either TCCON or models, have the potential reduce the differences between the four results.

Summary



have investigated various methods of accelerating radiative transfer calculations in Part III of the thesis. The general motivation for fast RT methods was laid out as a result of logistic considerations regarding the large number of measurements from current and future space-based missions. I have to emphasise that the discussed fast RT schemes do not improve or alter the established ways of solving the RT equation. Rather, they reduce the number of needed RT computations by replacing the thousands of line-by-line calculations by a small set of representative ones. The results of these representative calculations are then mapped onto all other spectral points to yield an approximation to a high-accuracy line-by-line calculation. Speed-ups of two orders of magnitude are common, and thus, these techniques make the processing of large amounts of retrievals feasible in the first place. Approximations to full line-by-line calculations are required to be highly accurate. Despite the fact that radiances are reconstructed with accuracies better than 0.1% in most cases, I could demonstrate that the retrieved XCO₂ can show significant differences depending on the utilised fast RT method.

I have reviewed three methods that are presently used in XCO_2 retrievals from the SWIR - LSI, linear-k and the PCA-based method - in Chapter 11. As part of the research for this thesis, I have further developed the PCA-based method in cooperation with V. Natraj⁵. The novel addition to the method is the inclusion of aerosol extinction coefficients $q^{(sca)}$ into the optical state matrix. Rather than just using the aerosol scattering properties from the band centre, the quantities required for the RT calculations (phase function expansion coefficients and the phase matrix) are reconstructed from the $q^{(sca)}$. This addition solves the issues related to systematic residuals induced by the spectrally varying scattering properties of aerosols. While in the case of the UoL-FP retrieval algorithm, the spectral variation is of a linear nature, the scheme could deal with non-linear wavelength dependencies in broadband spectral ranges. With the PCA-method now being able to intrinsically correct for spectrally varying aerosol properties, I was able to fully implement the method into the UoL-FP algorithm.

⁵Jet Propulsion Laboratory, Pasadena, CA, USA

I demonstrated the capability of the PCA-based method in Section 11.5. For a large set of real GOSAT scenes, I performed full line-by-line RT calculations and compared them to approximations from the PCA-based method. I then analysed the residuals and subsequently translated them into XCO_2 errors (specific for the GOSAT instrument) by virtue of the linear error analysis framework. From the analysis, I conclude that XCO_2 errors induced by the fast RT method are predicted to be below 0.6 ppm for both ocean or land scenes when two or more EOFs are used. So far, this analysis is the most in-depth investigation of this fast RT method and has been published in **Somkuti**, Boesch, Natraj, et al. (2017a).

Linear error analysis only provides an accurate estimate of the retrieval error if the retrieval itself is highly linear. For XCO_2 retrievals, this is not always the case. Thus, to quantify the impact of the fast RT scheme, I conducted a more involved exercise. Both LSI and linear-k methods were implemented into the UoL-FP algorithm to allow for a fully consistent comparison of the three different fast RT schemes. With the standard UoL-FP retrieval set up, I find that linear-k is a consistent outlier when comparing the retrieved XCO_2 . The underlying cause for the large departure from the other runs is related to the atmospheric weighting functions that impact the aerosol extinction profiles. Due to the smaller reconstruction accuracy for weighting functions in which the MS contributions constitute a much larger part compared to the SS ones, the aerosol-related Jacobians drive the retrieval into a different region of the state vector space. Because aerosols and XCO_2 are correlated to some degree, the final XCO_2 value as well as the retrieved aerosol profiles depart significantly from the other three runs. To my knowledge, the analysis of XCO₂ retrievals using three different fast RT approaches that are consistently integrated into one single algorithms has not been published.

Final Remarks and Outlook

To conclude the thesis, I will summarise the conducted research and key findings. The two main research areas, solar-induced chlorophyll fluorescence and fast radiative transfer, revolve around the measurement of atmospheric carbon dioxide - in particular the forward model. Space-based measurements of the column-averaged dry-air volume mixing ratio of CO_2 can be used inform flux inversion systems, which then allow the deduction of regional carbon flux budgets on a global scale. XCO_2 measurements are not only of great interest to carbon cycle science, but also to the monitoring of anthropogenic emissions to facilitate the meeting of emission targets set by the COP21 / Paris Agreement. It is imperative that XCO_2 retrievals meet the precision and accuracy requirements needed in order to fulfil these goals. Both SIF and fast radiative transfer are two components of the utilised retrieval algorithm, which have a significant impact on the final XCO_2 value and therefore need to be understood and characterised.

Using an established method for the retrieval of solar-induced chlorophyll fluorescence, I have generated a SIF data-set spanning the GOSAT record from April 2009 until December 2016. SIF is present in the O_2 A-band, which is utilised in XCO_2 retrievals. Any contribution to the forward model through a non-zero SIF radiance prior will therefore influence the retrieved CO_2 concentrations. Due to the seasonal nature of vegetation and SIF, this influence thus exhibits a similar seasonal cycle. I have processed four years of GOSAT XCO_2 retrievals, both with and without the SIF influence, and directly compared them. I observe significant differences up to about 0.5 ppm when averaged over subcontinental-scale regions. The differences are explained due to a change in retrieved surface pressure as well as a change in retrieved aerosol extinction. Unfortunately, it is not trivial to translate the spatio-temporal pattern of the differences to estimate the effect on flux inversions without actually performing them.

Apart from being a present contribution in the O_2 A-band that impacts XCO_2 retrievals, SIF is of great interest due to its inherent relationship to photosynthesis and plant status. I have revisited the empirical relationship between SIF and GPP via a process-based model (CASA) and measurements (FluxNet). I find that the derived

factor relating SIF and CASA-GPP is comparable to the values published in literature. Interestingly, however, I find that the seasonal cycle of GPP in tropical regions is out of phase when compared to SIF. This is apparently already known in the OCO-2 flux inversion community, yet unpublished. They find that posterior fluxes over tropical regions are equally phase shifted compared to the CASA prior. When deriving the same factor between GPP and SIF using FluxNet measurements, I find a significantly lower value than the CASA-derived one. Considering the average footprint size of flux towers (a few km²), I conclude that GOSAT's sampling pattern is the main cause of the difference, and I show this by investigating the sampling bias using CASA model data. As a final item, I study the intense 2012 North American drought using the SIF retrievals. I show that this large-scale drought event is evident in the SIF measurements, regardless of the used micro-window, polarisation or calibration technique. 2012 was special in that an early warm spring not only caused larger productivity before the onset of summer, but was also responsible for exacerbating the drought due to early depletion of water stored in the soil. I can observe the early spring through SIF, while also relating the SIF anomaly to anomalous carbon uptake.

For the second part of the thesis, I investigate fast radiative transfer methods. Being crucial for making full-physics XCO_2 retrievals possible in the first place, I give an overview of three state-of-the art methods. My main focus was then the PCA-based method. I contributed to the further development of this particular fast RT scheme by introducing a way to deal with spectrally varying aerosol scattering properties. As all approximation methods come with reconstruction errors, I have conducted a detailed study on the errors induced by the PCA-based method and applied linear error analysis to estimate the error on the retrieved XCO_2 . This work was published in **Somkuti**, Boesch, Natraj, et al. (2017a). While linear error analysis can produce realistic estimates, the often non-linear nature of XCO₂ retrievals cannot be taken into account this way. Thus, I decided to commit to the arduous task of implementing the three methods natively into the UoL-FP algorithm. This allowed to perform retrievals in a consistent manner by just changing the RT portion of the forward model, and leaving all other aspects unchanged. To demonstrate the effect that the fast RT methods have on the final XCO_2 value, I processed several hundred thousand OCO-2 soundings which took more than 70 000 CPU hours. The outcome of the analysis was somewhat surprising, as the linear-k method stood out as a significant outlier. While being the fastest of all three methods, it is also the technique with the least accurate reconstruction of multiple-scattering related atmospheric weighting functions. As a result, the retrieved aerosol properties diverge from those for the other methods. It is conceivable that this result is only particular to the UoL-FP algorithm, which retrieves a full extinction profile for every aerosol mixture. The differences between the methods are generally less than 1 ppm, making them significant - especially since

the differences are relatable to geophysical parameters such as aerosols or viewing geometry. Similar to the discussion of the influence of SIF on the XCO_2 retrievals, the same arguments hold. Sub-ppm biases can significantly change the outcome of flux inversions. But also here I must reiterate that bias correction procedures can alter the results and potentially reduce the differences and lessen the impact of the inaccuracies of the fast RT methods. The PCA-based method results in the smallest radiance errors for scenes with high aerosol loadings, making it particularly suitable for XCO_2 retrievals near cities with a high degree of air pollution. Cities will be a focus of future space-based missions to aid the monitoring of anthropogenic emissions.

From the results obtained in this thesis, I can formulate three main items of relevance that are worth investigating further.

- Effect of SIF and fast RT schemes on bias corrections Both Chapter 8 and Chapter 12 explore the effect of an aspect of the forward model on the final, retrieved XCO₂. While the results show that neither aspects can be neglected, it is unclear how these differences would impact the bias correction procedures which are routinely performed for public data releases. I expect the differences to be reduced as the XCO₂ data would be (per definition) corrected to match up with some (ground) truth. However, it would not be surprising if there would be differences left, and it would be interesting to see if either the fast RT method or the non-zero SIF prior have any lasting effect on XCO₂ biases.
- Future of Full-Physics XCO₂ Retrievals I have shown that the PCA-based method can provide the best approximation to full line-by-line MS calculations, when compared to two other contemporary and established models. The accuracy, however, comes at a computational cost. For the full GOSAT record with a few million cloud-free measurements, this is not a big issue given the high-performance computing resources that the University of Leicester has access to. The data volume of OCO-2 is about a hundred times larger, meaning that it would take approximately a day⁶ to process ~60 000 cloud-free measurements taken on a single day, even with a quicker fast RT scheme (such as linear-k). Future satellite missions will further increase the data volume to be closer to ~10 000 000 measurements per day, at which point the full-physics scheme must be revisited or replaced, unless the computing resources are scaled up by the same amount. Machine learning techniques could provide an interesting approach, in which a set of full-physics retrievals can be used to train a classifier. The classifier could then produce the XCO₂ value very quickly, given the

⁶Given the currently available computing resources at the University of Leicester and assuming maximum utilisation

measured radiances and other meteorological parameters. The issue is a rather pressing one, as space-based retrievals will play a big role in the monitoring and attribution of anthropogetic CO_2 emissions in support of COP21 (Paris Agreement). European contributions to emission monitoring in the near future are the MicroCarb mission (Pascal et al. 2017), to be launched in 2020, as well as the proposed Sentinel 7 mission, which is actually a constellation to allow for short revisit times.

• Fluorescence for Monitoring Vegetation Health In this thesis, I have exclusively focused on GOSAT for the retrieval of SIF. While many of the lessons learned can be generalised to other and future satellite missions, GOSAT was not designed to measure surface properties. As such, the GOSAT sampling pattern has always been a limiting factor to study the terrestrial biosphere. However GOSAT, along with GOME-2, have provided a historical set of measurements spanning over 6 years, and are still ongoing (as of April 2018). More recent (OCO-2, Sentinel 5-P) and future missions (FLEX) which can provide SIF measurements on a much denser spatial sampling need to be compared to each other and the long-term GOSAT record. This comparison however, needs to be comprehensive and fully taking into account the footprint extent of all involved instruments. Working exclusively on gridded L3 data, which is very common in the land surface community, is not sufficient in this case.

APPENDICES

Appendix A

The Discrete Ordinates Method

While a number of simplifications were applied to yield the scalar RTE in Equation 10.11 (Page 146), it is still an integro-differential equation in $I(\theta, \phi, \tau)$. The phase function P can be further simplified under the assumption of an optically isotropic and homogeneous medium. Rather than being a function of both incidence angle and outward scattered angle, P is only dependent on the scattering angle Θ :

$$\cos \Theta = \cos \theta' \cos \theta + \sin \theta' \sin \theta \cos(\phi - \phi'). \tag{A.1}$$

This specific structure of $P(\Theta)$ makes the so-called *discrete ordinates method* viable. Already laid out in Chandrasekhar (1960), the solutions to the RTE can be approximated by replacing the integral in the emission term by a sum. In this section, the implementation for the LIDORT radiative transfer model (Spurr 2008) is outlined, which uses the discrete ordinate method to calculate approximate solutions to the RTE.

The phase matrix $P(\Theta)$, now a function of the scattering angle Θ only, can be expanded into a finite series using Legendre polynomials P_l :

$$P(\Theta) = \sum_{l=0}^{LM} \beta_l P_l(\cos \Theta).$$
(A.2)

The equality only strictly holds if the number of Legendre moments $LM \rightarrow \infty$. In practical terms, many phase functions are well-represented with LM being on the order of a few hundred or a few thousand. Following e.g. Thomas et al. (2002), the *addition theorem for spherical harmonics* can be exploited to rewrite the Legendre polynomials in Equation A.2:

$$P_{l}(\cos \Theta) = P_{l}(\mu')P_{l}(\mu) + 2\sum_{m=1}^{l} \Lambda_{l}^{m}(\mu')\Lambda_{l}^{m}(\mu)\cos[m(\phi - \phi')], \qquad (A.3)$$

where again $\mu = \cos \theta \ (\mu' = \cos \theta')$, and

$$\Lambda_{l}^{m}(\mu) = \frac{(l-m)!}{(l+m)!} P_{l}^{m}(\mu), \tag{A.4}$$

with P_l^m being the associated Legendre polynomials¹. Substituting the expression for $P(\Theta)$ from Equation A.4 into Equation A.2 yields

$$P(\Theta) = \sum_{l=0}^{LM} \beta_l \left\{ P_l(\mu') P_l(\mu) + 2 \sum_{m=1}^l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \cos\left[m(\phi - \phi')\right] \right\}.$$
 (A.5)

The next step is to re-arrange Equation A.5 by re-ordering the two sums. First, the large bracket can be expanded to

$$P(\Theta) = \sum_{l=0}^{LM} \beta_l P_l(\mu') P_l(\mu) + \sum_{l=0}^{LM} \sum_{m=1}^l 2\beta_l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \cos\left[m(\phi - \phi')\right],$$
(A.6)

after which one can make use of the property

$$\sum_{l=0}^{LM} \sum_{m=1}^{l} A_{lm} = \sum_{m=1}^{LM} \sum_{l=m}^{LM} A_{lm}$$
(A.7)

to obtain

$$P(\Theta) = \sum_{l=0}^{LM} \beta_l P_l(\mu') P_l(\mu) + \sum_{m=1}^{LM} \sum_{l=m}^{LM} 2\beta_l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \cos\left[m(\phi - \phi')\right].$$
(A.8)

The outer sum of the second term in Equation A.8 starts at m = 1, however that sum can be written to start from m = 0, while also subtracting the corresponding term to keep the equation balanced. Here, some fundamental properties of the associated Legendre polynomials are used: $P_l^0 = P_l$, which leads to $\Lambda_l^0 = P_l$ (see Equation A.4) and thus

$$\sum_{l=m}^{LM} 2\beta_l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \cos\left[m(\phi - \phi')\right] \xrightarrow{m=0} \sum_{l=0}^{LM} 2\beta_l P_l(\mu') P_l(\mu), \tag{A.9}$$

where the second term is just twice the first term in Equation A.8. Hence, Equation A.8 becomes

$$P(\Theta) = -\sum_{l=0}^{LM} \beta_l P_l(\mu') P_l(\mu) + \sum_{m=0}^{LM} \sum_{l=m}^{LM} 2\beta_l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \cos\left[m(\phi - \phi')\right].$$
(A.10)

Finally, the two sums in Equation A.10 can be pulled together by making use of the same trick: when m = 0, the right sum essentially turns into the left one. The final representation of the phase function is

$$P(\Theta) = \sum_{m=0}^{LM} (2 - \delta_{m,0}) \beta_l \cos[m(\phi - \phi')] \sum_{l=m}^{LM} \Lambda_l^m(\mu') \Lambda_l^m(\mu).$$
(A.11)

The form in Equation A.11 is made compact using the Kronecker-Delta symbol $(\delta_{a,b} = 1 \text{ for } a = b, \text{ otherwise } 0)$. The dependence of the phase function has been

¹The *m* in P_l^m is not to be read as an exponent, but as an additional index.

separated into a polar part (μ, μ') and an azimuthal part $(\phi - \phi')$. Equation A.11 has the form of a Fourier cosine series in terms of the *relative azimuth*. Similar to the decomposition of the phase function, the intensity can be formally written as a series as well:

$$I(\theta,\phi,\tau) = \sum_{m=0}^{LM} I^m(\theta,\tau) \cos[m(\phi_0-\phi)].$$
(A.12)

Using this representation, the RTE can be written in terms of the Fourier components $I^m(\theta, \tau)$. Within the integration over all incoming ray angles $\int_{4\pi} d\Omega'$, the phase function is multiplied with the intensity, so using the Fourier representation (Equations A.11 and A.12), the emission part of the RTE reads:

$$\int_{4\pi} \sum_{m=0}^{LM} (2 - \delta_{m,0}) \cos[m(\phi - \phi')] \cos[m(\phi_0 - \phi)] \sum_{l=m}^{LM} \beta_l \Lambda_l^m(\mu') \Lambda_l^m(\mu) \,\mathrm{d}\Omega' \,. \tag{A.13}$$

Equation A.13 finally makes evident why the chosen representation of the phase function and the intensity is useful. The integration over $d\Omega'$ can now be separated fully into azimuthal and polar components. The azimuthal dependence of the RTE is reduced to a product of two cosine terms with two relative azimuth angles:

$$\int_{0}^{2\pi} \sum_{m=0}^{LM} \cos[m(\phi - \phi')] \cos[m(\phi_0 - \phi)] \,\mathrm{d}\phi' \tag{A.14}$$

$$\times \int_{-1}^{1} \sum_{m=0}^{LM} (2 - \delta_{m,0}) \sum_{l=m}^{LM} \beta_l I^m(\theta', \tau) \Lambda_l^m(\mu') \Lambda_l^m(\mu) \, \mathrm{d}\mu' \,. \tag{A.15}$$

Azimuthal integration can be carried out using some elemental trigonometry,

$$\int_{0}^{2\pi} \sum_{m=0}^{LM} \cos[m(\phi - \phi')] \cos[m(\phi_0 - \phi)] \,\mathrm{d}\phi' = \sum_{m=0}^{LM} \pi \cos[m(\phi_0 - \phi)]. \tag{A.16}$$

The RTE for each of the m the Fourier components is then

$$\mu \frac{\mathrm{d}}{\mathrm{d}\tau} I^{m}(\theta,\tau) = -I^{m}(\theta,\tau) + \frac{\varpi}{2} \sum_{l=m}^{LM} \beta_{l} \Lambda_{l}^{m}(\mu) \int_{-1}^{1} I^{m}(\theta',\tau) \Lambda_{l}^{m}(\mu') \,\mathrm{d}\mu' \,. \tag{A.17}$$

Before Equation A.17 can be further simplified, the boundary conditions regarding the intensity and its Fourier components have to be considered. Separating the intensity into down-welling (I^+ , towards the surface) and up-welling (I^- , away from the surface) components, Spurr (2008) lays out three conditions which the intensity field has to satisfy, when the model atmosphere is constructed as a set of N_{lay} optically homogeneous, plane-parallel layers:

1. At the top of the atmosphere (upper boundary of layer n = 1), there is no diffuse intensity:

$$I_1^{m,+}(\theta,\phi,0) = 0.$$
 (A.18)

2. At layer boundaries between layer n - 1 and layer n, continuity of I is imposed:

$$I_{n-1}^{m,\pm}(\tau_{n-1}) = I_n^{m,\pm}(0).$$
(A.19)

3. The lower layer boundary of the last layer $n = N_{\text{lay}}$ (the surface) relates the down-welling and up-welling intensities through the *bi-directional reflectance distribution function* (BRDF):

$$I_{N_{\text{lay}}}^{m,-}(\theta,\phi,\tau_n) = R(\theta,\phi,\theta',\phi') I_{N_{\text{lay}}}^{m,+}(\theta',\phi',\tau_n).$$
(A.20)

In order to solve RTE, the BRDF $R(\theta, \phi, \theta', \phi')$ has to be represented as a Fourier cosine series as well.

To set up a way of numerically solving Equation A.17, the discrete ordinates method relies on the *Gauss-Legendre* method (called *Gauss's formula* in Chandrasekhar (1960), Chapter 22). For a function $f(\mu)$, which is well-approximated through a polynomial of order 2n - 1 or less, the following relationship holds:

$$\int_{-1}^{1} f(\mu) \,\mathrm{d}\mu = \sum_{i=1}^{n} w_i f(\mu_i). \tag{A.21}$$

Thomas et al. (2002) established that the application of the Gaussian quadrature rule as above is potentially problematic near the surface and the top of the atmosphere due to discontinuities. This is the reason why $I(\theta, \tau)$ was separated into up-welling and down-welling components. Using this *Double-Gauss method*, the intensity (or its Fourier components) integration is performed like:

$$\int_{-1}^{1} I(\mu) \,\mathrm{d}\mu = \int_{0}^{1} I^{+}(\mu) + \int_{0}^{1} I^{-}(\mu) = \sum_{i=1}^{n} w_{i} I^{+}(\mu_{i}) + \sum_{i=1}^{n} w_{i} I^{-}(\mu_{i}). \tag{A.22}$$

Note that the weights w_i and the root μ_i are the same for both components of $I(\mu)$. Comparing the Double Gauss method to Equation A.21 one can note the change of the lower integration limit, such that the calculation of the weights is slightly modified.

Replacing the polar angle integration by the Gauss-Legendre sum, the RTE for a Fourier component m is obtained:

$$\pm \mu_i \frac{\mathrm{d}}{\mathrm{d}\tau} I^{\pm,m}(\mu_i,\tau) = \mp I^{\pm,m}(\mu_i,\tau) + \frac{\varpi}{2} \sum_{l=m}^{LM} \beta_l \Lambda_l^m(\pm \mu_i)$$

$$\times \sum_{j=1}^n w_j \left[I^{+,m}(\theta_j,\tau) \Lambda_l^m(\mu_j) + I^{-,m}(\theta_j,\tau) \Lambda_l^m(-\mu_j) \right].$$
(A.23)

Equation A.23 is a system of coupled first-order linear differential equations to be solved for $I^{\pm,m}(\mu_i,\tau)$. The discrete polar angles μ_i (i = 1 ... n) are often referred to

as stream directions. The system consists of 2n equations, and the numerical effort of solving the radiative transfer problem now becomes more apparent. The accuracy of the solution generally increases with the chosen number of stream directions (or streams), but also increases the computational effort, as the system of equations grows in size.

Skipping the details of the solution methods here, the solutions are commonly found through eigenvalue methods following an exponential ansatz. Ultimately, finding the approximate intensities that satisfy the RTE means having to solve an eigenproblem for each and every discrete ordinate, Fourier mode, layer and wavelength.

The abbreviated derivation of the discrete ordinate method presented here skipped a few details. For one, the scalar RTE in Equation 10.11 did not include the scattering of the attenuated solar ray, which merely adds another term

$$Q(\theta,\phi,\tau) = \frac{\varpi}{4\pi} P(\theta,-\theta_0,\phi-\phi_0) I_0(\theta,\phi,\tau)$$
(A.24)

with θ_0 , ϕ_0 and $I_0(\theta, \phi, \tau)$ being the solar polar angle, the solar azimuth angle, and the solar irradiance at the top of the atmosphere, respectively. In terms of the derivation of Equation A.23, this inhomogeneous source term does not change the strategy with which the discrete ordinate method was derived, especially since the homogeneous and inhomogeneous part of the problem are solved separately. The symmetry properties of phase matrices **P** and phase functions *P* was not further discussed, apart from the fact they are dependent only on the scattering angle Θ . Both **P** and *P* are usually defined for the scattering plane, rather than the meridian plane in which the Stokes vector is represented. The phase matrix or phase function therefore needs to be transformed into the correct coordinate system using an appropriate rotation or transformation matrix. Finally, the intensity $I(\mu, \tau)$ is considered to be azimuthally averaged.

Bibliography

- Allen, C. W., et al. 1964. "Astrophysical Quantities." *Physics Today* 17 (3): 66. (Cited on page 37).
- Alonso, L., et al. 2008. "Improved Fraunhofer Line Discrimination Method for Vegetation Fluorescence Quantification." *IEEE Geoscience and Remote Sensing Letters* 5 (4): 620–624. (Cited on page 69).
- Ambartzumian, V. 1936. "The Effect of the Absorption lines on the Radiative Equilibrium of the Outer Layers of the Stars." *Publ. Obs. Astron. Univ., Leningrad* 6:7–18. (Cited on page 149).
- Amoros-Lopez, J., et al. 2008. "Evaluation of Remote Sensing of Vegetation Fluorescence by the Analysis of Diurnal Cycles." *International Journal of Remote Sensing* 29 (17-18): 5423-5436. (Cited on page 94).
- Anav, A., et al. 2015. "Spatiotemporal Patterns of Terrestrial Gross Primary Production: A Review." *Reviews of Geophysics* 53, no. 3 (September 1): 2015RG000483. (Cited on page 126).
- Anderson, E., et al. 2007. *LAPACK Users' Guide. SIAM, Philadelphia, PA, 1992.* (Cited on page 161).
- Andrew D., F., et al. 2006. "FLUXNET and Modelling the Global Carbon Cycle." Global Change Biology 13, no. 3 (August 7): 610–633. (Cited on page 9).
- Baldocchi, D., et al. 2001. "FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities." *Bulletin of the American Meteorological Society* 82, no. 11 (November 1): 2415–2434. (Cited on page 9).
- Beer, C., et al. 2010. "Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate." *Science* 329, no. 5993 (August 13): 834–838. pmid: 20603496. (Cited on page 126).

- Benner, D. C., et al. 2016. "Line Parameters Including Temperature Dependences of Air- and Self-Broadened Line Shapes of ¹²C¹⁶O₂: 2.06-Um Region." *Journal of Molecular Spectroscopy*, New Visions of Spectroscopic Databases, Volume I, 326 (August 1): 21–47. (Cited on page 17).
- Boesch, H., et al. 2006. "Space-Based near-Infrared CO₂ Measurements: Testing the Orbiting Carbon Observatory Retrieval Algorithm and Validation Concept Using SCIAMACHY Observations over Park Falls, Wisconsin." *Journal of Geophysical Research: Atmospheres* 111 (D23). (Cited on page 42).
- Boesch, H., et al. 2013. "HDO/H₂O Ratio Retrievals from GOSAT." Atmos. Meas. *Tech.* 6, no. 3 (March 7): 599–612. (Cited on page 168).
- Brown, L. R., et al. 2000. "Experimental Line Parameters of the Oxygen A Band at 760 Nano Meter." *Journal of Molecular Spectroscopy* 199, no. 2 (February 1): 166–179. (Cited on page 15).
- Bucholtz, A. 1995. "Rayleigh-Scattering Calculations for the Terrestrial Atmosphere." *Applied Optics* 34, no. 15 (May 20): 2765–2773. (Cited on page 37).
- Buchwitz, M., M. Reuter, O. Schneising, W. Hewson, et al. 2017. "Global Satellite Observations of Column-Averaged Carbon Dioxide and Methane: The GHG-CCI XCO₂ and XCH₄ CRDP3 Data Set." *Remote Sensing of Environment*, Earth Observation of Essential Climate Variables, 203 (December 15): 276–295. (Cited on pages 102, 181).
- Buchwitz, M., et al. 2007. "First Direct Observation of the Atmospheric CO₂ Year-to-Year Increase from Space." *Atmos. Chem. Phys.* 7, no. 16 (August 20): 4249-4256. (Cited on page 11).
- Buchwitz, M., et al. 2010. Carbon Monitoring Satellite (CarbonSat) Mission Requirements Document (MRD), Version 1.2. (Cited on page 175).
- Chandrasekhar, S. 1960. *Radiative Transfer*. Dover, Mineola, NY. (Cited on pages 37, 142 sq., 213, 216).
- Chedin, A., et al. 2002. "Annual and Seasonal Variations of Atmospheric CO₂, N₂O and CO Concentrations Retrieved from NOAA/TOVS Satellite Observations." *Geophysical Research Letters* 29, no. 8 (April 30): 110–1. (Cited on page 11).
- Chevallier, F., et al. 2009. "On the Accuracy of the CO_2 Surface Fluxes to Be Estimated from the GOSAT Observations." *Geophysical Research Letters* 36, no. 19 (October 14). (Cited on page 11).

- Chevallier, F., et al. 2016. GHG-CCI: User Requirements Document for the GHG-CCI Project of ESA's Climate Change Initiative version 2.1. October 19. (Cited on pages 11, 149).
- Chevallier, F., et al. 2007. "Contribution of the Orbiting Carbon Observatory to the Estimation of CO2 Sources and Sinks: Theoretical Study in a Variational Data Assimilation Framework." *Journal of Geophysical Research: Atmospheres* 112 (D9): D09307. (Cited on pages 11, 110, 205).
- Ciais, P., et al. 2000. "Regional Biospheric Carbon Fluxes as Inferred from Atmospheric CO₂ Measurements." *Ecological Applications* 10, no. 6 (December 1): 1574–1589. (Cited on page 10).
- Cogan, A. J., et al. 2012. "Atmospheric Carbon Dioxide Retrieved from the Greenhouse Gases Observing SATellite (GOSAT): Comparison with Ground-based TCCON Observations and GEOS-Chem Model Calculations." *Journal of Geophysical Research: Atmospheres* 117 (D21). (Cited on page 173).
- Connor, B., et al. 2016. "Quantification of Uncertainties in OCO-2 Measurements of XCO₂: Simulations and Linear Error Analysis." *Atmos. Meas. Tech.* 9, no. 10 (October 27): 5227–5238. (Cited on page 43).
- Cooper, G. M., et al. 2013. *The Cell: A Molecular Approach*. Sinauer Associates. (Cited on page 61).
- Cox, C., et al. 1954. "Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter." *Josa* 44 (11): 838–850. (Cited on page 159).
- Crisp, D., et al. 2012. "The ACOS CO₂ Retrieval Algorithm Part II: Global XCO₂ Data Characterization." Atmos. Meas. Tech. 5, no. 4 (April 2): 687–707. (Cited on page 77).
- Crisp, D., et al. 2017. "The On-Orbit Performance of the Orbiting Carbon Observatory-2 (OCO-2) Instrument and Its Radiometrically Calibrated Products." *Atmos. Meas. Tech.* 10, no. 1 (January 5): 59–81. (Cited on pages 22, 175, 198).
- Crisp, D., et al. 2010. OCO-2 Level 2 Full Physics Algorithm Theoretical Basis Document. Version 1.0 Rev. 4, November 10. (Cited on page 45).
- Crisp, D., et al. 2004. "The Orbiting Carbon Observatory (OCO) Mission." Advances in Space Research, Trace Constituents in the Troposphere and Lower Stratosphere, 34, no. 4 (January 1): 700–709. (Cited on pages 11, 149).
- Demtröder, W. 2013. *Experimentalphysik 2: Elektrizität Und Optik*. 6th ed. Springer-Lehrbuch. Berlin Heidelberg: Springer-Verlag. (Cited on page 36).

Detmers, R. G., et al. System Quality Assurance Document. (Cited on page 33).

- Devi, V. M., et al. 2016. "Line Parameters Including Temperature Dependences of Self- and Air-Broadened Line Shapes of ¹²C¹⁶O₂: 1.6-Micron Region." Journal of Quantitative Spectroscopy and Radiative Transfer 177 (July 1): 117–144. (Cited on page 17).
- Donlon, C., et al. 2012. "The Global Monitoring for Environment and Security (GMES) Sentinel-3 Mission." *Remote Sensing of Environment*, The Sentinel Missions
 New Opportunities for Science, 120 (Supplement C): 37–57. (Cited on page 60).
- Dorigo, W., et al. 2017. "ESA CCI Soil Moisture for Improved Earth System Understanding: State-of-the Art and Future Directions." *Remote Sensing of Environment*, Earth Observation of Essential Climate Variables, 203 (December 15): 185–215. (Cited on page 127).
- Drouin, B. J., et al. 2017. "Multispectrum Analysis of the Oxygen A-Band." Journal of Quantitative Spectroscopy and Radiative Transfer, Satellite Remote Sensing and Spectroscopy: Joint ACE-Odin Meeting, October 2015, 186 (January 1): 118–138. (Cited on pages 17 sq.).
- Duan, M., et al. 2005. "A Fast Radiative Transfer Model for Simulating High-Resolution Absorption Bands." *Journal of Geophysical Research: Atmospheres* 110 (D15). (Cited on page 153).
- Duan, Z., et al. 2013. "First Results from Version 7 TRMM 3B43 Precipitation Product in Combination with a New Downscaling-calibration Procedure." *Remote Sensing* of Environment 131 (April 15): 1–13. (Cited on page 127).
- Efremenko, D., et al. 2014. "Optical Property Dimensionality Reduction Techniques for Accelerated Radiative Transfer Performance: Application to Remote Sensing Total Ozone Retrievals." *Journal of Quantitative Spectroscopy and Radiative Transfer* 133:128–135. (Cited on page 160).
- Eguchi, N., et al. 2007. "Characteristics of Cirrus Clouds from ICESat/GLAS Observations." *Geophysical Research Letters* 34, no. 9 (May 16): L09810. (Cited on pages 39, 168).
- Eldering, A., et al. 2015. OCO-2 Level 1B Algorithm Theoretical Basis. (Cited on page 175).
- European Space Agency. 2011a. CO_2 Observation Capability of the Sentinel-5 Mission. (Cited on page 175).

——. 2011b. GMES Sentinels 4 and 5 Mission Requirements Document. (Cited on page 175).

- Farr, T. G., et al. 2007. "The Shuttle Radar Topography Mission." *Reviews of Geophysics* 45, no. 2 (May 19). (Cited on page 32).
- Frankenberg, C., A. Butz, et al. 2011. "Disentangling Chlorophyll Fluorescence from Atmospheric Scattering Effects in O₂ A-Band Spectra of Reflected Sun-Light." *Geophysical Research Letters* 38, no. 3 (February 1): L03801. (Cited on pages 70, 102).
- Frankenberg, C., et al. 2012. "Remote Sensing of Near-Infrared Chlorophyll Fluorescence from Space in Scattering Atmospheres: Implications for Its Retrieval and Interferences with Atmospheric CO₂ Retrievals." *Atmos. Meas. Tech.* 5, no. 8 (August 28): 2081–2094. (Cited on pages 60, 102, 104 sq.).
- Frankenberg, C., et al. 2015. "The Orbiting Carbon Observatory (OCO-2): Spectrometer Performance Evaluation Using Pre-Launch Direct Sun Measurements." *Atmos. Meas. Tech.* 8, no. 1 (January 14): 301–313. (Cited on pages 27, 44).
- Frankenberg, C., J. B. Fisher, et al. 2011. "New Global Observations of the Terrestrial Carbon Cycle from GOSAT: Patterns of Plant Fluorescence with Gross Primary Productivity." *Geophysical Research Letters* 38, no. 17 (September 1): L17706. (Cited on pages 1, 59, 70, 76, 84, 91 sq., 94, 114 sqq., 121, 138).
- Geddes, A. G. 2015. "Development of a Fabry-Perot Etalon Spectrometer for High-Resolution Aerosol Observations in the Oxygen A-Band." Thesis, University of Leicester. (Cited on page 16).
- "GES DISC." 2018. Accessed March 29. https://disc.gsfc.nasa.gov/dataset s?project=OCO. (Cited on pages 32, 181).
- "GES DISC ABSCO." 2018. Accessed March 28. https://disc.gsfc.nasa.gov/ information/glossary?title=OCO-2%20ABSCO. (Cited on pages 15, 17).
- "GOSAT Data Archive Service." 2018. Accessed April 4. https://data2.gosat. nies.go.jp/index_en.html. (Cited on page 32).
- Griffiths, P. R., et al. 2007. *Fourier Transform Infrared Spectrometry*. John Wiley & Sons, March 16. (Cited on page 23).
- Guanter, L., et al. 2015. "Potential of the TROPOspheric Monitoring Instrument (TROPOMI) Onboard the Sentinel-5 Precursor for the Monitoring of Terrestrial Chlorophyll Fluorescence." *Atmospheric Measurement Techniques* 8 (3): 1337–1352. (Cited on page 59).
- Guanter, L., et al. 2012. "Retrieval and Global Assessment of Terrestrial Chlorophyll Fluorescence from GOSAT Space Measurements." *Remote Sensing of Environment* 121 (June 1): 236–251. (Cited on pages 94, 114).

- Guanter, L., et al. 2014. "Global and Time-Resolved Monitoring of Crop Photosynthesis with Chlorophyll Fluorescence." *Proceedings of the National Academy of Sciences* 111, no. 14 (August 4): E1327–E1333. pmid: 24706867. (Cited on pages 68, 119 sqq.).
- Haken, H., et al. 2003. Molekülphysik Und Quantenchemie: Einführung in Die Experimentellen Und Theoretischen Grundlagen. 4th ed. Springer-Lehrbuch. Berlin Heidelberg: Springer-Verlag. (Cited on page 13).
- Hasekamp, O., et al. 2008. "Efficient Calculation of Intensity and Polarization Spectra in Vertically Inhomogeneous Scattering and Absorbing Atmospheres." *Journal* of Geophysical Research: Atmospheres 113 (D20). (Cited on pages 156 sq., 177, 183, 193).
- Hasekamp, O., et al. 2016. "Algorithm Theoretical Basis Document for the RemoTeC XCO₂ Full Physics Product v2.3.8." http://www.esa-ghg-cci.org/?q= webfm_send/366. (Cited on pages 179, 183).
- Hoerling, M., et al. 2013. "Causes and Predictability of the 2012 Great Plains Drought." Bulletin of the American Meteorological Society 95, no. 2 (July 2): 269–282. (Cited on page 127).
- Hourdin, F., et al. 2006. "The LMDZ4 General Circulation Model: Climate Performance and Sensitivity to Parametrized Physics with Emphasis on Tropical Convection." *Climate Dynamics* 27 (7-8): 787–813. (Cited on page 32).
- Ingmann, P., et al. 2012. "Requirements for the GMES Atmosphere Service and ESA's Implementation Concept: Sentinels-4/-5 and -5p." *Remote Sensing of Environment*, The Sentinel Missions - New Opportunities for Science, 120 (May 15): 58–69. (Cited on page 175).
- IPCC. 2014. Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, March 24. (Cited on pages 7 sq.).
- Jaacks, R. G., et al. 1989. "Double Pendulum Michelson Interferometer with Extended Spectral Resolution." *Applied Optics* 28, no. 1 (January 1): 29–30. (Cited on page 23).
- Joiner, J., et al. 2011. "First Observations of Global and Seasonal Terrestrial Chlorophyll Fluorescence from Space." *Biogeosciences* 8 (3): 637–651. (Cited on page 59).

- Joiner, J., et al. 2013. "Global Monitoring of Terrestrial Chlorophyll Fluorescence from Moderate Spectral Resolution Near-Infrared Satellite Measurements: Methodology, Simulations, and Application to GOME-2." *Atmospheric Measurement Techniques* 6 (2): 2803–2823. (Cited on page 59).
- Joiner, J., et al. 2014. "The Seasonal Cycle of Satellite Chlorophyll Fluorescence Observations and Its Relationship to Vegetation Phenology and Ecosystem Atmosphere Carbon Exchange." *Remote Sensing of Environment* 152 (September): 375–391. (Cited on page 59).
- Joiner, J., et al. 2012. "Filling-in of near-Infrared Solar Lines by Terrestrial Fluorescence and Other Geophysical Effects: Simulations and Space-Based Observations from SCIAMACHY and GOSAT." *Atmospheric Measurement Techniques* 5 (4): 809. (Cited on page 59).
- Kahn, R. A., et al. 2015. "An Analysis of Global Aerosol Type as Retrieved by MISR: MISR Aerosol Type." *Journal of Geophysical Research: Atmospheres* 120, no. 9 (May 16): 4248-4281. (Cited on page 39).
- Kahn, R., et al. 2001. "Sensitivity of Multiangle Imaging to Natural Mixtures of Aerosols over Ocean." *Journal of Geophysical Research: Atmospheres* 106 (D16): 18219–18238. (Cited on page 189).
- Keeling, C. D., et al. 2001. Exchanges of Atmospheric CO₂ and ¹³CO₂ with the Terrestrial Biosphere and Oceans from 1978 to 2000. Vol. 1-6. Global aspects, SIO Reference Series. San Diego: Scripps Institution of Oceanography. (Cited on page 7).
- Kerr, R. A. 1977. "Carbon Dioxide and Climate: Carbon Budget Still Unbalanced." Science 197, no. 4311 (September 30): 1352–1353. pmid: 17746998. (Cited on page 7).
- Kikuchi, N., et al. 2016. "An Advanced Retrieval Algorithm for Greenhouse Gases Using Polarization Information Measured by GOSAT TANSO-FTS SWIR I: Simulation Study." *Journal of Geophysical Research: Atmospheres* 121, no. 21 (November 16). (Cited on page 145).
- Köhler, P., et al. 2015. "A Linear Method for the Retrieval of Sun-Induced Chlorophyll Fluorescence from GOME-2 and SCIAMACHY Data." *Atmos. Meas. Tech.* 8, no. 6 (June 26): 2589–2608. (Cited on pages 34, 92 sqq., 138).
- Kopparla, P., et al. 2016. "A Fast and Accurate PCA Based Radiative Transfer Model: Extension to the Broadband Shortwave Region." *Journal of Quantitative Spectroscopy* and Radiative Transfer 173 (April 1): 65–71. (Cited on page 159).

- Kopparla, P., et al. 2017. "PCA-Based Radiative Transfer: Improvements to Aerosol Scheme, Vertical Layering and Spectral Binning." Journal of Quantitative Spectroscopy and Radiative Transfer 198 (September 1): 104–111. (Cited on pages 3, 184).
- Krause, G. H., et al. 1991. "Chlorophyll Fluorescence and Photosynthesis: The Basics." *Annual review of plant biology* 42 (1): 313–349. (Cited on page 65).
- Kuze, A., et al. 2009. "Thermal and near Infrared Sensor for Carbon Observation Fourier-Transform Spectrometer on the Greenhouse Gases Observing Satellite for Greenhouse Gases Monitoring." *Applied Optics* 48, no. 35 (December 10): 6716–6733. (Cited on page 22).
- Lamarche, C., et al. 2017. "Compilation and Validation of SAR and Optical Data Products for a Complete and Global Map of Inland/Ocean Water Tailored to the Climate Modeling Community." *Remote Sensing* 9, no. 1 (January 11): 36. (Cited on page 124).
- Lambert, J. H. 1760. *Photometria Sive de Mensura et Gradibus Luminis, Colorum et Umbrae.* Klett. (Cited on page 141).
- Le Quéré, C., et al. 2016. "Global Carbon Budget 2016." *Earth System Science Data* 8, no. 2 (November 14): 605–649. (Cited on page 8).
- Le Quéré, C., et al. 2018. "Global Carbon Budget 2017." *Earth System Science Data* 10, no. 1 (March 12): 405–448. (Cited on pages 8 sq.).
- Lee, J.-E., et al. 2013. "Forest Productivity and Water Stress in Amazonia: Observations from GOSAT Chlorophyll Fluorescence." *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1761 (June 22). pmid: 23760636. (Cited on page 59).
- Lee, P. M. 2012. *Bayesian Statistics: An Introduction*. John Wiley & Sons, June 25. (Cited on page 49).
- Liu, X., et al. 2006. "Principal Component-Based Radiative Transfer Model for Hyperspectral Sensors: Theoretical Concept." *Applied Optics* 45, no. 1 (January 1): 201–209. (Cited on page 157).
- Lommel, E. 1889. "Die Photometrie Der Diffusen Zurückwerfung." Annalen der Physik 272:473–502. (Cited on pages 141, 143).
- Luus, K. A., et al. 2017. "Tundra Photosynthesis Captured by Satellite-Observed Solar-Induced Chlorophyll Fluorescence." *Geophysical Research Letters* 44, no. 3 (February 16): 2016GL070842. (Cited on page 59).

- Magney, T. S., et al. 2017. "Connecting Active to Passive Fluorescence with Photosynthesis: A Method for Evaluating Remote Sensing Measurements of Chl Fluorescence." *New Phytologist.* (Cited on page 67).
- Mallya G. et al. 2013. "2012 Midwest Drought in the United States." Journal of Hydrologic Engineering 18, no. 7 (July 1): 737-745. (Cited on page 127).
- Mishchenko, M. I., et al. 1996. "T-Matrix Computations of Light Scattering by Nonspherical Particles: A Review." Journal of Quantitative Spectroscopy and Radiative Transfer, Light Scattering by Non-Spherical Particles, 55, no. 5 (May 1): 535–575. (Cited on page 38).
- Mishchenko, M. I., et al. 2006. Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering. Cambridge University Press. (Cited on pages 46, 141 sq.).
- Mobley, C. 2018. "Ocean Optics." Accessed January 13. www.oceanopticsbook. info. (Cited on page 146).
- Monteith, J. L. 1972. "Solar Radiation and Productivity in Tropical Ecosystems." Journal of Applied Ecology 9 (3): 747–766. JSTOR: 2401901. (Cited on page 67).
- Moreno, J., et al. 2015. *Report for Mission Selection: FLEX.* ESA SP-1330/2. Noordwijk, The Netherlands: European Space Agency. (Cited on page 60).
- Müller, P., et al. 2001. "Non-Photochemical Quenching. A Response to Excess Light Energy." *Plant Physiology* 125 (4): 1558–1566. JSTOR: 4279788. (Cited on page 65).
- Nassar, R., et al. 2017. "Quantifying CO₂ Emissions From Individual Power Plants From Space." *Geophysical Research Letters* 44 (October 1): 10. (Cited on page 30).
- Natraj, V., et al. 2005. "Application of Principal Component Analysis to High Spectral Resolution Radiative Transfer: A Case Study of the O₂ A Band." Journal of Quantitative Spectroscopy and Radiative Transfer 95 (4): 539–556. (Cited on pages 157, 159, 161).
- Natraj, V., et al. 2007. "A Fast Linearized Pseudo-Spherical Two Orders of Scattering Model to Account for Polarization in Vertically Inhomogeneous Scattering-absorbing Media." *Journal of Quantitative Spectroscopy and Radiative Transfer* 107, no. 2 (September 1): 263–293. (Cited on pages 31, 147, 160).
- Natraj, V., et al. 2008. "Retrieval of XCO₂ from Simulated Orbiting Carbon Observatory Measurements Using the Fast Linearized R-2OS Radiative Transfer Model." *Journal of Geophysical Research: Atmospheres* 113 (D11): D11212. (Cited on page 148).

- Natraj, V., et al. 2010. "On the Use of Principal Component Analysis to Speed up Radiative Transfer Calculations." *Journal of Quantitative Spectroscopy and Radiative Transfer* 111 (5): 810–816. (Cited on page 159).
- NOAA. 2018. "ESRL Global Monitoring Division Global Greenhouse Gas Reference Network." Accessed March 28. https://www.esrl.noaa.gov/gmd/ccgg/ trends/gr.html. (Cited on page 6).
- NOAA Earth System Research Laboratory. 2016. "Cooperative Global Atmospheric Data Integration Project." http://dx.doi.org/10.15138/G3059Z. (Cited on page 10).
- O'Dell, C. W., et al. 2012. "The ACOS CO₂ Retrieval Algorithm Part 1: Description and Validation against Synthetic Observations." *Atmos. Meas. Tech.* 5, no. 1 (January 11): 99–121. (Cited on pages 20, 53, 55, 169).
- O'Dell, C. W. 2010. "Acceleration of Multiple-Scattering, Hyperspectral Radiative Transfer Calculations via Low-Streams Interpolation." *Journal of Geophysical Research: Atmospheres* 115 (D10). (Cited on pages 149, 151, 154 sq., 177 sq., 183, 193).
- Osterman, G. B., et al. 2017. "Warn Level, Bias Correction, and Lite File Product Description." https://disc.gsfc.nasa.gov/information/documents? title=0C0-2%20Documents. (Cited on page 185).
- Pan, Y., et al. 2011. "A Large and Persistent Carbon Sink in the World's Forests." Science (July 14): 1201609. pmid: 21764754. (Cited on page 12).
- Parazoo, N. C., et al. 2014. "Terrestrial Gross Primary Production Inferred from Satellite Fluorescence and Vegetation Models." *Global Change Biology* 20 (10): 3103–3121. (Cited on pages 59, 114 sq., 126).
- Pascal, V., et al. 2017. "An Improved Microcarb Dispersive Instrumental Concept for the Measurement of Greenhouse Gases Concentration in the Atmosphere." In *International Conference on Space Optics — ICSO 2014*, vol. 10563, 105633K. International Society for Optics and Photonics, November 17. (Cited on page 211).
- Plascyk, J. A. 1975. "The MK II Fraunhofer Line Discriminator/FLD-II/ for Airborne and Orbital Remote Sensing of Solar-Stimulated Luminescence." *Optical Engineering* 14 (4): 339–346. (Cited on page 69).
- Plascyk, J. A., et al. 1975. "The Fraunhofer Line Discriminator MKII-an Airborne Instrument for Precise and Standardized Ecological Luminescence Measurement." *IEEE Transactions on Instrumentation and Measurement* 24 (4): 306–313. (Cited on page 69).
- Polonsky, I. N., et al. 2014. "Performance of a Geostationary Mission, geoCARB, to Measure CO₂, CH₄ and CO Column-Averaged Concentrations." *Atmos. Meas. Tech.* 7, no. 4 (April 11): 959–981. (Cited on page 149).
- Porcar-Castell, A., et al. 2014. "Linking Chlorophyll a Fluorescence to Photosynthesis for Remote Sensing Applications: Mechanisms and Challenges." *Journal of Experimental Botany* 65, no. 15 (August 1): 4065–4095. (Cited on page 67).
- Potter, C., et al. 2003. "Continental-Scale Comparisons of Terrestrial Carbon Sinks Estimated from Satellite Data and Ecosystem Modeling 1982–1998." *Global and Planetary Change* 39, no. 3 (November 1): 201–213. (Cited on page 114).
- Poulter, B., et al. 2014. "Contribution of Semi-Arid Ecosystems to Interannual Variability of the Global Carbon Cycle." *Nature* 509, no. 7502 (May): 600–603. (Cited on page 9).
- "Primary Mauna Loa CO₂ Record | Scripps CO₂ Program." 2018. Accessed March 28. http://scrippsco2.ucsd.edu/data/atmospheric_co2/primary_mlo_ co2_record. (Cited on page 7).
- Rayner, P. J., et al. 2001. "The Utility of Remotely Sensed CO₂ Concentration Data in Surface Source Inversions." *Geophysical Research Letters* 28, no. 1 (January 1): 175–178. (Cited on pages 11 sq.).
- Reichler, T., et al. 2003. "Determining the Tropopause Height from Gridded Data." *Geophysical Research Letters* 30, no. 20 (October 1): 2042. (Cited on page 34).
- Reuter, M., et al. 2014. "Satellite-Inferred European Carbon Sink Larger than Expected." Atmos. Chem. Phys. 14, no. 24 (December 22): 13739–13753. (Cited on page 110).
- Rodgers, C. D. 2000. Inverse Methods for Atmospheric Sounding: Theory and Practice.World Scientific. (Cited on pages 47, 49 sq., 168).
- Schwalm, C. R., et al. 2012. "Reduction in Carbon Uptake during Turn of the Century Drought in Western North America." *Nature Geoscience* 5, no. 8 (July 29). (Cited on page 132).
- Somkuti, P., H. Boesch, V. Natraj, et al. 2017a. "Application of a PCA-Based Fast Radiative Transfer Model to XCO₂ Retrievals in the Shortwave Infrared." *Journal* of Geophysical Research: Atmospheres 122, no. 19 (October 12): 10, 477–10, 496. (Cited on pages 2 sq., 177, 207, 209).

- Souza, R. P., et al. 2004. "Photosynthetic Gas Exchange, Chlorophyll Fluorescence and Some Associated Metabolic Changes in Cowpea (Vigna Unguiculata) during Water Stress and Recovery." *Environmental and Experimental Botany* 51 (1): 45–56. (Cited on page 66).
- Spurr, R., et al. 2001. "A Linearized Discrete Ordinate Radiative Transfer Model for Atmospheric Remote-Sensing Retrieval." *Journal of Quantitative Spectroscopy and Radiative Transfer* 68, no. 6 (March 15): 689–735. (Cited on pages 31, 147).
- Spurr, R., et al. 2011. "A Linearized Two-Stream Radiative Transfer Code for Fast Approximation of Multiple-Scatter Fields." *Journal of Quantitative Spectroscopy and Radiative Transfer* 112, no. 16 (November 1): 2630–2637. (Cited on pages 31, 147, 160).
- Spurr, R., et al. 2013. "Linearization of the Principal Component Analysis Method for Radiative Transfer Acceleration: Application to Retrieval Algorithms and Sensitivity Studies." *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (August 1): 1–17. (Cited on pages 159, 167).
- Spurr, R. 2008. "LIDORT and VLIDORT: Linearized Pseudo-Spherical Scalar and Vector Discrete Ordinate Radiative Transfer Models for Use in Remote Sensing Retrieval Problems." In *Light Scattering Reviews 3*, 229–275. Springer Praxis Books. Springer, Berlin, Heidelberg. (Cited on pages 147, 160, 213, 215).
- Spurr, R., et al. 2014. "On the Generation of Atmospheric Property Jacobians from the (V)LIDORT Linearized Radiative Transfer Models." *Journal of Quantitative Spectroscopy and Radiative Transfer* 142 (July 1): 109–115. (Cited on pages 45 sq.).
- Sun Ying et al. 2015. "Drought Onset Mechanisms Revealed by Satellite Solar-induced Chlorophyll Fluorescence: Insights from Two Contrasting Extreme Events." *Journal* of Geophysical Research: Biogeosciences 120, no. 11 (November 2): 2427–2440. (Cited on page 137).
- Sun, Y., et al. 2017. "OCO-2 Advances Photosynthesis Observation from Space via Solar-Induced Chlorophyll Fluorescence." Science 358 (6360). (Cited on page 59).
- Taiz, L., et al. 2010. Plant Physiology: International Edition. 5th Edition. Sinauer Associates, July 9. (Cited on page 61).
- "TANSO-FTS GOSAT JAXA." 2018. Accessed March 29. http://www.eorc. jaxa.jp/GOSAT/instrument_1.html. (Cited on pages 25, 27).
- "TCCON Tccon-Wiki." 2018. Accessed March 29. https://tccon-wiki.caltech. edu/. (Cited on page 189).

- "The Keeling Curve." 2018. Accessed March 28. https://scripps.ucsd.edu/ programs/keelingcurve/. (Cited on page 6).
- "The Transcom Experiment." 2018. Accessed March 30. http://transcom.proje ct.asu.edu/. (Cited on page 76).
- Thomas, G. E., et al. 2002. *Radiative Transfer in the Atmosphere and Ocean*. Cambridge University Press, January 28. (Cited on pages 46, 143, 150, 213, 216).
- Van der Walt, S., et al. 2011. "The NumPy Array: A Structure for Efficient Numerical Computation." *Computing in Science & Engineering* 13 (2): 22–30. (Cited on page 161).
- Wang, P., et al. 2008. "FRESCO+: An Improved O₂ A-Band Cloud Retrieval Algorithm for Tropospheric Trace Gas Retrievals." Atmos. Chem. Phys. 8, no. 21 (November 14): 6565–6576. (Cited on page 95).
- Wark, D. Q., et al. 1965. "Absorption in the Atmosphere by the Oxygen "A" Band." Applied Optics 4, no. 7 (July 1): 839–845. (Cited on page 15).
- Webb, A. J. 2017. "Evaluating CH₄ Concentrations and Emissions in the Amazon Basin Using the Greenhouse Gases Observing Satellite and Dedicated CH₄ Models." Thesis, University of Leicester. (Cited on page 28).
- Wigley, T. M. L., et al. 2005. *The Carbon Cycle*. Cambridge University Press, August 22. (Cited on page 7).
- Willits, D. H., et al. 2001. "Measurement of Chlorophyll Fluorescence as a Heat Stress Indicator in Tomato: Laboratory and Greenhouse Comparisons." *Journal of the American Society for Horticultural Science* 126 (2): 188–194. (Cited on page 66).
- Wolf, S., et al. 2016. "Warm Spring Reduced Carbon Cycle Impact of the 2012 US Summer Drought." *Proceedings of the National Academy of Sciences* 113, no. 21 (May 24): 5880-5885. pmid: 27114518. (Cited on pages 133, 137).
- Wunch, D., P. O. Wennberg, G. C. Toon, et al. 2011. "A Method for Evaluating Bias in Global Measurements of CO₂ Total Columns from Space." *Atmos. Chem. Phys.* 11, no. 23 (December 9): 12317–12337. (Cited on page 185).
- Wunch, D., et al. 2017. "Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO₂ Measurements with TCCON." Atmos. Meas. Tech. 10, no. 6 (June 13): 2209–2238. (Cited on pages 185, 188, 205).

- Wunch, D., G. C. Toon, et al. 2011. "The Total Carbon Column Observing Network." Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 369, no. 1943 (May 28): 2087–2112. pmid: 21502178. (Cited on page 184).
- Yokota, T., et al. 2009. "Global Concentrations of CO2 and CH4 Retrieved from GOSAT: First Preliminary Results." *SOLA* 5:160–163. (Cited on page 22).
- Young, A. T. 1980. "Revised Depolarization Corrections for Atmospheric Extinction." *Applied Optics* 19, no. 20 (October 15): 3427–3428. (Cited on page 37).
- Zarco-Tejada, P. J., et al. 2006. "FluorMODgui V3.0: A Graphic User Interface for the Spectral Simulation of Leaf and Canopy Chlorophyll Fluorescence." *Computers & Geosciences* 32, no. 5 (June 1): 577–591. (Cited on page 66).
- Zarco-Tejada, P. J., et al. 2013. "Spatio-Temporal Patterns of Chlorophyll Fluorescence and Physiological and Structural Indices Acquired from Hyperspectral Imagery as Compared with Carbon Fluxes Measured with Eddy Covariance." *Remote Sensing of Environment* 133 (Supplement C): 102–115. (Cited on page 68).
- Zhang, Y., et al. 2017. "A Global Moderate Resolution Dataset of Gross Primary Production of Vegetation for 2000–2016." *Scientific Data* 4 (October 24). (Cited on page 126).
- Zhao, M., et al. 2005. "Improvements of the MODIS Terrestrial Gross and Net Primary Production Global Data Set." *Remote Sensing of Environment* 95, no. 2 (March 30): 164–176. (Cited on page 68).

Glossary

- **2OS** 2 Orders of Scattering. A radiative transfer model that computes the full Stokes vector for up to two orders of scattering.
- **ABSCO** Tabulated absorption coefficients or cross sections.
- **AOD** Aerosol optical depth, in some literature also AOT (aerosol optical thickness). The (column-integrated) optical depth due aerosol extinction (both scattering and absorption).
- **BOA** Bottom of the atmosphere. Usually refers to the surface layer or level in a model atmosphere.
- **CAMS** Copernicus Atmosphere Monitoring Service. Follow-up to MACC, a service to provide near-real-time analysis and forecast data related to atmospheric composition.
- **CASA** Carnegie-Ames-Stanford-Approach. A process-based global terrestrial biosphere model.
- **CCI** Climate Change Initiative. A program by the European Space Agency to facilitate the generation of high-quality long-term climate datasets.
- **ECMWF** European Centre for Medium-range Weather Forecast.
- **EOF** Empirical Orthogonal Function. The eigenvectors of the covariance matrix that is analysed during PCA.
- **ESA** European Space Agency.
- **EVI** Enhanced Vegetation Index. Similar to NDVI, a reflectance-based vegetation greenness index, which is more robust due to atmospheric corrections.
- **GOME-2** A moderate resolution spectrometer onboard the MetOp-A and MetOp-B satellites, featuring a channel in the O_2 A-band for SIF measurements.

- **GOSAT** Greenhouse gases observing satellite, the first dedicated satellite mission to measure atmospheric CO₂; operated by the Japanese Aerospace Exploration Agency (JAXA).
- **GPP** Gross Primary Production. The total amount of carbon uptake by vegetation, soil or bacteria.
- **ILS** Instrument line shape function, also instrument spectral response function (ISRF).
- **IQR** Interquartile Range. Difference between the 75th and the 25th percentile.
- **LIDORT** Linearized Discrete Ordinates, a general purpose multiple-scattering RT solver, created by R. Spurr.
- **LSI** Low Streams Interpolation. A fast radiative transfer method involving a double interpolation scheme.
- **Multiple Scattering (MS)** Sometimes also diffuse radiation. Contributions to the total radiance due to secondary and higher-order scattering events in the gaseous medium, caused by e.g. aerosols.
- **NDVI** Normalised Differential Vegetation Index. A reflectance-based vegetation greenness index involving radiances in the visual-red and near-infrared ranges.
- **OCO-2** Orbiting Carbon Observatory-2 mission, operated by the National Aeronautics and Space Administration (NASA).
- **PCA** Principal component analysis. The decomposition of a matrix into linearly independent components, ranked by order of explained variance. The principal components are calculated by projecting the original matrix onto the EOFs (eigenvectors of the covariance matrix).
- **PDF** Probability density function.
- **RT** Radiative transfer. An algorithm to describe the propagation of light in an absorbing and/or scattering (atmospheric) medium.
- **RTE** Radiative transfer equation.
- **SIF** Solar-induced Chlorophyll Fluorescence is a byproduct of plant photosynthesis. Radiation in the visible to near-infrared is emitted as a means of reducing surplus excitation energy that would otherwise damage the plant.

- **Single Scattering (SS)** Contributions to the total radiance due to first-order scattering events.
- **SNR** Signal-to-noise ratio.
- **Solar Zenith Angle (SZA)** Polar angle between the surface normal (zenith) and the sun.
- **SWIR** Short-wave infrared region of the electromagnetic spectrum, approximately between $0.9 2.5 \mu m$ (definition varies highly).
- **TANSO-FTS** Thermal and near-infrared sensor for carbon observations. The Fourier-transform spectrometer onboard GOSAT.
- **TCCON** Total column carbon observing network. A global network of ground stations that regularly measure trace gases using Fourier-transform spectrometers.
- **TOA** Top of the atmosphere. Usually refers to the topmost (altitude-wise) layer or level in a model atmosphere.
- **TWOSTR (or 2STREAM)** A fast linearised, two-stream multiple-scattering radiative transfer code, written by R. Spurr and V. Natraj.
- **UoL-FP** University of Leicester Full-Physics retrieval algorithm.
- **Viewing Zenith Angle (VZA)** Polar angle between the surface normal (zenith) and the satellite.
- **XCO₂** Column-averaged dry-air volume mixing ratio of carbon dioxide.