

MONITORING THE CARBON CYCLE: EVALUATION OF TERRESTRIAL
BIOSPHERE MODELS AND ANTHROPOGENIC GREENHOUSE GAS
EMISSIONS WITH ATMOSPHERIC OBSERVATIONS

by

Brendan Byrne

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

© Copyright 2018 by Brendan Byrne

Abstract

Monitoring the carbon cycle: Evaluation of terrestrial biosphere models and anthropogenic greenhouse gas emissions with atmospheric observations

Brendan Byrne

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2018

Reliable projections of climate change will require terrestrial biosphere models (TBMs) that produce robust projections of changes in the exchange of CO₂ between the atmosphere and terrestrial biosphere. In this thesis, atmospheric CO₂ observations are used to evaluate TBMs.

First, the sensitivity of several observing systems to surface fluxes of CO₂ is characterized. This analysis identifies the spatiotemporal scales over which atmospheric CO₂ observations provide significant constraints on net ecosystem exchange (NEE) fluxes.

Second, constraints from atmospheric CO₂ and solar-induced fluorescence (SIF) observations are combined to evaluate the seasonality of NEE, gross primary productivity (GPP) and ecosystem respiration (R_e) fluxes over the northern mid-latitudes for a set of TBMs. It is shown that model-based seasonal cycles of R_e exhibit systematic differences from optimized R_e constrained by atmospheric CO₂ and SIF measurements, with the models overestimating R_e during June-July and underestimating R_e during the fall. Further analysis suggests that the differences could be due to seasonal variations in the carbon use efficiency and to seasonal variations in the leaf litter and fine root carbon pool.

Finally, the ability of TBMs to simulate interannual variability (IAV) in NEE is evaluated. IAV in NEE produced by a set of TBMs and CO₂ flux inversions is compared to proxies of IAV in the carbon cycle, including temperature anomalies, SIF anomalies, and the Palmer drought index. It is shown that CO₂ flux inversions that assimilate

observations from the Greenhouse Gases Observing Satellite (GOSAT) out-perform most TBMs in recovering NEE anomalies driven by climate anomalies, suggesting that GOSAT CO₂ flux inversions can be used to evaluate NEE anomalies produced by TBMs on large scales.

This thesis also describes the installation of an open-path Fourier transform infrared spectroscopy (OP-FTIR) system in downtown Toronto. This system provides continuous observations of CO₂, CO, CH₄ and N₂O, which, in combination with other observing stations, will provide valuable top-down constraints on GHG emissions from Toronto. An initial evaluation of this instrument is performed and comparisons of the observed gases with meteorological observations and CO₂, CO, and CH₄ measurements at a nearby site are presented.

Acknowledgements

Funding for this research was provided by the Canadian Space Agency, the Natural Sciences and Engineering Research Council of Canada, and Environment and Climate Change Canada.

I would like to express my gratitude to Dylan Jones and Kim Strong for supervising my thesis research. I'm grateful that Kim and Dylan afforded me a lot of freedom to develop my own research questions, but also provided thoughtful guidance in how to address these questions effectively. I greatly appreciate the time and effort Kim and Dylan dedicated to mentoring me. I am also indebted to Debra Wunch, who has been very generous with her time in discussing much of the research in this thesis with me. In addition, I have benefited from many thoughtful discussions with Saroja Polavarapu about the research in this thesis. Much of the credit for the installing and maintaining the open path system at U of T is owed to Orfeo Colebatch who has dedicated much of his time to this project. I would also like to thank Jennifer Murphy for the thoughtful feedback she has provided as a member of my committee.

I thank the researchers who co-authored publications and/or presentations with me based on the studies presented in this thesis: Feng Deng, Zhao-Cheng Zeng, Junjie Liu, Ian Baker, Philipp Köhler, Christian Frankenberg, Joanna Joiner, Vivek Arora, Bakr Badawy, Anna Harper, Thorsten Warneke, Christof Petri, Rigel Kivi, Coleen Roehl, Saroja Polavarapu, David Baker, Shamil Maksyutov, Pierre Fogal, Richard Mittermeier, David Griffith, and Doug Worthy. I would particularly like to thank Feng Deng for helping me with the GEOS-Chem adjoint model and Ian Baker who traded many emails with me when I was trying to learn about terrestrial biosphere models. I would also like to thank the many other researchers who provided feedback on the studies presented in this thesis: Felix Vogel, Aldona Wiacek, Li Li, Andy Jacobson, Paul Wennberg, Randy Kawa, and Jim Collatz. And those who provided me with datasets or research tools: Alex Geddes, Cheol-Heon Jeong, and Kenneth Schuldt.

I have enjoyed my time as a member of the Earth, Atmospheric and Planetary Physics group at U of T, and would like to thank all of the members of this group for making my time here enjoyable. I have had many engaging discussions with the graduate students and postdocs at U of T about this research. I would also like to thank Ana Sousa and Krystyna Biel for being very helpful with administrative tasks.

Finally, I thank my family for supporting me in all of my endeavors. Particular thanks is owed to my wife, Adeeba, who was willing to move across the country so that I could pursue this PhD.

Contents

1	Introduction	1
1.1	The carbon cycle	2
1.1.1	The terrestrial biosphere	2
1.1.2	Ocean	7
1.1.3	Atmosphere	7
1.2	Top-down flux estimation	8
1.2.1	History of flux estimation	9
1.2.2	Review of urban GHG studies	18
1.2.3	State of the science	22
1.3	Thesis overview	23
1.3.1	Key results	23
1.3.2	Outline	25
1.3.3	Contributions	26
2	Theory, models and datasets	28
2.1	Chemical transport model	28
2.1.1	GEOS-Chem model	28
2.2	Bayesian data assimilation	30
2.2.1	Remote sounding approach	32
2.2.2	4-D variational data assimilation	33
2.3	CO ₂ datasets	35
2.3.1	Surface measurements	35
2.3.2	Remote sounding measurements	35
3	Sensitivity of CO₂ surface flux constraints to observational coverage	39
3.1	Introduction	39
3.2	Measurements	41
3.2.1	Pseudo-data	42

3.2.2	Discussion of observational coverage	44
3.3	Sensitivity experiments	50
3.3.1	Sensitivity metrics	50
3.3.2	Experiment set-up	51
3.3.3	Results	52
3.4	Flux inversion OSSEs	60
3.4.1	Experiment set-up	60
3.4.2	Inversion OSSE results	61
3.4.3	Inversion OSSE discussion	63
3.5	Conclusions	66
4	Evaluating GPP and respiration estimates over northern mid-latitude ecosystems using solar induced fluorescence and atmospheric CO₂ measurements	68
4.1	Introduction	68
4.2	Data and methods	71
4.2.1	Terrestrial biosphere models	71
4.2.2	Flux inversions	73
4.2.3	GOME-2 SIF	74
4.2.4	Atmospheric X _{CO₂}	76
4.2.5	FLUXCOM	77
4.2.6	Surface air and soil temperature	77
4.3	Results	78
4.3.1	Model fluxes	78
4.3.2	Comparing model GPP and GOME-2 SIF	80
4.3.3	Comparing model NEE and TCCON X _{CO₂}	80
4.3.4	Comparison of GPP and NEE	84
4.3.5	Estimating R_e	87
4.4	Discussion	95
4.4.1	Remaining challenges	97
4.5	Conclusions	98
5	On what scales can GOSAT flux inversions constrain anomalies in terrestrial ecosystems?	101
5.1	Introduction	101
5.2	Data and methods	105
5.2.1	FLUXCOM NEE data	105

5.2.2	Proxies	105
5.2.3	Inversion analyses	106
5.2.4	Terrestrial biosphere models	109
5.2.5	Anomalies and correlations	110
5.3	Results	111
5.3.1	Tropics	111
5.3.2	Northern extratropics	118
5.3.3	Observing system simulation experiments	121
5.4	Discussion	125
5.4.1	Implications of correlations between flux inversions and proxies . .	125
5.4.2	Scales constrained	128
5.4.3	Influence of the inversion configuration	129
5.5	Conclusions	132
6	OP-FTIR	134
6.1	Introduction	134
6.2	Instrumentation	135
6.2.1	Bruker 125M	135
6.2.2	Telescope	137
6.2.3	Retro-reflector	137
6.3	Instrument set-up	138
6.3.1	Location	138
6.4	Data collection	141
6.4.1	Spectra processing	142
6.5	Trace gas retrievals	144
6.5.1	Radiative transfer theory	144
6.5.2	MALT	145
6.6	Meteorological data	146
6.7	Results	150
6.7.1	Sensitivity of GHGs to meteorology	150
6.7.2	Comparison to Hanlan’s Point station	152
6.7.3	Potential for flux estimation	156
6.8	Conclusions	158
7	Conclusions	160
7.1	Identifying deficiencies in TBMs	160
7.2	Monitoring emission targets	162

A	Appendix	165
A.1	GOME-2 SIF errors	165
A.2	Seasonality of SIF and TBMs	165
A.3	Model transport errors	167
A.4	SiB3 and CASA comparison for individual vegetation types	169
A.5	Continental scales	171
B	Initial side-by-side comparisons of the OP-FTIR and LGR	172
C	Processing of OP-FTIR data	175
C.1	Matlab processing	175
C.2	Subtract shutter raised and lowered spectra	176
C.3	Copy files to KS-XENA	176
C.4	Run retrieval	177
	Bibliography	178

List of Tables

1.1	Examples of large city emission monitoring projects.	19
2.1	TCCON sites used in this study.	37
4.1	Terrestrial biosphere models used in this study.	72
5.1	Set-up of GEOS-Chem flux inversions. Differences are in model transport resolution, prior fluxes, and prior errors.	107
5.2	Slope and coefficient of determination (R^2) for linear regressions for anomalies across the entire tropics.	115
5.3	Slope and coefficient of determination (R^2) for linear regressions for regional anomalies during JJA in the northern extratropics.	120
5.4	Slope and coefficient of determination (R^2) for linear regressions for OSSE experiments.	124

List of Figures

1.1	Simplified schematic of the global carbon cycle. Numbers represent reservoir mass (in PgC) and annual carbon exchange fluxes (in PgC yr ⁻¹). Black numbers and arrows indicate reservoir mass and exchange fluxes estimated for the time prior to the Industrial Era, about 1750. Red arrows and numbers indicate annual ‘anthropogenic’ fluxes averaged over the 2000–2009 time period. These fluxes are a perturbation of the carbon cycle during Industrial Era post 1750. Reprinted from Figure 6.1 of Chapter 6 (Ciais et al., 2014) of the Working Group I Contribution to the Fifth Assessment Report (AR5) of the International Panel on Climate Change (IPCC).	6
1.2	Atmospheric CO ₂ concentration at Mauna Loa Observatory as a function of time (Scripps Institute of Oceanography, https://scripps.ucsd.edu). . .	8
3.1	Number of measurements per day for each observing system. Each subplot shows the number density of measurements for a given observing system and season. Each row, from top to bottom, is for a different observing system: surface, TCCON, GOSAT, and OCO-2 measurements. Each column, from left to right, is for a different meteorological season: SON, DJF, MAM, JJA.	43
3.2	GOSAT and OCO-2 measurements for each viewing mode. The time series for each viewing mode are shown for (a) GOSAT and (c) OCO-2. The spatial distribution of measurement number density for each viewing mode and season are also shown for (b) GOSAT and (d) OCO-2.	46
3.3	Number density of OCO-2 measurements for various WL ranges over four seasons. WLs decrease from the top to bottom.	48
3.4	Number of GOSAT (red) and OCO-2 (blue) measurements per year after aggregation as a function of model grid size (degrees).	49

3.5	Spatial distribution of the sensitivity metric (β) for observing system (rows) and each season (columns).	53
3.6	Spatial distribution of (top) seasonal surface fluxes and (bottom) the contribution metric (ψ) for each observing system (rows) and each season (columns).	54
3.7	Sensitivity metric for OCO-2 and GOSAT viewing modes. Sensitivity of (a) all GOSAT viewing modes, and of (b) M-Gain nadir, H-Gain nadir, and glint measurements separately. Sensitivity of (c) all OCO-2 measurements, and of (d) land nadir, land glint, and ocean glint separately.	56
3.8	Spatial distribution of the relative contribution to the sensitivity metric from OCO-2 observations in various WL ranges for each season (columns).	57
3.9	Seasonal differences (relative to SON) in sensitivity due to observational coverage and transport. For the surface data (top row), the observational coverage for SON is fixed over the seasonal cycle, so the seasonal differences are due only to transport. For TCCON, GOSAT, and OCO-2 data, the top panels show differences due to the changing observational coverage and to transport, whereas the bottom panels show differences due to transport (with fixed SON observational coverage).	59
3.10	Results of idealized OSSEs. (a) Spatial distribution of the true surface fluxes (in $\text{kg CO}_2 \text{ m}^{-2} \text{ s}^{-1}$) for each season. (b) The ratio of a posteriori to true fluxes for each observing system and season. (c) Difference between a posteriori and a priori surface fluxes for each observing system (in $\text{kg CO}_2 \text{ m}^{-2} \text{ s}^{-1}$).	62
3.11	(left section of panels) Differences (in Pg/yr) between the true and a priori fluxes (black), true and surface a posteriori fluxes (gray), true and TCCON a posteriori fluxes (green), true and GOSAT a posteriori fluxes (red), and true and OCO-2 a posteriori fluxes (blue) for each season, and (right section of panels) the annual mean of the absolute seasonal differences shown in the left section for (a) all land, (b) Europe, (c) Australia, (d) tropical Asia, (e) temperate east Asia, (f) boreal east Asia, (g) north Africa, (h) south Africa, (i) temperate South America, (j) tropical South America, (k) temperate North America, and (l) boreal North America. Shown in (m) is the estimated annual net flux for each TransCom land region.	64

4.1	MODIS IGBP vegetation classification (at a horizontal resolution of $2^\circ \times 2.5^\circ$). Coloring indicates that the given vegetation type makes up more than 50% of the vegetation type in the gridcell. The vegetation regions are: ENF, DNF, southern mixed forests, northern mixed forests, grasslands and croplands. Red circles show the locations of the four TCCON sites examined in this study: Park Falls (45.9°N , 90.3°W), Orléans (48.0°N , 2.1°E), Białystok (53.2°N , 23.0°E), and Sodankylä (67.4°N , 26.6°E). . . .	75
4.2	Mean seasonal cycles of (a) GPP, (b) R_e , and (c) NEE simulated by SiB3 (green), CASA (blue), CTEM-CRU (dashed salmon), CTEM-GEM (dashed orange), JULES (dashed purple), ANN (dotted gray), MARS (dotted purple-gray), and RF (dotted cyan) between 39°N and 65°N . CT2016 (solid red) and GOSAT-Inv NEE (dash-dotted red) are also plotted.	78
4.3	Normalized seasonal cycles of GOME-2 SIF (NASA and GFZ), and model GPP (SiB3, CASA, CTEM-CRU, CTEM-GEM, JULES and FLUXCOM ANN, FLUXCOM MARS, and FLUXCOM RF) for six vegetation regions. For each panel, the upper plot shows the seasonal cycle of GPP and SIF scaled so that the integral over the season equals one. The lower plot shows the difference between scaled TBM or FLUXCOM GPP and scaled NASA GOME-2 SIF. Grey shaded regions show the uncertainty estimate of the SIF seasonal cycle.	81
4.4	Five-year mean X_{CO_2} seasonal cycle at Sodankylä (top row), Białystok (second row), Orléans (third row) and Park Falls (bottom row). TCCON X_{CO_2} is shown in black. The columns show, from left to right: GOSAT-Inv, CT2016, CASA, SiB3, CTEM-CRU, CTEM-GEM, and JULES. Shaded regions indicate the uncertainty in the functional fit.	83
4.5	The RMS difference between normalized GPP and NASA GOME-2 SIF (averaged across all vegetation regions) versus the RMS difference between simulated X_{CO_2} and TCCON (averaged across the four TCCON sites). .	84
4.6	GPP $\cdot(-1)$, R_e , and NEE fluxes over 39° – 65° N. (a) CASA GPP $\cdot(-1)$, R_e and scaled $f(T)$ (Eq. 4.3). (b) SiB3 GPP $\cdot(-1)$, R_e , and scaled $f(T)$ (Eq. 4.3). (c) SiB3 and CASA NEE.	86

4.7	Left column shows (a) model $GPP \cdot (-1)$, (c) model R_e , (e) $optR_{inv-mod}$, and (g) the difference between optimized and model R_e for SiB3, CASA, and FLUXCOM (ANN, MARS, and RF). $optR_{CT2016-mod}$ (based on CT2016 NEE) is represented by solid lines, whereas $optR_{GOSATinv-mod}$ (based on GOSAT-Inv) is indicated by dashed lines. Right column shows the normalized seasonal cycles of (b) model GPP, (d) model R_e , (f) $optR_{inv-mod}$, and (h) the difference between optimized and model R_e . In all panels the solid, thick black line represent the mean of all the curves shown.	89
4.8	$optR_{inv-mod}$ (solid lines), after scaling GPP over a range of values, for SiB3, CASA and FLUXCOM. $optR_{inv-mod}$ are calculated after first scaling GPP to (a) 23 $Pg\ yr^{-1}$, (b) 26 $Pg\ yr^{-1}$, (c) 29 $Pg\ yr^{-1}$, and (d) 32 $Pg\ yr^{-1}$. R_e produced by SiB3, CASA, and FLUXCOM without scaling is indicated by the dotted lines. Colors are as in Figure 4.7.	90
4.9	(a, b) $optR_{inv-mod}$, (c, d) optimized R_H , and (e, f) optimized $R_H/f(T)$ using parameterizations from Randerson et al. (1996) (left column) and from (Denning et al., 1996) (right column). The solid black line shows the mean seasonal cycle and shaded gray region shows the range of optimized R_e for all NEE and GPP. The dotted black line shows $f(T)$ for CASA (a) and for SiB3 (b) scaled to fit plot area.	92
4.10	Normalized seasonal cycles of (a–c) model GPP, (d–f) model R_e , (g–i) $optR_{inv-mod}$, and (j–l) the difference between $optR_{inv-mod}$ and model R_e for (left column) North America, (middle column) Europe, and (right column) Asia. For subplots g–l, $optR_{CT2016-mod}$ is represented by solid lines, whereas $optR_{GOSATinv-mod}$ is indicated by dashed lines. In all panels the solid, heavy black line represent the mean of all the curves shown.	94
4.11	Maximum rate of drawdown ($gC\ m^{-2}\ day^{-1}$) during the growing season for (a) CT2016, (b) GOSAT-Inv, and (c) GOSAT-Inv minus CT2016.	98
5.1	Sub-continental regions in (a) the extratropics and (b) the tropics. In the tropics, we generate three continents by combining the regions in the Americas, Africa and the Middle East, and the Asia-Pacific and Indian sub-continent.	111

- 5.2 Correlation over a range of scales for models and inversions with (top) NINO 3.4 index, (second) $(-1)\times\text{SIF}$, (third) scPDSI, (fourth) T_{soil} , and (bottom) FLUXCOM NEE in the tropics. Squares represent terrestrial ecosystem models: VISIT (cyan), JULES (blue), CASA GFED CMS (green), CASA GFED 4.1 (magenta) and the black circle shows the mean correlation of the models. GOSAT flux inversions are represented by: GOSAT L4 (cyan up-triangle), $\text{GC}_{4\times 5-44\%-\text{IAV}}$ (green up-triangle), $\text{GC}_{4\times 5-100\%-\text{IAV}}$ (green down-triangle), $\text{GC}_{4\times 5-44\%}$ (red up-triangle), $\text{GC}_{4\times 5-100\%}$ (red down-triangle), $\text{GC}_{2\times 2.5-66\%}$ (orange up-triangle), and $\text{GC}_{2\times 2.5-200\%}$ (orange down-triangle). The green star shows CT2016. The gray circle and line show the correlation with the NINO 3.4 index. Dashed black lines indicate the correlation required for an α of 0.05, therefore, all correlations greater than the dashed black line indicate $P < 0.05$ 112
- 5.3 Correlations of monthly anomalies over tropical land at $4^\circ \times 5^\circ$ spatial resolution. Columns show coefficient of correlation (R) of (left) $\text{GC}_{2\times 2.5-200\%}$, (center) NINO 3.4 index, and (right) the difference between the two with (top row) the NINO 3.4 index, (second row) $(-1)\times\text{SIF}$, (third row) scPDSI, (fourth row) T_{soil} , and (bottom row) FLUXCOM NEE. 116
- 5.4 NEE anomalies ($\text{gC m}^{-2} \text{ day}^{-1}$) for FLUXCOM and $\text{GC}_{2\times 2.5-200\%}$ in the tropics. (left column) Monthly anomalies, (center column) smoothed (3-month running mean) monthly anomalies, and (right column) continental anomalies minus the scaled mean tropical anomalies for (a–b) the entire tropics, (c–e) the Americas, (f–h) Africa and the Middle East, and (i–k) the Asia Pacific and Indian sub-continent. For each panel, R^2 shows the coefficient of determination between $\text{GC}_{2\times 2.5-200\%}$ and FLUXCOM NEE anomalies within the panel. 117
- 5.5 Northern extratropical anomalies during JJA. Anomalies for (top row) $(-1)\times\text{SIF}$, (second) scPDSI, (third) T_{soil} , (fourth) FLUXCOM NEE, and (bottom) $\text{GC}_{2\times 2.5}$ over JJA for (left to right) 2010–2013. Black boxes highlight major climate anomalies: the 2010 Russian heat wave, 2011 drought in Mexico and southern USA, the 2012 North American drought, and the 2013 California drought. 119

5.6	Mean correlation coefficient (R) with the truth over a range of spatial scales for CT2016 (white star), OSSE _{4x5-100%-IAV} (white down-triangle), OSSE _{CT2016-44%} (grey up-triangle), OSSE _{CT2016-100%} (grey down-triangle), OSSE _{JULES-44%} (black up-triangle), and OSSE _{JULES-100%} (black down-triangle).	122
5.7	Monthly NEE anomalies ($\text{gC m}^{-2} \text{ day}^{-1}$) for OSSE _{JULES-100%} (red), OSSE _{CT2016-100%} (green), OSSE _{4x5-100%-IAV} (blue) and truth (black) in the tropics. (left column) Monthly anomalies, (center column) smoothed (3-month running mean) monthly anomalies, and (right column) continental anomalies minus the scaled mean tropical anomalies for (a–b) the entire tropics, (c–e) the Americas, (f–h) Africa and the Middle East, and (i–k) the Asia Pacific and Indian sub-continent.	123
5.8	Correlation between FLUXCOM MARS GPP anomalies and SIF anomalies at $2^\circ \times 2.5^\circ$ spatial resolution.	127
5.9	Comparison of GC _{IAV} posterior and prior IAV. (a) Correlation coefficient (R) between the posterior and prior IAV in the tropics at the spatial scale of $4^\circ \times 5^\circ$. (b) Mean correlation coefficient (R) between posterior and prior IAV in the tropics for different degrees of spatial aggregation. (c) Northern extratropical anomalies during JJA for (top) prior and (bottom) posterior NEE for (left–right columns) 2010–2013.	131
6.1	Optical path of the Bruker IFS 125M spectrometer. Reprinted from the IFS 125M manual.	136
6.2	The OP-FTIR telescope and movable mirror mounted within a custom-built frame.	137
6.3	View of the top of the telescope. Ray tracing (a) from the globar to the retro-reflector and (b) from the retro-reflector into the telescope are shown. The shading indicates the path of the IR beam.	139
6.4	Light path for the entire optical system. The light path is shown as a collimated beam emerging from the OAP at the top of the telescope, traveling to and returning from the retro-reflector, being focused by the telescope and collimated again before being directed into the Bruker 125M.	140
6.5	Path between the instrument and retro-reflector. (Google Earth V 7.3.2.5483. (June 22, 2018). Toronto, Canada, $43^\circ 39' 37.13''$ N $79^\circ 23' 49.77''$ W, eye alt 246 m.)	140

6.6	Location of the OP-FTIR system in downtown Toronto, with buildings which are 100 m or greater in height indicated by markers (http://skyscraperpage.com/cities/m)	
6.7	a) OP-FTIR absorption spectra recorded with the global shutter raised (blue) and lowered (red). b) The difference between spectra recorded with the shutter raised and lowered.	143
6.8	(a) Measured, (b) fitted, and (c) residual spectra in the wavenumber range 2132.52–2234.97 cm^{-1} for 1 April 2018 at 22:25 local time. The fitted spectrum is obtained by optimizing CO_2 , CO , N_2O , and H_2O dry-air mole fractions.	147
6.9	Scatter plots of ECCC vs. McLennan met-station data. The panels show (a) temperature, (b) pressure, (c) wind direction, and (d) wind speed after averaging to hourly temporal resolution. The black line indicates the 1:1 line and the red line shows the linear regression.	148
6.10	Scatter plots of Wallberg vs. McLennan met-station data. The panels show (a) temperature, (b) pressure, (c) wind direction, and (d) wind speed after averaging to hourly temporal resolution. The black line indicates the 1:1 line and the red line shows the linear regression.	149
6.11	Time series of (a) H_2O (b) CO , (c) CO_2 , (d) N_2O , and (e) CH_4 dry-air mole fractions retrieved from OP-FTIR measurements. Gases are shown in blue for retrievals using 2132.52–2234.97 cm^{-1} and red for 2900–3027 cm^{-1} . Pastel colors indicate hourly means and darker colors indicate daily means.	151
6.12	Number density for OP-FTIR measurements of (a) CO , (b), CO_2 , (c) N_2O , and (d) CH_4 dry-air mole fractions as a function of wind speed for November 2017 through June 2018. Blue dots show the mean dry-air mole fractions for each wind speed. Meteorological data is taken from the McLennan met-station.	152
6.13	Pollution roses for OP-FTIR measurements of (a) CO , (b), CO_2 , (c) N_2O , and (d) CH_4 spanning November 2017 through June 2018. Meteorological data is taken from the McLennan met-station.	153
6.14	Locations of the OP-FTIR and Hanlan’s Point station within Toronto, Ontario. Google Maps (2018).	154
6.15	Pollution roses of the gradient between the OP-FTIR and Hanlan’s Point for (a) CO , (b), CO_2 , and (c) CH_4 spanning November 2017 through June 2018. Meteorological data is taken from the McLennan met-station, but the wind direction must also be in the same within the same 45° wind direction bin for the Toronto Island ECCC met-station.	155

6.16	Number of occurrences of hourly gradients in (a) CO ₂ , (b) CO, and (c) CH ₄ for wind from the north (blue) and from the south (green) and for wind speeds between 1.5 m s ⁻¹ and 4.5 m s ⁻¹	157
A.1	Seasonal cycle of NASA GOME-2 SIF over the years 2008-2013 for the ENF vegetation region. Also plotted is the mean seasonal cycle averaged over this time period.	166
A.2	(a) Start of season (day when SIF (GPP) reaches 25% of the maximum value), (b) end of season (day when SIF (GPP) reaches below 25% of the maximum value), and (c) length of the season (in days).	166
A.3	Mean X _{CO₂} seasonal cycle at a) Sodankylä, b) Białystok, c) Orléans and d) Park Falls. Solid lines show the mean X _{CO₂} seasonal cycle from TCCON observations (black) and simulated by GEOS-Chem (red) over the period 2008-2012 with minimal data screening. Dashed lines show the mean X _{CO₂} seasonal cycle from TCCON observations after additional data screening was performed (black) and simulated by TM5 (red) over the period 2008-2014 (also with additional data screening).	168
A.4	Day of year at which the peak GPP (black), peak <i>R_e</i> (light gray) and peak NEE (dark gray) occur for (a) SiB3 and (b) CASA, and (c) the difference in timing between SiB3 and CASA.	170
A.5	Seasonal cycles of (a–c) inversion NEE (d–f) model GPP, (g–i) model <i>R_e</i> , (j–l) <i>optR_{inv-mod}</i> , and (m–o) the difference between <i>optR_{inv-mod}</i> and model <i>R_e</i> for (left column) North America, (middle column) Europe, and (right column) Asia. For subplots a–c, CT2016 NEE is the solid line and GOSAT-Inv is the dash-dot line. For subplots m–n, <i>optR_{CT2016-mod}</i> is represented by solid lines, whereas <i>optR_{GOSATinv-mod}</i> is indicated by dashed lines. In all panels the solid, heavy black line represent the mean of all the curves shown.	171
B.1	Measurements of (a) CO ₂ and (c) CH ₄ from the (red) OP-FTIR and (black) 3 second averages and (green) 6 minute medians with the LGR positioned on the Burton Tower balcony on September 4th, 2018. Scatter plot and linear regression of OP-FTIR measurements against LGR measurements for (b) CO ₂ and (d) CH ₄	173

B.2 Measurements of (a) CO₂ and (c) CH₄ from the (red) OP-FTIR and (black) 3 second averages and (green) 6 minute medians with the LGR positioned on the Galbraith roof on September 7th, 2018. Scatter plot and linear regression of OP-FTIR measurements against LGR measurements for (b) CO₂ and (d) CH₄. 174

Acronyms

4D-Var Four Dimensional Variational Assimilation

ACOS Atmospheric CO₂ Observations from Space

AGL Above Ground Level

ANN Artificial Neural Network

AR5 Fifth Assessment Report

ATP Adenosine Triphosphate

BEACO₂N Berkeley Atmospheric CO₂ Observation Network

BETHY Biosphere Energy Transfer Hydrology

CASA Carnegie Ames Stanford Approach

CCCma Canadian Centre for Climate Modeling and Analysis

CDIAC Carbon Dioxide Information Analysis Center

CMS Carbon Monitoring System

CRU Climate Research Unit

CT2016 CarbonTracker 2016

CTEM Canadian Terrestrial Ecosystem Model

CTM Chemical Transport Model

CUE Carbon Use Efficiency

DNF Deciduous Needle-leaf Forests

DJF December-January-February

ECCC Environment and Climate Change Canada

ECMWF European Centre for Medium-Range Weather Forecasts

ENSO El Niño–Southern Oscillation

ENF Evergreen Needle-leaf Forests

ESRL Earth System Research Laboratory

fPAR fraction of Photosynthetically Active Radiation

FTIR Fourier Transform Infrared

FTS Fourier Transform Spectroscopy

GCM General Circulation Model

GEM-MACH-GHG Global Environmental Multi-scale - Modeling Air Quality and
Chemistry - Greenhouse Gas

GEOS Goddard Earth Observing System

GFED Global Fire Emissions Database

GHG Greenhouse Gas

GISS Goddard Institute for Space Studies

GMAO Global Modeling and Assimilation Office

GOSAT Greenhouse Gases Observing Satellite

GOME-2 Global Ozone Monitoring Experiment-2

GPCP Global Precipitation Climatology Project

GPP Gross Primary Productivity

H-Gain High Gain

HadISST Hadley Centre Sea Ice and Sea Surface Temperature

IAV Interannual Variability

IR Infrared

IFS Infrared Fourier Spectrometer

INFLUX Indianapolis Flux Experiment

ITCZ Intertropical Convergence Zone

IPCC International Panel on Climate Change

IGBP International Geosphere-Biosphere Program

JCDAS JMA Climate Data Assimilation System

JJA June-July-August

JMA Japan Meteorological Agency

JRA-25 Japan 25-year reanalysis

JULES Joint UK Land Environment Simulator

L-BFGS-B Limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm for Bound-constrained optimization

LAI Leaf Area Index

LUE Light Use Efficiency

LGR Los Gatos Research

M-Gain Medium Gain

MALT Multiple Atmospheric Layer Transmission

MAM March-April-May

MAP Maximum A posteriori Probability

MERRA Modern Era Retrospective-analysis for Research and Applications

MetOp-A Meteorological Operational Satellite-A

MODIS Moderate Resolution Imaging Spectroradiometer

MARS multivariate regression spline

NASA National Aeronautics and Space Administration
NCEP National Centers for Environmental Prediction
NEE Net Ecosystem Exchange
NEMA National Electrical Manufacturers Association
NDVI Normalized Differential Vegetation Index
NIES National Institute for Environmental Studies
NIES-TM NIES global atmospheric tracer transport model
NOAA National Oceanic and Atmospheric Administration
NPP Net Primary Productivity
OAP Off-Axis-Paraboloid
OCO-2 Orbiting Carbon Observatory
OP-FTIR Open Path Fourier Transform Infrared
OPD Optical Path Difference
OSSE Observing System Simulation Experiment
PAR Photosynthetically Active Radiation
PBL Planetary Boundary Layer
RF Random Forest
RMS Root Mean Square
scPDSI Monthly Self-calibrated Palmer Drought Severity Index
SDBM Simple Diagnostic Biospheric Model
SiB Simple Biosphere model
SIF Solar Induced Fluorescence
SOCAAR Southern Ontario Centre for Atmospheric Aerosol Research
SON September-October-November

SST Sea Surface Temperature

SWIR Shortwave Infrared

TANSO-FTS Thermal and Near-infrared Sensor for Carbon Observation Fourier Transform Spectrometer

TBM Terrestrial Biosphere Model

TCCON Total Carbon Column Observing Network

TransCom Transport Comparison Project

VISIT Vegetation Integrative Simulator for Trace Gases

VOC Volatile Organic Compound

WL Warn Level

WMO World Meteorological Organization

Chapter 1

Introduction

Anthropogenic climate change is primarily driven by the emission of greenhouse gases (GHGs) to the atmosphere through anthropogenic activities (IPCC, 2013). Climate change is impacting the Earth system in a multitude of ways with significant negative impacts for human societies (IPCC, 2014). This has motivated recent international agreements to try to limit the growth of atmospheric GHGs by adopting GHG emission reduction targets. However, the rate of emission reductions required to limit future warming to some given threshold are not well quantified. This is partially because of poorly understood feedbacks within the Earth system that could impact the atmospheric concentrations of GHGs. These feedbacks are of particular concern for CO₂. Over the industrial era, uptake of anthropogenic CO₂ by terrestrial ecosystems and the oceans has reduced the atmospheric growth rate such that only ~44% of anthropogenic CO₂ remains in the atmosphere. Thus, carbon cycle feedbacks that impact the airborne fraction of anthropogenic CO₂ could have significant implications for the anthropogenic emission pathways needed to limit warming of the climate system. Consequently, improving our ability to predict the atmospheric CO₂ growth for different emission scenarios is of high importance. The focus of this thesis is to evaluate terrestrial biosphere models (TBMs) that will be critical for predicting the airborne fraction of CO₂ into the future, and to develop GHG emission monitoring techniques for urban environments to monitor whether emission reduction targets are being met.

This chapter reviews the current understanding of Earth's carbon cycle (Sec. 1.1) and how atmospheric CO₂ observations have informed our current understanding of the carbon cycle (Sec. 1.2). This is followed by an overview of the research presented in this thesis (Sec. 1.3).

1.1 The carbon cycle

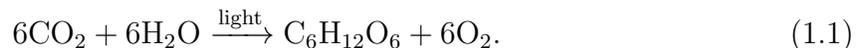
The carbon cycle describes the movement of carbon between various reservoirs in the Earth system. On the timescales of interest (< 1000 years), the carbon cycle can be approximated as a closed system with atmospheric, biospheric, and oceanic carbon reservoirs. However, since the Industrial Revolution, anthropogenic activities have perturbed this system with the transfer of carbon from the geologic to atmospheric reservoir through the burning of fossil fuels and cement production. In this section, the carbon cycle of Earth is described, with emphasis on recent perturbations to the carbon cycle as a result of anthropogenic activities. Throughout this thesis, a co-ordinate system is used where fluxes to the atmosphere are positive and fluxes out of the atmosphere are negative.

1.1.1 The terrestrial biosphere

The terrestrial biosphere refers to all terrestrial living and dead organic matter and contains 450–650 PgC in vegetation, 1500–2400 PgC in soils, and ~ 1700 PgC in permafrost (Ciais et al., 2013). The terrestrial biosphere exchanges carbon with both the atmospheric and oceanic reservoirs. Exchange between the terrestrial biosphere and atmosphere is primarily due to gross primary productivity (GPP) and ecosystem respiration (R_e). Biomass burning and carbon emissions to the atmosphere in the form of CO, CH₄, and volatile organic compounds (VOCs) play a smaller role. Carbon can move from the terrestrial biosphere to the oceanic reservoir through the leaching of carbon from soils into rivers that flow into the ocean.

Gross primary productivity

GPP represents the gross uptake of CO₂ through photosynthesis, wherein CO₂ is reduced to make organic matter using solar energy:



The sequestered carbon (given as glucose in Equation 1.1) can then be used to synthesize more complicated organic molecules or be respired. On an annual basis, GPP draws more than 100 PgC from the atmosphere (Ciais et al., 2013).

The spatiotemporal distribution of GPP is sensitive to the availability of H₂O and light to drive photosynthesis (Equation 1.1). Annual total GPP is highest in the tropics, where sunlight and moisture are available year round (Anav et al., 2015). In the subtropics, monsoonal regions have high productivity, but many other regions have low GPP

due to water unavailability (e.g., the Sahara, Australia). Temperate regions have very strong seasonality in GPP, with large fluxes from May through September and small fluxes in the winter. This seasonality is partially due to variations in light availability but is also due to temperature variations (Berry and Bjorkman, 1980). GPP is limited by temperature at low and high extremes. At low temperatures, photosynthesis is limited due to the slower rate of chemical reactions (Chapin et al., 2002). At high temperatures, photosynthesis declines due to increased photorespiration and, under extreme conditions, enzyme inactivation and destruction of photosynthetic pigments (Chapin et al., 2002). The optimum temperature for GPP is quite variable between plant species, as they have evolved different temperature sensitivities to suit their habitat. GPP can also be impacted by nutrient availability. In particular, nitrogen and phosphorus availability are important variables. Nitrogen is a component of RuBisCo, an enzyme involved in carbon fixation, and phosphorus is a component of adenosine triphosphate (ATP) used in intracellular energy transfer.

Ecosystem respiration

R_e represents the gross release of CO_2 by an ecosystem. R_e occurs using carbon that was sequestered in GPP, and thus the spatial distribution of R_e fluxes closely mirrors that of GPP. R_e can be subdivided into autotrophic respiration (R_a) and heterotrophic respiration (R_H). R_a is respiration by autotrophs in the maintenance and synthesis of living tissue in the leaves, stem and roots (Waring and Running, 2010). Maintenance respiration has an approximately exponential dependence on temperature, whereas synthesis respiration is proportional to the carbon incorporated in new tissue. The fraction of GPP that is not used in R_a can be characterized by the carbon use efficiency (CUE), defined as

$$\text{CUE} = \frac{\text{GPP} + R_a}{\text{GPP}}. \quad (1.2)$$

Note that GPP values are negative and R_a values are positive. In many TBMs, the CUE is assumed to be a constant, typically 0.5 or 0.6. However, observational studies find a wider range of values for the CUE. In a synthesis of 26 studies, DeLucia et al. (2007) showed that CUE is variable between different forest types, ranging over 0.23–0.83. They showed that CUE decreased with increasing forest age, increased with the ratio of leaf mass-to-total mass, and was sensitive to the forest type. Furthermore, many studies have shown that CUE varies throughout the year. In particular, the “Kok effect”, wherein light inhibits mitochondrial respiration in leaves, can result in seasonal variation in CUE (Heskel et al., 2013; Wehr et al., 2016).

R_H is respiration by heterotrophic organisms during the decomposition of nonliving organic matter. The majority of R_H occurs in soils and is primarily performed by soil animals and microbes. The rate of R_H is dependent on the amount and quality of substrate available. The labile (leaf litter and fine root) carbon pool is the most easily available carbon pool and provides the substrate for the majority of R_H . Only a small fraction of R_H is derived from decomposition of older, more recalcitrant carbon compounds, which generally have long turnover times of more than 100 years (Trumbore, 2000; Ryan and Law, 2005). Sources of leaf litter and fine root carbon to the labile pool can have significant seasonal variability. Randerson et al. (1996) showed that seasonal variations in this carbon pool can have a large impact on estimates of model R_H . In addition to substrate dependence, R_H is dependent on soil temperature and moisture.

Net ecosystem exchange

NEE is the residual between GPP and R_e :

$$\text{NEE} = \text{GPP} + R_e, \quad (1.3)$$

where fluxes to the atmosphere are positive, such that $\text{GPP} \leq 0$ and $R_e \geq 0$. NEE is a useful concept because it describes the net flux of CO_2 between the biosphere and atmosphere due to living organisms. Furthermore, NEE dominates the exchange of CO_2 between the terrestrial biosphere and atmosphere on seasonal timescales.

Other fluxes

Although most of the carbon (more than 100 PgC yr^{-1}) sequestered through GPP is released back to the atmosphere through R_e , there are other mechanisms by which carbon can leave an ecosystem. Biomass burning releases $\sim 2.16 \text{ PgC yr}^{-1}$ to the atmosphere every year (van der Werf et al., 2017). Although the annual net biomass burning flux is much smaller than GPP and R_e , the interannual variability (IAV) in biomass burning is substantial. van der Werf et al. (2017) estimate a standard deviation in IAV of biomass burning of 0.32 PgC yr^{-1} . For comparison, Peylin et al. (2013) report a standard deviation in IAV of the total land flux (excluding fossil fuels) of 1.06 PgC yr^{-1} .

In addition to biomass burning, carbon can be emitted from the terrestrial biosphere to the atmosphere in the form of CH_4 , CO , and VOCs. These compounds are then oxidized to CO_2 in the atmosphere. Randerson et al. (2002) estimate that this amounts to $\sim 1.5 \text{ PgC yr}^{-1}$.

Carbon can also be removed from the terrestrial biosphere through lateral carbon fluxes (Randerson et al., 2002; Ciais et al., 2008). These include leaching of dissolved organic carbon and dissolved inorganic carbon from soils into rivers (1.7–2.7 PgC yr⁻¹, Ciais et al., 2013). Once in rivers, 0.2–0.6 PgC yr⁻¹ is buried in aquatic sediments, 0.8–1.2 PgC yr⁻¹ returns to the atmosphere as CO₂, and ~0.9 PgC yr⁻¹ is delivered to the ocean. Agricultural carbon harvested in croplands also undergoes lateral displacement to feedlots and urban areas (~1.3 PgC yr⁻¹, Ciais et al., 2007). Randerson et al. (2002) quantify non-CO₂ losses from terrestrial ecosystems to be 2.8–4.9 PgC yr⁻¹.

Terrestrial carbon balance

The terrestrial carbon balance describes the net flux of carbon in and out of terrestrial ecosystems. It is important to note that this is different from the net annual flux of CO₂ between the atmosphere and land, typically referred to as the “residual land sink”. The residual land sink (F_{residual}) is defined as:

$$F_{\text{residual}} = \Delta_{\text{atm}} - (F_{\text{ff}} + F_{\text{luc}} + F_{\text{ocn}}) \quad (1.4)$$

where Δ_{atm} is the atmospheric CO₂ growth rate, F_{ff} is fossil fuel emissions, F_{luc} is land use change emissions, and F_{ocn} is net ocean CO₂ flux. Therefore, the residual land sink will implicitly include all mechanisms by which carbon moves from the terrestrial biosphere to CO₂ in the atmosphere:

$$F_{\text{residual}} = \text{NEE} + F_{\text{BB}} + F_{\text{non-CO}_2}, \quad (1.5)$$

where F_{BB} is biomass burning CO₂ emissions, and $F_{\text{non-CO}_2}$ is carbon emissions as CO, CH₄ and VOCs, which are oxidized to CO₂ in the atmosphere. Ciais et al. (2013) give the residual land sink as -1.5 ± 1.1 PgC yr⁻¹ for 1980–1989 and -2.6 ± 1.2 PgC yr⁻¹ for 1990–2009. To calculate the terrestrial carbon balance from the residual land sink, exchange between the terrestrial biosphere and other reservoirs must be accounted for. After accounting for losses from the terrestrial biosphere to the ocean (~0.9 PgC yr⁻¹, Ciais et al., 2013), the residual land sink suggests that the amount of terrestrial biomass increased by 0.6 ± 1.1 PgC yr⁻¹ for 1980–1989 and 1.7 ± 1.2 PgC yr⁻¹ for 1990–2009. A schematic of carbon fluxes for the global carbon cycle is shown in Fig. 1.1 (reprinted from Ciais et al. (2014)).

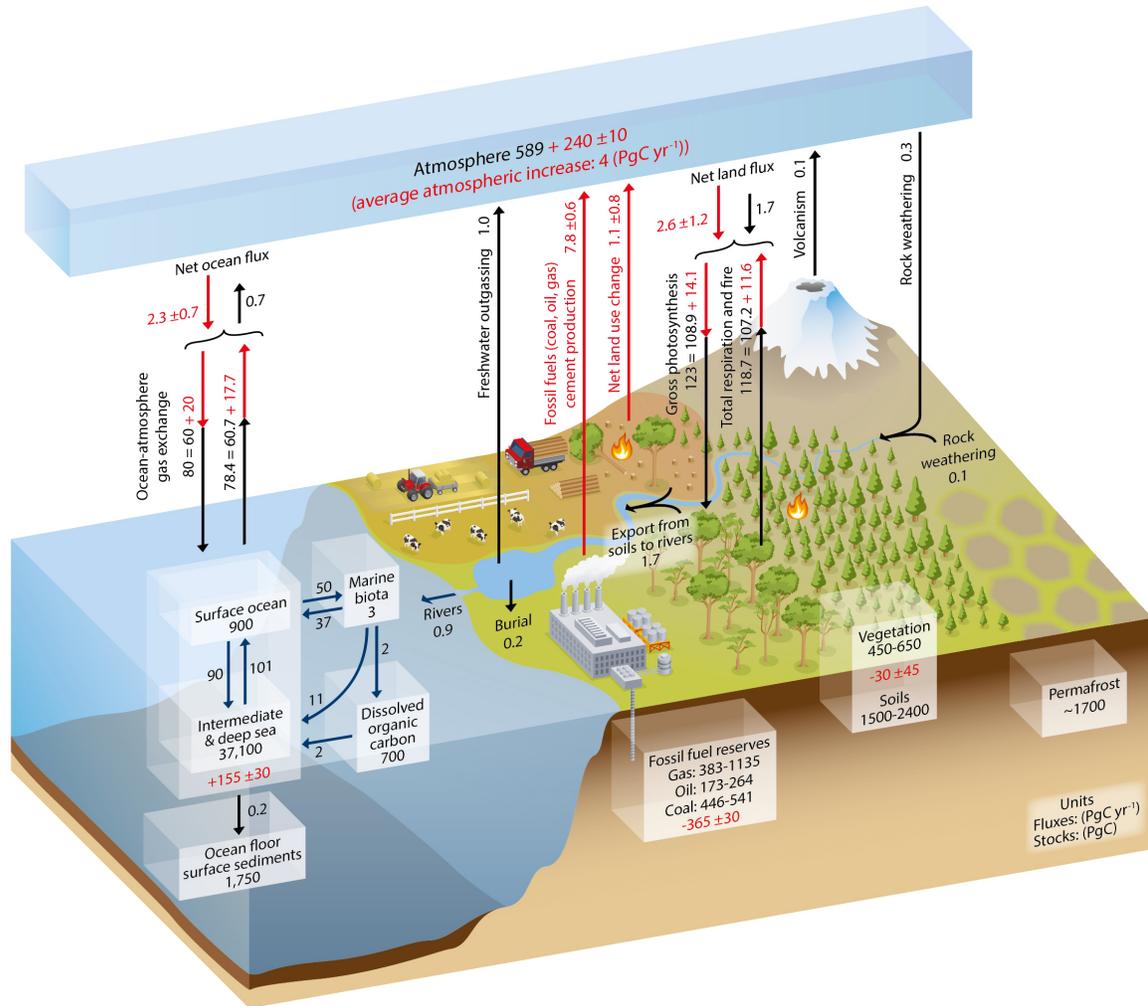


Figure 1.1: Simplified schematic of the global carbon cycle. Numbers represent reservoir mass (in PgC) and annual carbon exchange fluxes (in PgC yr⁻¹). Black numbers and arrows indicate reservoir mass and exchange fluxes estimated for the time prior to the Industrial Era, about 1750. Red arrows and numbers indicate annual ‘anthropogenic’ fluxes averaged over the 2000–2009 time period. These fluxes are a perturbation of the carbon cycle during Industrial Era post 1750. Reprinted from Figure 6.1 of Chapter 6 (Ciais et al., 2014) of the Working Group I Contribution to the Fifth Assessment Report (AR5) of the International Panel on Climate Change (IPCC).

1.1.2 Ocean

The ocean contains vast quantities of carbon. Roughly 38,000 PgC is stored as dissolved inorganic carbon, 700 PgC as dissolved organic carbon, and 3 PgC as marine biota (Ciais et al., 2013). The net flux of carbon between the atmosphere and ocean is driven by the partial pressure difference. Therefore, there are large one-way fluxes between the atmosphere and ocean of $\sim 80 \text{ PgC yr}^{-1}$, while the net uptake of carbon by the ocean is much smaller ($2.3 \pm 0.7 \text{ PgC yr}^{-1}$ for 2000–2009, Ciais et al., 2013).

Spatially, the partial pressure of ocean CO_2 is higher in upwelling regions, such as the eastern boundary of basins and in equatorial zones (McKinley et al., 2017), resulting in an out-gassing of CO_2 to the atmosphere. The equatorial Pacific is reported as a source of $0.44 \pm 0.41 \text{ PgC yr}^{-1}$. In contrast, there is significant uptake in western boundary currents (mainly due to the subduction of anthropogenic carbon-laden waters during deep vertical convection in winter) and at high latitudes (McKinley et al., 2017). In particular, strong uptake is reported for the North Atlantic ($0.3\text{--}0.59 \text{ PgC yr}^{-1}$ for $18\text{--}76^\circ \text{ N}$), North Pacific ($0.47 \pm 0.13 \text{ PgC yr}^{-1}$ for $18\text{--}60^\circ \text{ N}$), and Southern Ocean ($0.34 \pm 0.20 \text{ PgC yr}^{-1}$ south of 44° S) (McKinley et al., 2017).

IAV in ocean uptake is estimated to be $0.27\text{--}0.35 \text{ PgC yr}^{-1}$ (Le Quéré et al., 2018). It is largest in the equatorial Pacific and is strongly influenced by El Niño–Southern Oscillation (ENSO) variability (Landschuetzer et al., 2016; Chatterjee et al., 2017), characterized by a standard deviation of $0.15\text{--}0.20 \text{ PgC yr}^{-1}$ (McKinley et al., 2017). In extratropical regions, variability in uptake is largest on decadal scales. Decadal-scale variability in ocean uptake likely drives decadal variations in the combined uptake of CO_2 by the ocean and terrestrial biosphere (Landschuetzer et al., 2016).

1.1.3 Atmosphere

Atmospheric carbon is primarily in the form of CO_2 , with an atmospheric burden of $\sim 829 \text{ PgC}$ of CO_2 . Although the timescale of exchange with the surface for individual CO_2 molecules is on the order of a few years, net fluxes of CO_2 are mostly balanced on these timescales (Ciais et al., 2013). Net fluxes of CO_2 generally occur on much longer timescales of decades to millennia (Ciais et al., 2013). This long atmospheric lifetime implies that CO_2 is well-mixed in the atmosphere. Therefore, observations of the atmospheric CO_2 concentration at an individual location, distant from large local sources or sinks, can provide a constraint on the global mean concentration. The Mauna Loa Observatory, located at an altitude of 3397 m, has provided such observations since 1958. Figure 1.2 shows a timeseries of atmospheric CO_2 observations at this site. These observations

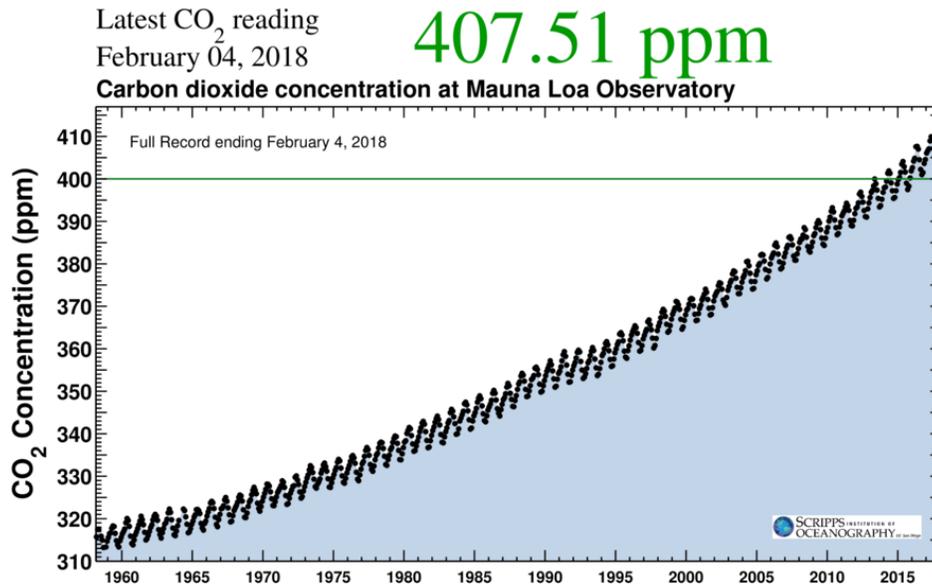


Figure 1.2: Atmospheric CO₂ concentration at Mauna Loa Observatory as a function of time (Scripps Institute of Oceanography, <https://scripps.ucsd.edu>).

show a secular year-on-year increase in atmospheric CO₂, which is primarily attributed to anthropogenic emissions of CO₂ to the atmosphere. The growth of atmospheric CO₂ accounts for about 44% of anthropogenic CO₂ emissions (mean for 1960–2010, Ciais et al., 2013), with the remainder sequestered into the ocean and terrestrial biosphere.

Figure 1.2 demonstrates that there are also variations in atmospheric CO₂ of ≤ 10 ppm on timescales of months to years. These variations are driven by the exchange of carbon between the atmosphere, biosphere and ocean. The largest source of variability is due to the seasonal cycle of the northern extratropical terrestrial biosphere (Sec. 1.1.1), whereby uptake by the biosphere during the growing season (May–September) reduces the atmospheric CO₂ concentration, while CO₂ is released to the atmosphere outside the growing season when GPP is low but R_e is still substantial. On interannual timescales, atmospheric CO₂ variability is primarily driven by the response of the ocean and terrestrial ecosystems to internal climate variability.

1.2 Top-down flux estimation

Due to the fact that anthropogenic emissions of CO₂ are reasonably well quantified relative to the natural fluxes, measurements of atmospheric CO₂ can be used to infer natural sources and sinks at Earth’s surface. Furthermore, on seasonal to interannual

timescales, variations in atmospheric CO₂ are dominated by natural fluxes, particularly from the terrestrial biosphere. Therefore, measurements of atmospheric CO₂ can be used to constrain carbon fluxes between the terrestrial biosphere and atmosphere, and thus better constrain the magnitude and timing of CO₂ fluxes originating from processes within the terrestrial biosphere.

1.2.1 History of flux estimation

Early investigations

In contrast to the present day, much of the early interest in observing CO₂ was motivated by its use as an atmospheric tracer. The influential Swedish meteorologist, Carl-Gustav Rossby had been a strong advocate of CO₂ measurements, with the intention of using these observations to track air masses (Keeling, 1998). Rossby started a Swedish program for CO₂ observations in 1955, which was a major factor in triggering interest in measuring CO₂ in the USA a few years later. However, the measurements obtained from the Swedish program were so imprecise (150–450 ppm) as to be useless and the program was abandoned (Keeling, 1998). The first study to reliably monitor atmospheric CO₂ was Keeling (1960), who deployed continuous atmospheric CO₂ monitoring stations in Antarctica, Hawaii, and California in combination with ship and aircraft based flask sampling. Keeling (1960) found that monthly variability in atmospheric CO₂ could be explained by the seasonal cycle of uptake and release of CO₂ by terrestrial land ecosystems and secular year-on-year growth due to anthropogenic emissions. These observations also demonstrated that atmospheric CO₂ showed little small-scale variability at locations distant from local sources or sinks, implying that atmospheric CO₂ variability is primarily due to large-scale fluxes coupled with atmospheric transport.

Not long after these first continuous CO₂ monitoring stations were established, attempts were made to quantify sources and sinks of CO₂. Bolin and Keeling (1963) analyzed the meridional seasonal cycle of CO₂ using data from the sites in Keeling (1960), and a new continuous measurement station at Point Barrow, Alaska. They applied an inversion method to constrain CO₂ fluxes and atmospheric transport. A forward model was constructed, giving the time rate of change of CO₂ (q) as a function of surface fluxes ($Q(\mu, t)$) and a vertically-averaged one-dimensional diffusion model transporting CO₂ meridionally:

$$\frac{\partial q}{\partial t} = \frac{K}{a^2} \frac{\partial}{\partial \mu} \left[(1 - \mu^2) \frac{\partial q}{\partial \mu} \right] + Q(\mu, t), \quad (1.6)$$

where K is a latitude-independent exchange coefficient, $\mu = \sin(\alpha)$, α being latitude, and

a is the radius of the Earth. Anthropogenic emissions were prescribed while both K and the natural sources and sinks of CO_2 were optimized; the latter was approximated with a truncated series of Legendre polynomials. Bolin and Keeling (1963) found a large source of CO_2 in the tropics ($\sim 11.0 \text{ PgC yr}^{-1}$) and a sink split evenly between the northern and southern hemispheres poleward of 30° ($\sim 5.5 \text{ PgC yr}^{-1}$), but conceded that these estimates were unreliable. They argued that greater observational coverage of CO_2 was required for robust estimates and concluded that CO_2 is an excellent tracer for the study of atmospheric mixing processes. Junge and Czeplak (1968), using a similar forward model, showed that simulated CO_2 concentrations were not very sensitive to variations of K or prescribed biospheric fluxes with latitude and concluded that the inverse problem was intractable. However, Enting and Mansbridge (1989) showed that large errors are introduced by the simple vertically-averaged one-dimensional diffusion model, with the largest source of error due to equating observed surface data to vertical averages in the model. Enting and Mansbridge (1989) showed that these errors are much reduced in a 2-D latitude-altitude model.

As well as performing the first CO_2 flux inversion, Bolin and Keeling (1963) showed that the rate of increase in atmospheric CO_2 was insufficient to account for all fossil fuel emissions. Previous studies, based on ^{14}C observations in wood, had already suggested that a significant fraction of fossil fuel emissions were likely absorbed by the ocean (Revelle and Suess, 1957; Bolin and Eriksson, 1959). Thus, at the time, it was generally believed that the reduced growth rate of atmospheric CO_2 was due to uptake by the ocean. However, it was also proposed that the terrestrial biosphere was sequestering carbon through the CO_2 fertilization effect (Bolin and Bischof, 1970).

The first 2-D model simulation of atmospheric CO_2 was performed by Machta (1972), using prescribed surface fluxes. Modeled CO_2 concentrations showed large differences from observed values. These differences were attributed to incorrect biospheric fluxes, and the author concluded that more realistic biospheric fluxes needed to be developed in order to produce more realistic simulations. Machta (1972) also noted that there seemed to be IAV present in the CO_2 growth rate. By the mid-1970s, IAV was clearly present in atmospheric CO_2 datasets (Keeling et al., 1976a,b), and this IAV was soon found to be closely associated with El Niño-Southern Oscillation (ENSO) variability (Bacastow, 1976; Bacastow et al., 1980).

In 1979, the so-called “Charney Report” was released (Charney et al., 1979). This was a report created by an ad hoc study group to assess the impact of increasing CO_2 on climate, and concluded that increasing CO_2 was likely to have substantial climate impacts. This report signifies the time around which the scientific community was re-

alizing that anthropogenic CO₂ emissions were resulting in significant global warming. Thus by this time, interest in monitoring atmospheric CO₂ was primarily motivated by the impact of rising CO₂ on Earth's climate.

Numerical modeling and inverse methods

In 1980, Pearman and Hyson (1980) used a 2-D model (latitude–altitude) to perform a quasi-inversion of atmospheric CO₂ data. A necessary assumption for this approach was that CO₂ observed at remote observing sites was representative of the zonal mean background. This assumption clearly broke down on many occasions, such as the phase shift between observations over the North Atlantic and at Barrow, Alaska (Pearman and Hyson, 1980). To perform their inversion, they first ran a forward model with prior biospheric fluxes. The fluxes were then iteratively adjusted based on a trial-and-error approach. From this approach, Pearman and Hyson (1980) obtained a rough estimate of the magnitude of uptake by the terrestrial biosphere in the northern extratropics during the growing season. This model was then further developed by Pearman et al. (1983) to estimate ocean uptake.

Around the same time, the first 3-D simulations of atmospheric CO₂ were published. The first 3-D CO₂ tracer transport study was performed by Fung et al. (1983) using the Goddard Institute for Space Studies (GISS) General Circulation Model (GCM) at $8^\circ \times 10^\circ$ horizontal resolution with nine vertical layers. They showed that zonal gradients exist in CO₂ and that the zonal mixing timescale is longer than the timescale for biospheric exchange. They suggested that IAV at monitoring stations could result from variability in atmospheric circulation, although later studies showed that it was primarily IAV in CO₂ sources and sinks that resulted in IAV in observed CO₂ (Law and Simmonds, 1996). Throughout the 1980s there were a number of studies that used tracer transport models to simulate atmospheric CO₂ fields based on prescribed surface fluxes (Fung, 1986; Fung et al., 1987; Keeling et al., 1989a; Taylor, 1989). These studies examined the spatiotemporal structures in atmospheric CO₂ and compared simulated atmospheric CO₂ with observations to evaluate estimated fluxes of the terrestrial biosphere.

In 1989, Enting and Mansbridge (1989) and Tans et al. (1989) used numerical inverse methods to assimilate surface CO₂ data and optimize surface fluxes using 2-D atmospheric tracer transport models. Both studies found that a large sink of CO₂ was required in the northern extratropics to reasonably reproduce the pole-to-pole CO₂ gradient. However, Tans et al. (1989) also argued that zonal variability in observed CO₂ precluded robust conclusions with a 2-D model. As an illustration of this technique, we will describe the method used by Tans et al. (1989). They first performed curve fitting on boundary layer

CO₂ data from background observing sites to obtain the meridional CO₂ distribution as a function of latitude. The fitted CO₂ was then used to prescribe the atmospheric CO₂ concentration in the lowest atmospheric layer. The model was then integrated forward in time and required surface fluxes were calculated.

The earliest CO₂ flux inversions performed using a 3-D transport model were by Keeling et al. (1989b) and Tans et al. (1990) using “synthesis” inversion approaches. In these studies, atmospheric CO₂ fields are simulated for a variety of sources (i.e., fossil fuel emission, NEE) individually. The linear combination of sources was estimated so that the calculated CO₂ fields matched the observed concentrations. Therefore, the overall strength of the sources could be varied but their spatiotemporal structure prescribed. Several years later, Enting et al. (1995) improved on this approach by applying a Bayesian technique to take into account prior uncertainties when weighting source contributions. Both Keeling et al. (1989b) and Tans et al. (1990) again found that a large sink in the northern extra-tropics was required to match the observed meridional CO₂ gradient, consistent with previous 2-D inversions. Tans et al. (1990) further argued that this sink must be due to the terrestrial biosphere, because the partial pressure difference in CO₂ between the surface ocean and atmosphere would have to be much larger than had been observed to provide an oceanic sink of this magnitude. They concluded that the global oceans sequestered less than 1 PgC yr⁻¹, while there must be a terrestrial sink of 2.0–3.4 PgC yr⁻¹. In a review of CO₂ uptake by the ocean, Siegenthaler and Sarmiento (1993) demonstrated that Tans et al. (1990) likely underestimated ocean uptake (Siegenthaler and Sarmiento (1993) suggested 2.0 ± 0.6 PgC yr⁻¹ and AR5 reported 2.0 ± 0.7 for 1980–1989 (Ciais et al., 2013)), but agreed that the atmospheric growth rate and uptake by the oceans are insufficient to account for CO₂ released by anthropogenic emissions, suggesting a terrestrial sink is in fact present. In the intervening years, further evidence has supported the idea that the CO₂ sink is split between the terrestrial biosphere and oceans, and this idea has generally gained favor over the earlier view that the oceans were almost entirely responsible for the uptake (Ciais et al., 2013).

Model errors

By the 1990s, the importance of employing 3-D transport models for interpreting observed atmospheric CO₂ was clear. Furthermore, increasing computational power had made such models more feasible than ever before. At the same time, it was recognized that there were considerable differences in inversion-based global CO₂ budgets arising from model transport (Law et al., 1996). To investigate differences in model transport, the CO₂ Transport Comparison Project (TransCom) was initiated in 1993. The results of a 12-

model inter-comparison were presented in Law et al. (1996). They found that models had substantial differences in the efficiency of the inter-hemispheric exchange, resulting in differences between models in the annual mean meridional gradient at the surface by a factor of two. This is a major problem for inferring the meridional distribution of annual net CO₂ sources and sinks. If the inter-hemispheric mixing is too vigorous, then the sink would be underestimated in the northern hemisphere and overestimated in the southern hemisphere (because the fossil fuel source is predominantly in the northern extra-tropics). Conversely, the opposite bias would be inferred if the mixing is too weak.

Modeled planetary boundary layer (PBL) dynamics were also found to have a significant impact on inferred fluxes by Denning et al. (1995). This is because insolation drives both mixing within the PBL and photosynthesis by vegetation, such that PBL height and NEE by the biosphere are highly correlated. Therefore, the PBL will be deeper during the day when net uptake of CO₂ by the biosphere is occurring. Conversely, the PBL will be shallow at night when there is net release of CO₂ by the biosphere. Under both circumstances, observed boundary layer CO₂ mole-fractions will be enhanced relative to a stationary mean PBL height. For this reason, this phenomena is commonly referred to as the “rectifier effect”. If variations in PBL height are not accurately modeled, systematic errors can be introduced. As an example, consider a model with a constant daily mean PBL height. If this model were to assimilate observations at a surface sampling site to optimize NEE, uptake would be underestimated during the day while release would be overestimated at night (Denning et al., 1999b). Furthermore, summer uptake would also be underestimated relative to winter release in extratropical regions (Denning et al., 1995).

In the second TransCom study, Denning et al. (1999a) performed SF₆ simulations with 11 tracer transport models. They found that most models were reasonably successful at simulating observed meridional gradients of SF₆ in the remote boundary layer, but found less agreement at continental sites. They also found that there were large differences in vertical transport between models. Differences in meridional gradients in SF₆ could not be solely attributed to meridional transport or inter-hemispheric mixing; instead a combination of meridional and vertical transport caused the differences. They suggest that differences among models were best explained by differences in sub-grid-scale parameterized vertical transport. The impact of vertical transport on flux inversions of surface data was further demonstrated by Stephens et al. (2007) and Yang et al. (2007). Stephens et al. (2007) used measurements of the vertical atmospheric CO₂ distributions at midday to evaluate flux inversion posterior CO₂ fields. They showed that, for many flux inversions, the vertical gradient in atmospheric CO₂ was too small during the sum-

mer, suggesting that these models had vertical mixing which was too strong and thus overestimated uptake by the biosphere when assimilating surface observations. In contrast, during the winter, the multi-model mean vertical CO₂ profile was consistent with observed concentrations. The models that best reproduced the observed CO₂ vertical profiles estimated weaker uptake by northern extratropical land ($-1.5 \pm 0.6 \text{ PgC yr}^{-1}$) in comparison to the multi-model mean ($-2.4 \pm 1.1 \text{ PgC yr}^{-1}$), which was compensated for by a smaller source in tropical land ($+0.1 \pm 0.8 \text{ PgC yr}^{-1}$) than the multi-model mean ($+1.8 \pm 1.7 \text{ PgC yr}^{-1}$).

Despite the significant difficulties related to model transport, the number of synthesis flux inversion studies increased dramatically through the 1990s and into the 2000s (Enting et al., 1995; Fan et al., 1998; Kaminski et al., 1999; Bousquet et al., 1999; Taguchi, 2000; Bousquet et al., 2000; Gurney et al., 2002; Peylin et al., 2002). Many of these studies attempted to constrain CO₂ fluxes on sub-continental-scale regions. However, the results were highly inconsistent in the net fluxes on regional scales. There are multiple reasons for this. Model transport errors play a role, but the inversion set-up and sparsity of observations also have large impacts on regional fluxes. Significant errors can be introduced through the aggregation of fluxes (Kaminski et al., 2001). In most inversions of the time, fluxes were optimized for a small number of large regions which helped regularize the inverse problem and save on limited computational resources. The problem with this method is that imposed structures within a spatial domain cannot be optimized. Kaminski et al. (2001) showed that the aggregation errors can be of the same order as the inferred flux. Similarly, Peylin et al. (2002) showed that aggregation errors can be introduced based on the temporal aggregation of optimized fluxes. Errors can also be introduced by the uneven spatiotemporal distribution of observations. It has been found that regional flux estimates are sensitive to the specific observing sites included in the analysis (Rödenbeck et al., 2003; Gurney et al., 2008; Bruhwiler et al., 2011). For example, Bruhwiler et al. (2011) found that the inclusion of newer European sites results in a large re-balancing of uptake from Europe to boreal Eurasia in comparison to an inversion with existing older sites. Unfortunately, model transport errors are still a major problem for flux inversions, and constraints on regional net fluxes remain sensitive to the inversion set-up.

Changes in the seasonal cycle amplitude

Around the same time as these developments in inverse modeling were occurring, a number of studies showed that observed CO₂ hinted at an increase in the amplitude of the seasonal cycle in the northern extratropics (Bacastow et al., 1985). Keeling et al.

(1996) showed that the amplitude of the seasonal CO₂ cycle had increased by 20% in Hawaii and 40% in the Arctic over three decades of measurements. They also found that the spring drawdown of CO₂ had advanced by about a week, suggesting a lengthening of the northern extratropical growing season. More recently, Graven et al. (2013) showed that the trend in the amplitude of the seasonal cycle of CO₂ has continued, increasing by ~50% between 45–90 °N.

The underlying mechanisms driving these changes in the seasonal cycle are not fully understood. Soon after the trend was discovered, it was suggested that it was partially explained by the CO₂ fertilization effect and/or increased winter R_e due to warming (Houghton, 1987; Kohlmaier et al., 1989). Myneni et al. (1997) showed that there is a positive trend in the normalized difference vegetation index (NDVI) over 1981–1991, consistent with a lengthening of the growing season between 45–70 °N and with increasing GPP. Similarly, increasing fall and early winter R_e has also been observed in northern ecosystems (Commane et al., 2017). Some studies link the trend primarily to boreal and Arctic ecosystems (Randerson et al., 1997; Graven et al., 2013), while others suggest that intensification of agriculture in cropland between 25°N and 60°N is the primary driver (Zeng et al., 2014).

Remote sensing of CO₂

During the 2000s, remote sensing instruments and techniques were developed and employed to measure CO₂ with sufficient precision and accuracy to provide useful information for carbon cycle science. These observations are made by fitting spectral absorption lines to retrieve abundances of atmospheric gases along an atmospheric path from a source to detector. Using the sun as a source and CO₂ absorption lines in the near infrared, this technique can be used to retrieve the column abundance of CO₂ with relatively uniform sensitivity throughout the troposphere. If O₂ is simultaneously retrieved, then the ratio of the column abundance of CO₂ to O₂ (scaled by the mean O₂ concentration) gives the column-averaged dry-air mole fraction of CO₂ (X_{CO_2}).

In anticipation of space-based spectrometers that would measure X_{CO_2} , the Total Carbon Column Observing Network (TCCON), a global network of ground-based Fourier transform spectrometers, was founded in 2004 (Wunch et al., 2011). This network was designed to be a validation network for space-based observing systems as well as for direct use in carbon cycle science, providing CO₂ observations with high precision (< 0.25% in CO₂). TCCON measurements are tied to the World Meteorological Organization (WMO) scale for in-situ measurements of CO₂ by performing comparisons with WMO-scale instrumentation aboard aircraft that measure atmospheric CO₂ profiles. Bias-corrected

TCCON measurements are of high accuracy, with biases of less than < 0.8 ppm (Wunch et al., 2010). Starting in the early 2000s, a number of space-based instruments were launched that could make atmospheric CO_2 measurements. However, it was not until the launch of the Greenhouse Gases Observing Satellite (GOSAT, 2009–present), that reliable space-based X_{CO_2} observations were made. Measurements of X_{CO_2} from GOSAT are performed in a similar manner to TCCON, and thus GOSAT observations can be validated using TCCON data in a direct manner. The main difference between these observing systems is that GOSAT observes solar radiation that is reflected off the Earth’s surface, rather than looking directly at the sun.

Constraints on surface fluxes provided by space-based observations of X_{CO_2} are substantially different from those provided by measurements within the boundary layer, which were widely used in earlier flux inversion studies. X_{CO_2} measurements reflect a greater influence of the large-scale, free tropospheric transport (Keppel-Aleks et al., 2011). In general, variability in X_{CO_2} is most strongly influenced by synoptic scale eddy-driven disturbances of the meridional CO_2 gradient (Keppel-Aleks et al., 2011). Therefore, an advantage of X_{CO_2} observations relative to surface observations is that they are somewhat less sensitive to model errors in PBL dynamics. A second advantage is that X_{CO_2} observations generally have smaller representativeness errors (caused by differences in temporal and spatial resolutions between the observations and the model). For surface observations, representativeness errors are typically much larger (~ 1 ppm, Rödenbeck et al., 2003) than instrumental errors (~ 0.2 ppm) and can be systematic. The main drawback of X_{CO_2} measurements is that they are difficult to validate. In situ observations can easily be calibrated with a standard gas, however, more challenging and expensive methods are required for X_{CO_2} validation. Furthermore, the observational coverage from satellites is so extensive that it is impractical to validate space-based X_{CO_2} observations directly. The method employed to evaluate space-based X_{CO_2} is through comparisons with TCCON, which is in turn validated with aircraft campaigns and Air-Core, a tube which collects ambient air samples while descending through the atmosphere from a high altitude (Karion et al., 2010). Unfortunately, small biases in retrieved X_{CO_2} can result in large biases in retrieved fluxes, which has made robust estimates of annual net fluxes on small scales challenging. Miller et al. (2018) calculate that the global mean absolute value of the atmospheric X_{CO_2} signal from biospheric fluxes is 0.5 ppm in February and 1.3 ppm in July, whereas satellite observations have systematic errors with a mean absolute value of 0.6 ppm in both February and July.

Space-based instruments have dramatically increased the number of measurements that can be assimilated in flux inversion studies. This increased data density has moti-

vated the development of more sophisticated flux inversion methods beyond the widely applied synthesis inversions. During the 2000s, both Kalman filter methods (Bruhwiler et al., 2005; Peters et al., 2005; Michalak, 2008; Feng et al., 2009) and variational assimilation methods (Chevallier et al., 2005; Baker et al., 2006b; Henze et al., 2007) were developed. These methods were then applied to GOSAT observations (Basu et al., 2013, 2014; Saeki et al., 2013; Maksyutov et al., 2013; Chevallier et al., 2014; Deng et al., 2014, 2016; Houweling et al., 2015). Many early inversions found substantial differences in annual net fluxes between inversions based on surface observations and those based on GOSAT observations (Chevallier et al., 2014; Houweling et al., 2015). Although differences are likely due to a variety of factors, it has become quite clear that there are regional-scale biases in the satellite observations that can result in net annual fluxes that are unphysical (Chevallier et al., 2014). Over time, the retrieval algorithms are being improved, so there is hope that many of the retrieval errors in the observations can be mitigated. Recent comparisons with aircraft observations suggest that GOSAT retrieval biases and model transport errors contribute roughly equally to errors in flux inversions (Frankenberg et al., 2016). Similar retrieval biases have been found with observations from the Orbiting Carbon Observatory-2 (OCO-2) (Wunch et al., 2017), which was launched in 2014. However, OCO-2 has much higher data density on small spatial scales than GOSAT. This high spatial density of OCO-2 measurements allows for more precise quantification of retrieval biases and may give insights into the causes of these biases.

Despite these challenges, space-based X_{CO_2} measurements have been successfully applied to address questions that are less sensitive to retrieval biases. An area of particular success is tropical carbon dynamics. Since the 1970s, it has been known that IAV in the atmospheric growth rate is primarily modulated by the impact of ENSO variability on the tropical carbon cycle. However, the poor spatial coverage of surface observing sites has meant that this variability is poorly observed. With the greater spatial coverage provided by GOSAT and OCO-2, variability in tropical CO_2 fluxes on regional scales has been investigated (Deng et al., 2016; Liu et al., 2017; Chatterjee et al., 2017; Bowman et al., 2017). Major advances have also occurred in the detection of large localized sources using X_{CO_2} from OCO-2 (Nassar et al., 2017; Schwandner et al., 2017).

Biosphere model optimization

An emerging area of research has been to optimize parameterizations within TBMs using flux constraints from atmospheric CO_2 observations. Wang et al. (2001) and Kaminski et al. (2002) were among the first to apply an inverse modeling approach to optimize

TBM parameters using CO₂ observations. Kaminski et al. (2002) coupled the Simple Diagnostic Biospheric Model (SDBM) to the TM2 CTM and tried to optimize the light use efficiency (LUE) and the Q₁₀ parameter, which is the factor by which respiration is enhanced for a 10 °C increase in temperature, for the 12 different biomes specified in the TBM. One of the main results of that study was that an increase in the LUE of the high-latitude deciduous forests and tundra biomes was required in the model to account for the observed seasonal cycle in the CO₂ data. The Kaminski et al. (2002) study was extended by Rayner et al. (2005), who replaced the SDBM model with the more complex Biosphere Energy Transfer Hydrology (BETHY) model, which incorporates a more process-based description of the terrestrial biosphere. This assimilation system has become known as the Carbon Cycle Data Assimilation System (CCDAS) and has been applied in a number of studies (Kaminski et al., 2010, 2012; Koffi et al., 2012, 2013; Ziehn et al., 2011, 2012). Kaminski et al. (2013) further developed CCDAS to assimilate soil moisture, fraction of absorbed photosynthetically active radiation and X_{CO₂} measurements within a single assimilation system. Independent of CCDAS, a number of other studies have attempted to optimize TBM parameters by assimilating atmospheric CO₂ measurements (Schuh et al., 2013; Chen et al., 2017).

1.2.2 Review of urban GHG studies

Urban emissions account for 37–49% of direct global GHG emissions, and a larger fraction for indirect emissions (Seto et al., 2014). This is despite the fact that urban areas only occupy 0.2–2.7% of ice-free land (Seto et al., 2014). Thus, mitigating GHG emissions from urban areas will play an important role in mitigating climate change. Bottom-up accounting of emissions will be required to monitor emission reductions, however, independent verification of these estimates will also be required. Measurements of GHG enhancements in urban areas provide an independent observational constraint to estimate emissions.

Over the past decade, a number of projects (Table 1.1) have been initiated to monitor GHG emissions in urban areas with the goal of providing top-down emission constraints. Monitoring urban GHG emissions using atmospheric observations is still at the experimental stage, with different approaches being applied for different projects. The differences in approaches are partially due to available resources and the specific problem addressed, but also because no approach is clearly superior to any other. A number of different measuring platforms have been used, including surface-based in-situ and flask measurements (Newman et al., 2013; McKain et al., 2015; Hopkins et al., 2016; Shuster-

Table 1.1: Examples of large city emission monitoring projects.

Project Name City	Website	References
INFLUX Indianapolis, U.S.A.	http://influx.psu.edu	Davis et al. (2017)
BEACO ₂ N California Bay Area, U.S.A.	http://beacon.berkeley.edu/	Shusterman et al. (2016)
Salt Lake City, U.S.A.	http://lair.utah.edu/urban.html	Pataki et al. (2006) Mitchell et al. (2018a) Bares et al. (2018)
Los Angeles, U.S.A.	http://megacities.jpl.nasa.gov	Newman et al. (2016) Verhulst et al. (2017)
CO ₂ -MegaParis Paris, France	http://co2-megaparis.lsce.ipsl.fr	Bréon et al. (2015) Stauffer et al. (2016) Xueref-Remy et al. (2018)

man et al., 2016; Miles et al., 2017; Verhulst et al., 2017; Bares et al., 2018; Xueref-Remy et al., 2018), remote sensing of solar absorption spectra (Wunch et al., 2009, 2016; Fu et al., 2014; Wong et al., 2015, 2016; Chen et al., 2016; Viatte et al., 2017; Hedelius et al., 2017, 2018), and aircraft-based in-situ observations (Wennberg et al., 2012; Cambaliza et al., 2015). Observations within the PBL are of particular relevance to the OP-FTIR system, and are the main focus of this section.

In-situ or flask observations within the PBL are the most widely applied observing technique and are used in all of the projects listed in Table 1.1. The Indianapolis Flux Experiment (INFLUX) employs 12 continuously sampling GHG sensors on towers ranging between 40 and 100 m above ground level (AGL), with two of the towers located outside of the city (Indianapolis) as background sites. The Berkeley Atmospheric CO₂ Observation Network (BEACO₂N) project employs a large number (~ 28) of moderate quality observing sites in the California Bay Area. The instrument sites were chosen in an ad-hoc approach based on site availability, with the majority of sites being deployed at schools and near the ground (< 10 m AGL) (Shusterman et al., 2016). The observing network in Salt Lake City has five continuous observing sites within the urban area and one background site, all of which are located near the surface (≤ 20 m AGL). These sites provide the longest timeseries of any urban GHG observing network (with multiple sites), allowing the investigation of decadal trends (Mitchell et al., 2018b). In Los Angeles, there are 16 instrument sites at locations chosen based on the analysis of Kort et al. (2013). The instruments are deployed at different heights for different sites: six sites have instruments at ≥ 50 m AGL, while six have instruments at ≤ 20 m AGL (Verhulst et al.,

2017). The network includes four background sites: two marine sites, one continental and one continental/mid-troposphere site. In Paris, instruments are deployed at six sites, with different sampling heights for different sites (4–317 m AGL) (Xueref-Remy et al., 2018).

Characterizing enhancements

A number of studies have investigated the atmospheric imprint of the urban emissions observed at urban sites relative to background sites. From this type of analysis, insights into the spatiotemporal structure of the urban GHG enhancements can be garnered. For Los Angeles, Verhulst et al. (2017) examined the enhancements of CO₂ and CH₄ at urban sites relative to marine background sites. They found that the marine background can be characterized to within ~ 1 ppm for CO₂ and ~ 10 ppb for CH₄, and that mid-afternoon enhancements near downtown Los Angeles reach ~ 15 ppm for CO₂ and ~ 80 ppb for CH₄. They concluded that analytical and background uncertainties are small relative to the urban enhancements, suggesting that local enhancements can be well observed. For INFLUX, Miles et al. (2017) examined the afternoon enhancements at urban tower sites relative to background sites during the dormant season and found significant urban enhancements of CO₂, CO and CH₄ relative to a background tower site. Similarly, in Paris, Xueref-Remy et al. (2018) found that sites nearer to the urban center have larger enhancements relative to the background sites, and low wind speeds result in larger enhancements for urban sites. These studies showed that boundary layer enhancements are generally significant and can be well characterized relative to background sites. However, there can still be significant differences between urban sites that are not directly related to emissions.

The sampling height of urban sites can have a significant impact on observed urban enhancements. Miles et al. (2017) examined the vertical gradient in GHGs for the subset of INFLUX towers that have multiple sampling heights. They found that differences in the sampling height affect the urban enhancements by up to 50% during the afternoon, and found that vertical gradients are largest at the downtown site. Similarly, Xueref-Remy et al. (2018) found significant differences in the diurnal cycle observed at the same station at 50 m and 180 m AGL in Paris. They attributed differences to strong stratification within the PBL at night, such that measurements higher in the PBL are less sensitive to nighttime emissions. They found that vigorous mixing results in a much reduced altitude gradient during the day (0.3 ppm). Haszpra et al. (2015) investigated the daytime vertical structure of CO₂ within the boundary layer using tall tower and aircraft measurements. They generated an expression to approximate the relative dry-

air mole fraction enhancement of CO₂ (C) within the boundary layer as a function of observing height (z):

$$C = 5.3 \frac{\overline{cw_s}}{w_*} \left(\frac{z}{z_i} \right)^{-\frac{1}{5}} - 0.7 \frac{\overline{cw_1}}{w_*} \left(1 - \frac{z}{z_i} \right)^{-1} + \text{constant}, \quad (1.7)$$

where $\overline{cw_s}$ is the vertical CO₂ flux at the surface, $\overline{cw_1}$ is the vertical CO₂ flux at the top of the mixed layer, w_* is the convective velocity scale, and z_i is the PBL height. The first term on the right-hand side represents the change in concentration due to sources/sinks at the surface, while the second term represents entrainment at the top of the boundary layer. Note that this equation cannot be applied very close to the surface or the top of the boundary layer as C approaches infinity (Haszpra et al., 2015). Equation (1.7) predicts that the urban GHG enhancement will increase rapidly for observing heights (z) approaching the surface, which is an important consideration for most urban observing networks.

Flux estimation

The main objective of monitoring urban GHG enhancements is to estimate GHG emissions. The simplest approach to estimate emissions is probably the mass-balance approach. In this approach, an air mass is observed before and after passing over a source. The change in GHG abundance is then inferred to be due to surface fluxes (Mays et al., 2009; Peischl et al., 2015; Super et al., 2017). This approach is most easily applied to aircraft measurements, where there is information on the vertical profile (Mays et al., 2009; Peischl et al., 2015), but has also been applied to ground-based in-situ measurements by Super et al. (2017) to estimate CO₂ and CO emissions from Rotterdam, the Netherlands. Super et al. (2017) deployed two surface in-situ sites, one upwind and one downwind of the city of Rotterdam, the Netherlands. Assuming the emissions were well-mixed within the PBL and neglecting entrainment, Super et al. (2017) estimated afternoon emissions consistent with city emission estimates for other cities. However, this type of analysis requires a number of assumptions, and significant biases could easily be present.

More sophisticated Bayesian inversion approaches have also been applied. Most urban flux estimation studies have employed a meso-scale inversion model. This technique has the potential to best utilize the information contained in GHG measurements, and has been applied previously to constrain urban CH₄ (Viatte et al., 2017) and CO₂ (Lauvaux et al., 2013, 2016) fluxes. However, as with global-scale flux inversions, errors in model transport can have a large impact on inferred fluxes (Deng et al., 2017). Furthermore,

dense observational coverage is required to fully constrain the many different sources of GHGs within an urban environment.

1.2.3 State of the science

In his reflections on his career, Charles “Dave” Keeling recalled that Roger Revelle, the director of the Scripps Institution of Oceanography in the 1950s, *“insisted that [a CO₂ sampling] aircraft project take priority over starting up measurements at Mauna Loa, because he believed that the main objective of [Keeling’s] program should be to gain a ‘snapshot’ of CO₂ around the world. He still held to the prevailing belief that the CO₂ concentration in air was spatially variable and that therefore sampling must be widespread to establish a reliable global average. Measurements should be repeated in, say, 20 years to see whether the global concentration had noticeably changed”* (Keeling, 1998). This quote illustrates how little was understood about the carbon cycle when this groundbreaking CO₂ observing project was being implemented. Significant progress in carbon cycle science has been made in the intervening 60 years. We now know that CO₂ is a well-mixed atmospheric species. Variability in seasonal to inter-annual timescales is driven by the terrestrial biosphere, while long-term growth in atmospheric CO₂ is driven by anthropogenic emissions. Furthermore, this growth is known to only account for roughly 44% of anthropogenic CO₂ emissions, with the remainder being sequestered by oceans and the terrestrial biosphere.

Yet many questions remain unanswered. There is now a lot of evidence supporting a terrestrial land sink large enough to imply that the amount of biomass is increasing, particularly in the northern extratropics. But it is unclear if this uptake is concentrated in certain ecosystems and climates. Furthermore, it is unclear why the increased uptake is occurring, and TBMs have large disagreements on the relative importance of different processes driving the uptake (Huntzinger et al., 2017). Questions also remain regarding the increase in the magnitude of the CO₂ seasonal cycle. For example, the relative contribution of boreal and Arctic ecosystems versus croplands to these changes is not well constrained.

Estimates of annual net fluxes on regional scales continues to be a major focus of inverse modelling studies. However, given that transport errors, biases in observations, and other factors continue to prevent reliable constraints on regional net fluxes, it may be time to reassess the questions that can be most effectively addressed with top-down flux estimations. In my opinion, the central question to answer is “How will the terrestrial carbon sink evolve in time?” To address this question, most top-down studies have

attempted to constrain the current spatial distribution of the terrestrial carbon sink. I suggest that using atmospheric CO₂ constraints to evaluate the performance of TBMs will yield more productive results.

Evaluating TBMs offers some advantages over estimating annual net fluxes. One advantage is that larger signals in the observations can be used. For example, model estimates of northern extratropical summer biospheric drawdown can be evaluated against observations. The signal of summer drawdown in observed CO₂ is large (~ 0.5 Pg week⁻¹) and thus constraints on this flux can be estimated with more confidence than smaller annual net fluxes (~ 0.02 Pg week⁻¹, assuming 1 Pg yr⁻¹ sink). These constraints can also be combined with independent constraints on GPP from SIF observations to evaluate both simulated GPP and R_e . A second advantage of evaluating TBMs is that they can be used to simulate changes in the terrestrial carbon sink into the future. The current spatial structure of the terrestrial carbon sink is likely transient. Increasing atmospheric CO₂, climate change, and nutrient availability are likely to impact the terrestrial carbon sink. Therefore, TBMs are required to understand the evolution of the land sink under climate change.

1.3 Thesis overview

This thesis has two aims: (1) to investigate the utility of atmospheric CO₂ observations in evaluating CO₂ fluxes in TBMs, and (2) to introduce a system to monitor anthropogenic GHGs from Toronto, Ontario.

The first objective involves three studies. The first study quantifies the sensitivity of different observing systems to surface fluxes of CO₂. The second study evaluates the mean seasonal cycle of GPP and R_e simulated by a set of TBMs. The final study quantifies how well flux inversions can constrain IAV within the carbon cycle, and whether these constraints can be used to evaluate TBMs.

For the second objective, the set-up of an open-path Fourier transform infrared spectroscopy (OP-FTIR) system in Toronto, Ontario is presented. Data collected from this system are compared with meteorological observations and a nearby GHG monitoring station run by Environment and Climate Change Canada (ECCC).

1.3.1 Key results

Four key results are obtained in this thesis:

1. Different CO₂ observing systems have large differences in their spatiotemporal sen-

- sitivity to surface fluxes. These differences have significant impacts on net annual CO₂ fluxes estimated from these observing systems, and these impacts must be considered when comparing flux inversions (Chapter 3).
2. Model-based seasonal cycles of R_e show systematic differences from optimized R_e constrained by atmospheric CO₂ and Solar Induced Fluorescence (SIF) measurements, with the models overestimating R_e during June–July and underestimating R_e during the fall. Further analysis suggests that biases may partially arise from neglecting seasonal variability in the CUE (Chapter 4).
 3. GOSAT flux inversions capture monthly NEE anomalies on continental-scales in the tropics and June–July–August anomalies on sub-continental-scales in the northern extra-tropics. Furthermore, GOSAT flux inversions capture anomalies in NEE better than most TBMs. Thus, constraints from flux inversions can be used to evaluate the response of TBMs to climate anomalies (Chapter 5).
 4. An OP-FTIR system for CO₂, CO, CH₄ and N₂O monitoring has been installed on the University of Toronto’s St. George Campus. The system has been recording spectra near continuously since November 2017 and will provide a valuable observational dataset which, in combination with other observing stations, can be used to provide a top-down constraint on GHG emissions from Toronto, ON (Chapter 6).

The significance of this research

It is well understood that TBMs are required to predict changes in the airborne fraction of CO₂, yet how to best evaluate and improve these models remains a topic of debate. The key results of this thesis demonstrate that top-down constraints on NEE are sufficient to evaluate TBMs, particularly when constraints from CO₂ observations are combined with SIF constraints on GPP. The results show that TBMs have significant deficiencies in simulating both the seasonal cycle of R_e and flux anomalies in NEE, motivating future investigations into how the model deficiencies can be corrected. Furthermore, the results highlight fundamental ecological processes that are currently missing for TBMs, particularly seasonal variability in the carbon use efficiency.

This thesis develops techniques that can be effectively used to identify deficiencies in TBMs in future studies. Using constraints from atmospheric CO₂ and SIF observations, it may be possible to evaluate modeled GPP and R_e on regional scales and this area of research should be further investigated. Furthermore, GOSAT flux inversions are found to provide a reasonable constraint on NEE anomalies over a range of scales. Future

studies can build on these results and use flux inversion NEE anomalies to evaluate model anomaly responses.

The OP-FTIR system installed in downtown Toronto is one of the first FTIR long-path systems for monitoring city emissions in the world. This system will provide an important component of a large GHG observation network being established in Toronto, and will be useful for monitoring emission reduction targets.

1.3.2 Outline

The remaining chapters of this thesis are as follows:

Chapter 2 provides an overview of the theory, models, and data-sets used in this thesis.

Chapter 3 examines the sensitivity of observations to surface CO₂ fluxes for the in-situ surface network, TCCON, GOSAT, and OCO-2. First, the spatial coverage of the observing systems is discussed in detail. Then the sensitivity of these observing systems to surface fluxes is calculated with the GEOS-Chem adjoint model. Finally, the implications for flux inversions are investigated with a series of observing system simulation experiments (OSSEs).

Chapter 4 evaluates the seasonal cycle of GPP, R_e and NEE produced by four TBMs and FLUXCOM over northern mid-latitude ecosystems. First, model GPP and NEE are evaluated using SIF and atmospheric CO₂ observations, respectively. Then an optimized seasonal cycle of R_e is calculated based on GPP and NEE constraints. Finally, optimized R_e is used to evaluate model R_e seasonal cycles.

Chapter 5 investigates how well GOSAT flux inversions can isolate NEE anomalies in comparison to TBMs. First, the agreement between GOSAT flux inversions and flux proxies, variables closely associated with NEE IAV, is quantified. Second, the spatial scales over which the GOSAT inversion constrains flux anomalies is examined. Finally, the sensitivity of the results to the inversion set-up is quantified.

Chapter 6 describes the installation of the OP-FTIR system, as-well as the measurements and data analysis.

Chapter 7 provides the conclusions of this thesis and suggestions for future work.

1.3.3 Contributions

The studies presented in this thesis were all performed by the author, but also benefited from the contributions of many co-authors. The contributions for each study are given here.

Chapter 3 was carried out by the author, Dylan Jones, Kim Strong, Zhao-Cheng Zeng, Feng Deng, and Junjie Liu. The study was designed by the author, Dylan Jones and Kim Strong, while the analysis was performed by the author. All co-authors provided feedback on the analysis. The adjustments to the GEOS-Chem adjoint model for these experiments were based on previous experiments by Junjie Liu and were based on her code. Changes from this code were performed by the author and Zhao-Cheng Zeng. The observation operator was adjusted from code developed by Feng Deng.

Chapter 4 was carried out by the author, Debra Wunch, Dylan Jones, Kim Strong, Feng Deng, Ian Baker, Philipp Köhler, Christian Frankenberg, Joanna Joiner, Vivek Arora, Bakr Badawy, Anna Harper, Thorsten Warneke, Christof Petri, Rigel Kivi, and Coleen Roehl. The study was designed by the author, Debra Wunch, Dylan Jones and Kim Strong, while the analysis was performed by the author. All co-authors provided feedback on the analysis. TBM fluxes were provided by Ian Baker, Vivek Arora, Bakr Badawy, and Anna Harper. GOME-2 SIF datasets were produced by Philipp Köhler and Joanna Joiner. Christian Frankenberg provided feedback on the application of these SIF datasets. TCCON data were provided by Thorsten Warneke, Christof Petri, Rigel Kivi, and Coleen Roehl.

Chapter 5 was carried out by the author, Dylan Jones, Kim Strong, Saroja Polavarapu, Anna Harper, David Baker, and Shamil Maksyutov. The study was designed by the author, Dylan Jones and Kim Strong, while the analysis was performed by the author. All co-authors provided feedback on the analysis. Joint UK Land Environment Simulator (JULES) fluxes were provided by Anna Harper.

Chapter 6 was carried out by the author, Kim Strong, Orfeo Colebatch, Debra Wunch, Dylan Jones, Pierre Fogal, Richard Mittermeier, David Griffith, and Doug Worthy. The OP-FTIR instrumentation was obtained on a long-term loan from ECCO. The initial planning for the use of the OP-FTIR system was performed by the author, Kim Strong and Orfeo Colebatch. The author and Orfeo Colebatch installed the system, aligned the components and have been running the system to record

spectra. Retrievals of recorded spectra have been performed by the author, as has all subsequent analysis. All co-authors have provided feedback on these steps.

Chapter 2

Theory, models and datasets

2.1 Chemical transport model

A chemical transport model (CTM) is a numerical model which simulates atmospheric chemistry. These models take meteorological fields as input and simulate chemicals as passive tracers in the atmosphere by solving the continuity equations for mass conservation. For simulations of CO₂, chemical reactions are generally neglected, as they have a relatively small impact on atmospheric CO₂ but are computationally expensive. Some studies include the atmospheric chemical production of CO₂ by oxidation of CO, CH₄ and VOCs by prescribing an atmospheric source (Nassar et al., 2010). For the studies in this thesis, no atmospheric chemical reactions are simulated nor is an atmospheric source of CO₂ prescribed.

2.1.1 GEOS-Chem model

The GEOS-Chem model (www.geos-chem.org) is a global 3-D chemical transport model driven by assimilated meteorology, which can run in several different configurations. For this work GEOS-Chem “Classic” is used, which is described below. It is driven by assimilated meteorology from the Goddard Earth Observing System of the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO). Both GEOS-5.2.0 and GEOS-FP (GEOS-5.7.2) reanalysis fields are used in this analysis. The GEOS-5.2.0 fields are generated on a $0.5^\circ \times 0.66^\circ$ rectilinear grid, while GEOS-FP reanalysis fields are generated with cubed-sphere c360 horizontal resolution (approximately $0.25^\circ \times 0.3125^\circ$ on a rectilinear grid) and a 7.5 min time step for advection and convection. For GEOS-Chem Classic, these fields are regridded to a rectilinear grid with 47 vertical levels and either $4^\circ \times 5^\circ$ or $2^\circ \times 2.5^\circ$ horizontal resolution (depending on

the specific model run). The fields are also temporally averaged to three-hour averages for 3-D fields and one-hour averages for 2-D fields. GEOS-Chem employs a 30 min time step for $4^\circ \times 5^\circ$ and 15 min time step for $2^\circ \times 2.5^\circ$.

To perform tracer transport, GEOS-Chem solves the equation of mass conservation for a scalar tracer, given by

$$\frac{\partial \rho q}{\partial t} + \nabla \cdot \rho q \mathbf{v} = F, \quad (2.1)$$

where q is the dry mixing ratio of a tracer, ρ is the air density, \mathbf{v} is the velocity vector, and F is the net source/sink. In GEOS-Chem, this is implemented using the numerical scheme developed by Lin and Rood (1996). A “pressure fixer” is also applied, whereby the horizontal velocity field is adjusted such that the vertically integrated divergence of mass matches the pressure tendency. This step is required because the temporal and spatial re-gridding of wind and pressure fields introduces a “mass-wind inconsistency”, where mass is not conserved (Jöckel et al., 2001).

Convection in GEOS-Chem is based on three-hourly net updraft and detrainment convective mass fluxes from the GEOS-5 simulation as described in Wu et al. (2007). PBL mixing in GEOS-Chem is instantaneous between the surface and boundary layer height. The PBL height is diagnosed from the archived GEOS-5 fields based on the bulk Richardson number with surface friction (Holtslag and Boville, 1993; Wu et al., 2007).

Transport differences from the native model

The lower spatiotemporal resolution of GEOS-Chem relative to GEOS-5/GEOS-FP is known to produce transport biases. Yu et al. (2018) showed that temporal averaging of archived meteorology results in a reduction of vertical transport (partially due to the loss of resolved convection) of up to 20% for GEOS-Chem at $0.25^\circ \times 0.3125^\circ$ relative to the GEOS-FP simulation. They also showed that significant errors are introduced from the remapping from cubed-sphere to rectilinear grid, and that reducing the resolution to $2^\circ \times 2.5^\circ$ results in further weakening of vertical transport. They found that ^{222}Rn concentrations (half life of ~ 4 days) are overestimated by up to 40% in the surface air and underestimated by 40% in the upper troposphere for $2^\circ \times 2.5^\circ$ relative to GEOS-FP c360 simulations.

Stanevich (2018a) showed that transport is further degraded between the $2^\circ \times 2.5^\circ$ and $4^\circ \times 5^\circ$ versions of GEOS-Chem. Based on comparisons between simulated and observed CH_4 fields, they found that differences were due to excessive mixing from enhanced numerical diffusion at $4^\circ \times 5^\circ$, which reduced the concentration gradients, particularly in the upper troposphere and lower stratosphere. However, the loss of resolved vertical

fluxes also resulted in reduced vertical mixing, particularly in mid-latitudes and northern latitudes.

No corrections have been made for these known biases in this thesis. Therefore, any systematic errors in these fields will impact the results. However, some sensitivity analysis has been performed by comparing flux inversion results with inversions using different transport models (Chapter 4) and by performing inversions at both $2^\circ \times 2.5^\circ$ and $4^\circ \times 5^\circ$ resolution (Chapter 5). The results of this thesis are found to be robust despite transport errors.

Prescribed CO₂ fluxes

The GEOS-Chem CO₂ simulation was first developed by Suntharalingam et al. (2004). This version of the CO₂ simulation contained no chemistry but included prescribed CO₂ fluxes from biomass burning, biofuel burning, fossil fuel burning and cement manufacture, ocean exchange and NEE. Nassar et al. (2010) expanded the GEOS-Chem CO₂ simulation emission inventories to include prescribed CO₂ surface emissions from shipping, 3-D spatially-distributed emissions from aviation, and 3-D chemical production of CO₂. However, in this thesis, no atmospheric chemical production of CO₂ is prescribed. For all simulations performed in this dissertation, the following CO₂ fluxes are used. Monthly ocean fluxes are from Takahashi et al. (2009), anthropogenic emissions are from Andres et al. (2016), and biomass burning emissions are from the Global Fire Emission Database GFEDv3 (van der Werf et al., 2006). Prescribed NEE fluxes are different for different experiments and are introduced in each chapter.

2.2 Bayesian data assimilation

Bayesian inversion methods are used in both remote sounding and flux inversion applications. This section provides an overview of the Bayesian data assimilation formulation and its application in remote sounding (Sec. 2.2.1) and 4D-Var flux inversions (Sec. 2.2.2). The description presented here follows from Rodgers (2000), Henze et al. (2007) and Connor et al. (2008).

In an inverse problem, the objective is to obtain improved knowledge of a state (\mathbf{x}) given a set of observations (\mathbf{y}) which are related by a forward model,

$$\mathbf{y} = H(\mathbf{x}) + \boldsymbol{\epsilon}, \quad (2.2)$$

where $\boldsymbol{\epsilon}$ represents measurement and forward model errors. Inverse problems are typically

ill-posed, such that the solution is underdetermined by the observational constraints. In this case, regularization of the problem is required to produce a unique solution and prevent overfitting of the data. Bayesian inference provides a mechanism to regularize inverse problems using prior knowledge of the system. By Bayes' theorem, the probability of \mathbf{x} given \mathbf{y} , $P(\mathbf{x}|\mathbf{y})$, is given by:

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}. \quad (2.3)$$

In this thesis, the probability density functions $P(\mathbf{y}|\mathbf{x})$ and $P(\mathbf{x})$ are assumed to be Gaussian, and are expressed as a function of prior knowledge of the state (\mathbf{x}^b):

$$P(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}_x^{-1}(\mathbf{x} - \mathbf{x}^b)\right), \quad (2.4)$$

$$P(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}))\right), \quad (2.5)$$

where \mathbf{B}_x and \mathbf{R} are the prior state and measurement error covariances matrices. With these expressions, Eq. 2.3 can be re-written as

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}_x^{-1}(\mathbf{x} - \mathbf{x}^b) - \frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}))\right). \quad (2.6)$$

From Eq. 2.6, the Bayesian cost function can be defined as:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}_x^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x})). \quad (2.7)$$

For this thesis, we calculate the maximum a posteriori (MAP) state (\mathbf{x}^a), which is the most probably state given a set of observations and prior constraints. The MAP state is obtained by maximizing $P(\mathbf{x}|\mathbf{y})$ and is calculated by minimizing the Bayesian cost function (Eq. 2.7). There are several ways to approach this minimization problem. In optimal estimation (OE), an analytic solution is obtained (Sec. 2.2.1). In 4D-Var, this is performed numerically using an adjoint model (Sec. 2.2.2).

2.2.1 Remote sounding approach

In remote sounding applications, the state vector typically consists of the atmospheric profiles of the desired gas (\mathbf{u}) and interfering species that are also retrieved (\mathbf{e}),

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{e} \end{pmatrix}. \quad (2.8)$$

Typically, the retrieval of the \mathbf{x}^a is performed using a variant of the OE method. In OE, the posterior probability distribution is assumed to take the form,

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^a)^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}^a)\right), \quad (2.9)$$

where \mathbf{P} is the posterior error covariance, and the forward model is assumed to be linear,

$$\mathbf{y} = H(\mathbf{x}) + \boldsymbol{\epsilon} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}, \quad (2.10)$$

where

$$\mathbf{H} = \frac{\partial H(\mathbf{x})}{\partial \mathbf{x}}. \quad (2.11)$$

In this case, an analytic solution can be obtained for \mathbf{x}^a by minimizing the cost function (Eq. 2.7),

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{A}(\mathbf{x} - \mathbf{x}^b), \quad (2.12)$$

where

$$\mathbf{A} = \mathbf{B}_x \mathbf{H}^T (\mathbf{H} \mathbf{B}_x \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}. \quad (2.13)$$

\mathbf{A} is the averaging kernel matrix and gives the weighting of the observations and prior information in calculating the posterior state. In many remote sensing applications, the total column dry-air mole-fraction of the desired gas (X_{gas}^a) is also calculated. To perform this calculation, a pressure weighting function (\mathbf{h}) is required, which performs the mapping from an atmospheric profile to X_{gas} :

$$X_{\text{gas}}^a = \mathbf{h}^T \mathbf{x}^a. \quad (2.14)$$

Now, to obtain X_{gas}^a from the subset of the state vector containing the gas of interest (\mathbf{u}), the column averaging kernel (\mathbf{a}_{gas}) can be defined as:

$$(\mathbf{a}_{\text{gas}})_j = \frac{\partial X_{\text{gas}}^a}{\partial \mathbf{u}_j} \frac{1}{\mathbf{h}_j} = (\mathbf{h}^T \mathbf{A})_j \frac{1}{\mathbf{h}_j}, \quad (2.15)$$

where j goes from 1 to the number of atmospheric levels. Then, X_{gas} is given by:

$$X_{\text{gas}}^a = X_{\text{gas}}^b + \mathbf{h}^T \mathbf{a}_{\text{gas}}^T (\mathbf{u} - \mathbf{u}^b), \quad (2.16)$$

where X_{gas}^b is the a priori column dry-air mole fraction used in the remote sounding retrieval (see Connor et al. (2008) for more details).

2.2.2 4-D variational data assimilation

In 4D-Var applications, \mathbf{x} typically represents the 3-D atmospheric state of a gas of interest. Now, consider a forward transport model (M) which transports the state forward in time from t to $t + 1$:

$$\mathbf{x}_{t+1} = M(\mathbf{x}_t, \mathbf{p}), \quad (2.17)$$

where \mathbf{p} signifies model parameters such as a surface flux. From this definition, Eq. 2.7 can be extended to be a function of a model parameter and include observations at multiple times,

$$J(\mathbf{x}_0, \mathbf{p}) = \frac{1}{2}(\mathbf{p} - \mathbf{p}^b)^T \mathbf{B}_{\mathbf{p}}^{-1}(\mathbf{p} - \mathbf{p}^b) + \sum_{t=t_0}^N \frac{1}{2}(\mathbf{y}_t - H(\mathbf{x}_t))^T \mathbf{R}^{-1}(\mathbf{y}_t - H(\mathbf{x}_t)), \quad (2.18)$$

where $\mathbf{B}_{\mathbf{p}}$ is the prior error covariance for parameter \mathbf{p} . Equation 2.18 gives the 4D-Var cost function for \mathbf{p} . The MAP for \mathbf{p} can now be obtained by minimizing $J(\mathbf{x}_0, \mathbf{p})$. To minimize a given cost function over an extended time domain with 4D-Var, the sensitivity of the cost function to the parameter being optimized must be known. The GEOS-Chem adjoint model provides a mechanism to calculate the gradient of a cost function with respect to a set of model parameters ($\nabla_{\mathbf{p}} J$) as described in Henze et al. (2007). This model calculates the local Jacobian of the forward model with respect to the model state or a model parameter around a given time step, t , given by:

$$\mathbf{M}_x^t = \frac{\partial M(\mathbf{x}^t, \mathbf{p})}{\partial \mathbf{x}^t} = \frac{\partial \mathbf{x}^{t+1}}{\partial \mathbf{x}^t}. \quad (2.19)$$

Using the chain rule, the adjoint of the Jacobian can be used to calculate sensitivities over multiple time steps from time step t to time step N :

$$\frac{\partial \mathbf{x}^N}{\partial \mathbf{x}^t} = (\mathbf{M}_x^t)^T (\mathbf{M}_x^{t+1})^T \dots (\mathbf{M}_x^{N-1})^T \left(\frac{\partial \mathbf{x}^N}{\partial \mathbf{x}^{N-1}} \right). \quad (2.20)$$

For 4D-Var, the adjoint model is used to calculate the gradient of Eq. 2.18 with respect to the parameter being optimized. As an example, consider a cost function that is only evaluated at the final time step. The gradient of this cost function will be given by:

$$\nabla_{\mathbf{p}} J = \lambda_{\mathbf{p}}^0 + \mathbf{B}_{\mathbf{p}}^{-1}(\mathbf{p} - \mathbf{p}^b), \quad (2.21)$$

where $\lambda_{\mathbf{p}}^0$ is the sensitivity of $J(\mathbf{x}^N, \mathbf{p})$ to model variable \mathbf{p} and is calculated by iteratively solving:

$$\lambda_{\mathbf{p}}^{t-1} = \left(\frac{\partial \mathbf{x}^{t-1}}{\partial \mathbf{p}} \right)^T \frac{\partial J(\mathbf{x}^N, \mathbf{p})}{\partial \mathbf{x}^{t-1}} + \lambda_{\mathbf{p}}^t. \quad (2.22)$$

The adjoint model minimizes Eq. 2.18 iteratively using the Limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm for Bound-constrained optimization (L-BFGS-B, Byrd et al., 1995). For the studies in this thesis, the GEOS-Chem adjoint model is employed to optimize CO₂ surface fluxes, namely NEE and oceanic fluxes. This is performed by optimizing monthly scaling factors of prior flux estimates. In this set-up, $\mathbf{B}_{\mathbf{p}}$ is given as a percentage of the prior flux estimate rather than an absolute value. Furthermore, both $\mathbf{B}_{\mathbf{p}}$ and \mathbf{R} are assumed to be diagonal for all of the inversions performed in this thesis.

Practical assimilation of observations in 4D-Var

Equation 2.18 has a term giving the difference between measured and modeled observations, referred to as the innovation,

$$\mathbf{y}_t - H(\mathbf{x}_t). \quad (2.23)$$

For remote sounding observations, this term would be the difference between observed spectra and modeled spectra based on the modeled atmospheric state. However, in the practical 4D-Var implementation, this term is calculated by taking the differences between a previously retrieved total column dry-air mole fraction of the desired gas, X_{obs}^a , and total column dry-air mole fraction calculated based on the model state, X_{model}^a .

This is done in several steps. First, the modeled gas profile (\mathbf{x}) is interpolated from the model pressure levels to the pressure levels used by the forward model of the remote sounding retrieval. The interpolated modeled gas profile is referred to as \mathbf{x}_H . Then X_{model}^a is retrieved:

$$X_{\text{model}}^a = X_{\text{obs}}^b + \mathbf{h}^T \mathbf{a}_{\text{gas}}^T (\mathbf{x}_H - \mathbf{x}_H^b), \quad (2.24)$$

where X_{obs}^b is the a priori column dry-air mole-fraction used in the remote sounding

retrieval, \mathbf{h} is the pressure weighting function, \mathbf{a}_{gas} is the column averaging kernel, and \mathbf{x}_H^b is the remote sounding prior profile. Finally, the innovation is calculated:

$$\mathbf{y}_t - H(\mathbf{x}_t) = X_{\text{obs}}^a - X_{\text{model}}^a. \quad (2.25)$$

2.3 CO₂ datasets

2.3.1 Surface measurements

Surface measurements can be performed either using an in-situ gas analyzer or by taking a flask sample, which is then returned to a lab and analyzed. A number of different groups from around the world collect surface CO₂ observations. In this thesis, surface measurements are taken from the Obspack PROTOTYPE package (Masarie et al., 2014; Project, 2013). This package incorporates data from many observing sites around the world. The PROTOTYPE product includes actual data as well as derived data (averages) specifically prepared for the CarbonTracker CO₂ data assimilation system (Peters et al., 2007). The product includes 190 data sets with contributions from 20 laboratories. Surface observations used in this thesis are taken from the PROTOTYPE product (<https://www.esrl.noaa.gov/gmd/ccgg/obspack/>).

There are additional surface measurements which are not included in the ObsPack dataset (nor in this thesis), but could be used to fill in some spatial gaps in the coverage of surface measurements. A notable example is the Japan-Russia Siberian Tall Tower Inland Observation Network of nine tower sites in Siberia (Sasakawa et al., 2010, 2013). There are also several surface sites in the Amazon which are not included in the ObsPack dataset (Molina et al., 2015).

2.3.2 Remote sounding measurements

Remote sounding methods retrieve CO₂ mole fractions from observations of absorption spectra. An illustration of this technique can be found in Chapter 6, where atmospheric mole fractions are retrieved from recorded infrared spectra.

TCCON

The TCCON is a network of ground-based Fourier transform spectrometers that record solar absorption spectra in the near-infrared, from which X_{CO_2} columns are retrieved (Wunch et al., 2011). The TCCON has adopted the Bruker Infrared Fourier Spectrometer (IFS) 125HR as its preferred instrument, as this is the most robust and stable

high-resolution FTS commercially available. TCCON sites make repeated measurements of solar spectra (every few minutes) under clear-sky conditions. The spectra are recorded with 45 cm maximum OPD ($\sim 0.02 \text{ cm}^{-1}$ spectral resolution). CO_2 abundances are retrieved using a non-linear least squares approach from absorption lines in the near infrared ($6180.00\text{--}6260.00 \text{ cm}^{-1}$ and $6297\text{--}6382 \text{ cm}^{-1}$ for CO_2 and $7765.00\text{--}8005.00 \text{ cm}^{-1}$ for O_2). The column-averaged dry-air mole fractions of CO_2 (X_{CO_2}) are retrieved by taking the ratio of the column abundance of CO_2 to O_2 (scaled by the mean O_2 concentration), resulting in high precision ($< 0.25\%$ in CO_2) X_{CO_2} measurements. TCCON measurements are tied to the World Meteorological Organization (WMO) scale by performing comparisons with WMO-scale instrumentation aboard aircraft that measure atmospheric CO_2 profiles. Bias-corrected TCCON measurements are of high accuracy, with biases of less than $< 0.8 \text{ ppm}$ (Wunch et al., 2010).

For this thesis, TCCON data were obtained from the TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center (CDIAC). However, the TCCON Data Archive has moved and is now hosted by CaltechDATA [[http https://tccodata.org](http://tccodata.org)]. Table 2.1 lists the sites used in this thesis.

GOSAT

GOSAT was launched in February 2009 in a sun-synchronous orbit, with a repeat cycle of 3 days that produces 44 separate ground track repeats (Yoshida et al., 2013). The Thermal And Near-infrared Sensor for carbon Observation Fourier Transform Spectrometer (TANSO-FTS) aboard GOSAT is a double-pendulum-type interferometer with two cube-corner reflectors. The instrument has an OPD of $\pm 2.5 \text{ cm}$, resulting in $\sim 0.2 \text{ cm}^{-1}$ spectral resolution (Kuze et al., 2009). The instrument measures reflected solar radiation, from which X_{CO_2} is estimated. The footprint of the GOSAT measurements has a diameter of about 10 km. The instrument has a pointing mechanism capable of rotating $\pm 35^\circ$ across-track and $\pm 20^\circ$ along-track, and observations can be obtained in nadir viewing mode, glint mode, and target mode. Since August 2010, in nadir mode, TANSO-FTS has been measuring with a three-point cross-track pattern with 263 km cross-track separation, resulting in a swath of 526 km. Measurements have an along-track separation of 283 km (Crisp et al., 2012). Over the ocean where surface reflectance is small, the pointing mechanism views a widely spread sun-glint area, where specular reflection occurs and reflectance is high. As the along-track viewing angle range is $< 20^\circ$, glint observation is limited to low and middle latitudes.

In this thesis, NASA Atmospheric CO_2 Observations from Space (ACOS) GOSAT retrievals are used. Depending on the study, either version 3.4 or 3.5 is used; further pro-

Table 2.1: TCCON sites used in this study.

Site Name	Lat	Lon	Reference
Eureka, Canada	80.05°N	86.42°W	Strong et al. (2017)
Sodankyla, Finland	67.37°N	26.63°E	Kivi et al. (2014)
Bialystok, Poland	53.23°N	23.03°E	Deutscher et al. (2014)
Bremen, Germany	53.10°N	8.85°E	Notholt et al. (2014)
Karlsruhe, Germany	49.10°N	8.44°E	Hase et al. (2014)
Paris, France	48.49°N	2.36°E	Te et al. (2014)
Orleans, France	47.97°N	2.11°E	Warneke et al. (2014)
Garmisch, Germany	47.48°N	11.06°E	Sussmann and Rettinger (2014)
Park Falls, USA	45.95°N	90.27°W	Wennberg et al. (2014b)
Rikubetsu, Japan	43.46°N	143.77°E	Morino et al. (2014b)
Lamont, USA	36.60°N	97.49°W	Wennberg et al. (2016)
Tsukuba, Japan	36.05°N	140.12°E	Morino et al. (2014a)
Dryden, USA	34.96°N	117.88°W	Iraci et al. (2016)
Caltech, USA	34.14°N	118.13°W	Wennberg et al. (2014a)
Saga, Japan	33.24°N	130.29°E	Kawakami et al. (2014)
Izana, Tenerife, Spain	28.3°N	16.5°W	Blumenstock et al. (2014)
Manaus, Brazil	3.21°S	60.60°W	Dubey et al. (2014)
Ascension Island	7.92°S	14.33°W	Feist et al. (2014)
Darwin, Australia	12.42°S	130.90°E	Griffith et al. (2014a)
Reunion Island	20.90°S	55.49°E	De Mazière et al. (2014)
Wollongong, Australia	34.41°S	150.88°E	Griffith et al. (2014b)
Lauder, New Zealand	45.04°S	169.68°E	Sherlock et al. (2014)

cessing of the data is described within each chapter. The ACOS algorithm uses three spectral bands, the O₂ A band (12,950–13,910 cm⁻¹), the weak CO₂ band (6166–6286 cm⁻¹), and the strong CO₂ band (4810–4897 cm⁻¹) (O’Dell et al., 2012). The column abundances of CO₂ and O₂ are retrieved using an OE approach. Column-averaged dry-air mole fractions of CO₂ are then obtained based on the ratio of retrieved CO₂ to O₂ abundances. Additional information on the ACOS retrieval algorithm is available in O’Dell et al. (2012) and Crisp et al. (2012). NASA ACOS GOSAT lite files are obtained from the CO₂ Virtual Science Data Environment (<https://co2.jpl.nasa.gov/#mission=ACOS>)

OCO-2

The OCO-2 spacecraft, launched in July 2014, carries and points a single instrument. The OCO-2 instrument incorporates three co-boresighted, long-slit imaging grating spectrometers optimized for the O₂ A band at 12,950–13,910 cm⁻¹ and the CO₂ bands at 6166–6286 cm⁻¹ and 4810–4897 cm⁻¹. OCO-2 has a swath of 10 km and collects eight adjacent spatially-resolved samples, giving a footprint of about 3 km². Measurements are recorded every 0.333 seconds, resulting in roughly 24 soundings per second. The spacecraft is in a sun-synchronous orbit with an equator crossing time near local noon (13:36, Crisp, 2015) and has a repeat cycle of 16 days, resulting in 233 separate ground track repeats. For routine science operations, the instrument’s bore sight is pointed to the local nadir or at the glint spot and measures reflected solar radiation. Although, OCO-2 originally alternated between glint and nadir modes on a 16-day repeat cycle, this was found to be inefficient. In early July 2015, this observation strategy was modified to alternate between glint and nadir observations on alternate orbits. OCO-2 also has a target model which allows the spacecraft to target selected surface calibration and validation sites to collect thousands of soundings as the spacecraft flies overhead.

Throughout this thesis, version 7 of the ACOS OCO-2 lite files are used. These data were downloaded from the CO₂ Virtual Science Data Environment (<https://co2.jpl.nasa.gov/#mission=OCO-2>). In addition to the quality flag, which filters data based on comparisons with truth proxies (i.e., TCCON, the Southern Hemisphere approximation, and small area analysis), ACOS OCO-2 measurements are provided with a Warn Level (WL) filter. WLs are designed to minimize the variance of X_{CO₂} in regions of small atmospheric variability (Mandrake et al., 2013), and are ordered in decreasing data quality from 0 being the best to 19 being the worst. The spatiotemporal distribution of OCO-2 observations for different WL cutoffs are discussed in Chapter 3.

Chapter 3

Sensitivity of CO₂ surface flux constraints to observational coverage

This chapter is adapted from:

Byrne, B., Jones, D. B. A., Strong, K., Zeng, Z.-C., Deng, F., and Liu, J. (2017). Sensitivity of CO₂ surface flux constraints to observational coverage. *J. Geophys. Res.-Atmos*, 112(12):6672–6694. <https://doi.org/10.1002/2016JD026164>

3.1 Introduction

Measurements of atmospheric CO₂ from a variety of observing systems (e.g., surface sites, TCCON, GOSAT and OCO-2) are commonly assimilated into inversion systems to optimize NEE fluxes. However, these observation systems strongly under-constrain surface fluxes on regional scales. This implies that the optimization of NEE will be sensitive to the spatiotemporal distribution of observations, and to the sensitivity of these observations to surface fluxes.

Previous flux inversion studies using surface in-situ observations have noted high sensitivity of flux inversion results to the distribution of observations. Several studies have found that including or excluding certain surface sites can result in systematic differences in regional flux estimates (Rödenbeck et al., 2003; Gurney et al., 2008; Bruhwiler et al., 2011). For example, Bruhwiler et al. (2011) found that the inclusion of newer European sites results in a large re-balancing of uptake from Europe to boreal Eurasia in comparison to an inversion with existing older sites. Rödenbeck et al. (2003) found that using different sets of surface stations in their 20-year flux inversion results in regional flux differences that show up predominately as shifts in the long-term mean flux, while the timing and amplitude of interannual flux anomalies remain comparatively consistent. Reuter et al.

(2014) argued that the heterogeneous distribution of surface observations over Europe, with sites concentrated in Western Europe where there is likely a weaker sink than Eastern Europe, results in a systematic underestimate of the net continental sink in comparison to that inferred from observations from GOSAT, although Feng et al. (2016) show that realistic biases in GOSAT data could also explain the disagreement.

Flux inversions using satellite data have also been shown to be susceptible to the spatio-temporal distribution of observations. Because satellites measure column CO_2 using reflected sunlight, the spatial distribution of observations by satellites shifts seasonally with the declination of the sun. Since fluxes also vary seasonally, this leads to systematic errors in annual fluxes. This was demonstrated by Liu et al. (2014) using an Observing System Simulation Experiment (OSSE). In their simulation, they defined a set of "true" fluxes and selected a priori fluxes with the same global annual net flux (-5.3 PgC), but with a different spatial and temporal distribution. Liu et al. (2014) retrieved a net flux that is 0.7 PgC lower than the true and a priori value by sampling simulated atmospheric CO_2 with the distribution of GOSAT observations. This bias was due to the preferential sampling of the summer hemisphere and due to the fact that GOSAT always makes measurements at the same local time (12:45-13:15). The time of day of measurements resulted in a bias because the a priori diurnal cycle was weaker than the true diurnal cycle (Liu et al., 2014). Liu et al. (2014) show that this bias can be introduced despite the fact that X_{CO_2} sampled at 13:00 p.m. (local time) has similar magnitude to daily average X_{CO_2} (Miller et al., 2007).

Liu et al. (2014) also examined the impact of GOSAT sampling on fluxes estimated for the regions used in the TransCom Project (Gurney et al., 2002). They found that even though the bias in the a posteriori monthly flux is reduced for every region during every month, the annual mean flux is degraded for some regions. This phenomena can be illustrated by considering a true flux that has a larger seasonal cycle than the a priori flux, such as the North American boreal region in Liu et al. (2014). During the summer months, the sensitivity of observations to fluxes is high, so that the a posteriori flux is pulled down to the true flux. In contrast, there is little sensitivity in the winter, so that the a posteriori fluxes are only slightly increased. In this case, the flux estimate will improve for each month but the annual mean flux will be biased low in comparison to the a priori estimate.

These studies demonstrate the importance of understanding the spatial and temporal sensitivity of observations to surface fluxes. Biases can be introduced in estimates of the mean fluxes when data are assimilated over only part of the diurnal or seasonal cycle. Thus, to interpret flux inversion results, we need to be conscious of the spatial and tem-

poral extent to which we should expect the observations to provide reliable constraints on surface fluxes. It is difficult to determine the patterns of sensitivity to surface fluxes with inversion results because the flux inversions are sensitive to the treatment of observation error covariances, a priori fluxes and flux error covariances, as well as the atmospheric transport model employed.

In this chapter, we assess the sensitivity of observing systems to CO₂ fluxes. We focus on observations from the in situ surface network, TCCON, GOSAT, and OCO-2. We use two metrics to relate the sensitivity of observations to surface fluxes for each observing system. The utility of these metrics is that they provide an indication of where we should expect observing systems to be able to constrain surface fluxes. The advantage of using a sensitivity metric over flux inversion experiments is that the sensitivity metric is independent of the choice of a priori fluxes and uncertainties, which strongly impact flux inversions. To better understand how the spatial and temporal distribution of observations will impact flux inversions, we also perform a set of simplified OSSEs to relate the sensitivity metrics to a posteriori fluxes.

This chapter is organized as follows. In Sec. 3.2, we describe the pseudo-measurements used for each observing system. We also discuss differences in the spatio-temporal distribution of observations and the impact of data screening for OCO-2 and GOSAT. In Sec. 3.3, we define sensitivity metrics to quantify the sensitivity of the various observing systems to surface fluxes. We then apply these metrics to the set of observations discussed in Sec. 3.2. In Sec. 3.4, we perform a simple OSSE experiment to examine how well the sensitivity metrics relate to inversion results for surface in-situ, GOSAT, and OCO-2 data. In Sec. 3.5, we discuss our results.

3.2 Measurements

We do not use observed CO₂ abundances, but instead generate pseudo-observations from the GEOS-Chem model. However, we do use actual measurement times, locations and averaging kernels (for the remote sensing measurements) to generate our set of artificial observations. Measurement times are aggregated into the 2° × 2.5° GEOS-Chem spatial grid with 1-hour time steps. This is done by determining if there is one or more measurements in a given grid cell for a given time step, and if there is, one pseudo-observation is generated in the given grid cell by sampling the forward model.

3.2.1 Pseudo-data

Surface measurements

To generate the artificial surface measurements, we obtain surface measurement times and locations from the Obspack PROTOTYPE package (Masarie et al., 2014; Project, 2013). We generate only surface flask, in-situ, and tower pseudo-data for our calculations. In generating the pseudo-data, we only use Obspack measurements with `obs_flag=1`, which indicates that the measurement has large spatial scale representation (rather than locally influenced). Surface in situ observations are not dependent on the solar zenith angle, as the remotely sensed data are, so they could theoretically have the same observation coverage throughout the year. To avoid the influence of data gaps due to technical and logistical issues, we assume this to be the case and use the same set of measurements for every season, based on the measurements in September-October-November (SON) of 2011. This results in 9,895 measurements per season for all seasons. The model is sampled at the location and times of these measurements to generate the pseudo-dataset.

TCCON

In this chapter, 21 sites are used, which are shown in Figure 3.1 and listed in Table 2.1. The Dryden and Caltech sites are in the same GEOS-chem grid box and thus we have combined their measurements. All of the sites that were considered in this work are still operational, except for the site at Manaus, which has been relocated.

Ideally, we would like to examine the sensitivity of an ideal TCCON network that performed to its capacity. In reality, weather and instrument issues often result in significant data gaps. Therefore, to maximize observational coverage for each station, we select measurements from the year that had the most measurements for each season at each station. This results in 6,315 measurements for SON, 4,650 measurements for December-January-February (DJF), 7,581 measurements for March-April-May (MAM), and 7,892 measurements for June-July-August (JJA).

GOSAT

We downloaded version v3.4 of the NASA ACOS GOSAT lite files from the CO₂ Virtual Science Data Environment (<https://co2.jpl.nasa.gov/#mission=ACOS>). We selected all measurements from the TANSO-FTS shortwave infrared channel, including ocean glint, high gain (H-Gain) nadir and medium gain (M-Gain) nadir, which pass the quality flag requirement. We used measurements from September 2011 to August 2012. This results

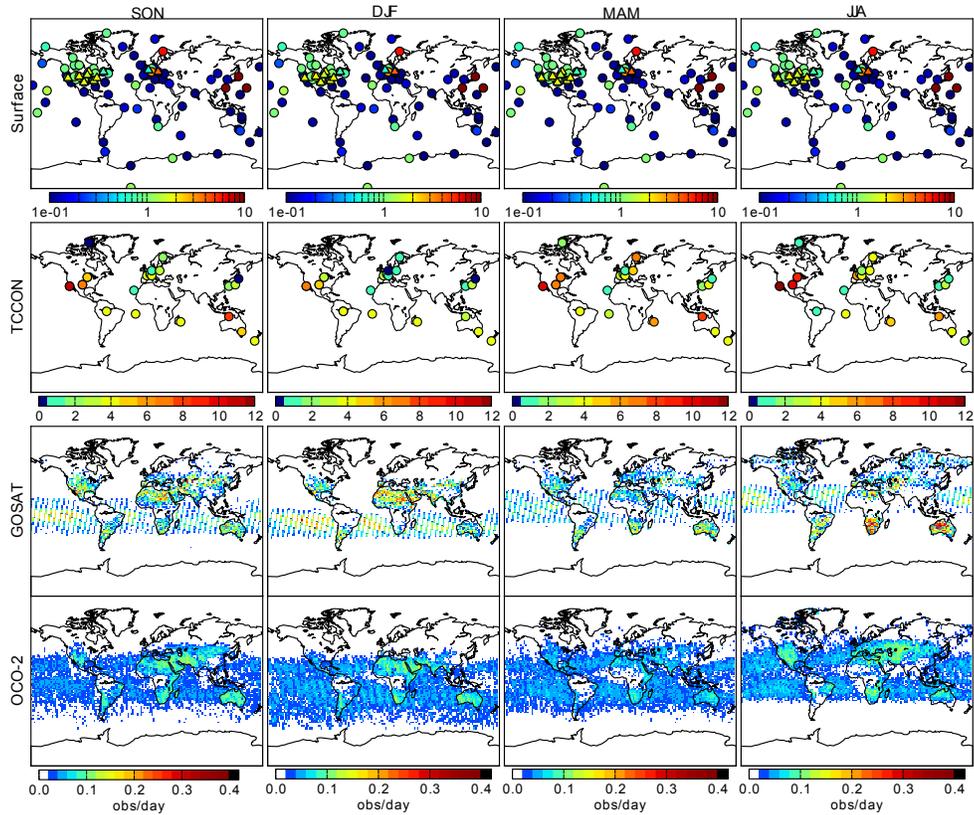


Figure 3.1: Number of measurements per day for each observing system. Each subplot shows the number density of measurements for a given observing system and season. Each row, from top to bottom, is for a different observing system: surface, TCCON, GOSAT, and OCO-2 measurements. Each column, from left to right, is for a different meteorological season: SON, DJF, MAM, JJA.

in 14,253 measurements for SON, 13,316 measurements for DJF, 11,005 measurements for MAM, and 12,813 measurements for JJA. As with the surface measurements, we sampled the model at the GOSAT measurement locations and times to generate the pseudo-data. For glint observations, we sampled the model at the glint spot. However, for GOSAT (and TCCON and OCO-2) we must also account for the vertical sensitivity of the measurements, as described in Section 3.1.

OCO-2

We used all ACOS OCO-2 measurements that pass the quality criteria and have a WL less than or equal to 10. We choose this WL range because this is the range used by the OCO-2 team in selecting the data used in computing the OCO-2 bias corrections (note that the bias corrections were then applied to all the data, including to data with warn levels greater than 10). We generate pseudo-observations from September 2014 to August 2015, which results in 20,095 measurements for SON, 22,758 measurements for DJF, 21,257 measurements for MAM, and 23,488 measurements for JJA. It should be noted that this is more restrictive data selection than for GOSAT, for which there is no WL screening available in version v3.4. Note, warn levels are provided with the v3.5 release of the GOSAT ACOS product.

3.2.2 Discussion of observational coverage

Figure 3.1 shows the number of measurements per day for each observing system. Surface sites are located over much of the globe. However, sites are much more densely concentrated in the Northern Hemisphere, particularly in North America and Europe. Furthermore, the sites in these regions generally sample much more frequently. Many of the more remote sites make less than one measurement a week, whereas the majority of North American sites make measurements at least daily.

TCCON sites are concentrated in North America, Europe, and East Asia. High latitude sites show clear seasonal differences in sampling, with many measurements in the summer and fewer in the winter. Because we are generating pseudo-data for the year with the most measurements for each site, the older sites tend to have more measurements in comparison to newer sites as they are more likely to have a year where the instrument was functioning well.

GOSAT measurements cover a much larger fraction of Earth than the surface and TCCON measurements. However, there is strong seasonality in their spatial coverage, particularly in the northern extratropics where there are many more measurements in the

Northern Hemisphere summer. Cloudiness in the tropics results in few measurements over tropical South America, tropical Africa, and southeast Asia. Regions in the tropics also show large seasonal differences in the number of measurements due to seasonal differences in cloud cover, driven by the meridional migration of the intertropical convergence zone (ITCZ). In addition to cloudiness, measurements are screened for high aerosol conditions, which reduced the number of observations over certain regions, particularly over the southern Sahara and India.

Figure 3.2 shows the spatio-temporal distribution of GOSAT measurements for different observing modes. The viewing mode is set to nadir over land and glint over the ocean. GOSAT nadir measurements are split into H-Gain, which occurs over most of the globe, and M-Gain, which are over bright surfaces. M-Gain measurements are most prominent over deserts, particularly North Africa during the Northern Hemisphere fall and winter. GOSAT glint measurements are only over water and are constrained to within 20° of the declination of the sun, thus there is a relatively narrow band of measurements over the oceans which migrates with the seasons. Due to difficulty in correcting for biases, ocean glint and M-gain nadir measurements were often excluded from flux inversions. However, it can be seen that these measurements provide a large fraction of the total measurements. Recently, Deng et al. (2016) have found that improved retrieval algorithms (ACOS v3.4) have corrected glint biases to manageable levels. Comparing flux estimates obtained with and without ocean glint measurements, Deng et al. (2016) found that the glint measurements were helpful for quantifying regional tropical and sub-tropical fluxes.

OCO-2 measurements share many of the same spatial and temporal features as GOSAT measurements. Extratropical measurements vary seasonally and cloudiness adversely impacts measurements in the tropics. OCO-2 nadir measurements generally have similar spatial coverage to GOSAT nadir observations, but are fewer in number when aggregated at $2^\circ \times 2.5^\circ$ scales (because OCO-2 land observations are performed in both nadir and glint modes and because the swath width is thin (see Section 2.2.2)). OCO-2's glint mode can make measurements over a much larger latitude range than GOSAT and can make measurements over land (Figure 3.2d). Therefore, glint measurements provide a much larger percentage of total measurements for OCO-2 in comparison to GOSAT. The temporal partitioning between nadir and glint mode has changed several times since OCO-2 was launched. Initially, the viewing mode alternated between glint and nadir measurements on consecutive 16-day ground-track repeat cycles, so that the entire sunlit hemisphere is sampled in both modes at 32-day intervals. However, beginning on 2 July 2015, the viewing mode has alternated every orbit between one orbit of nadir followed by an orbit in glint, except for orbits that are mostly over the ocean which are always

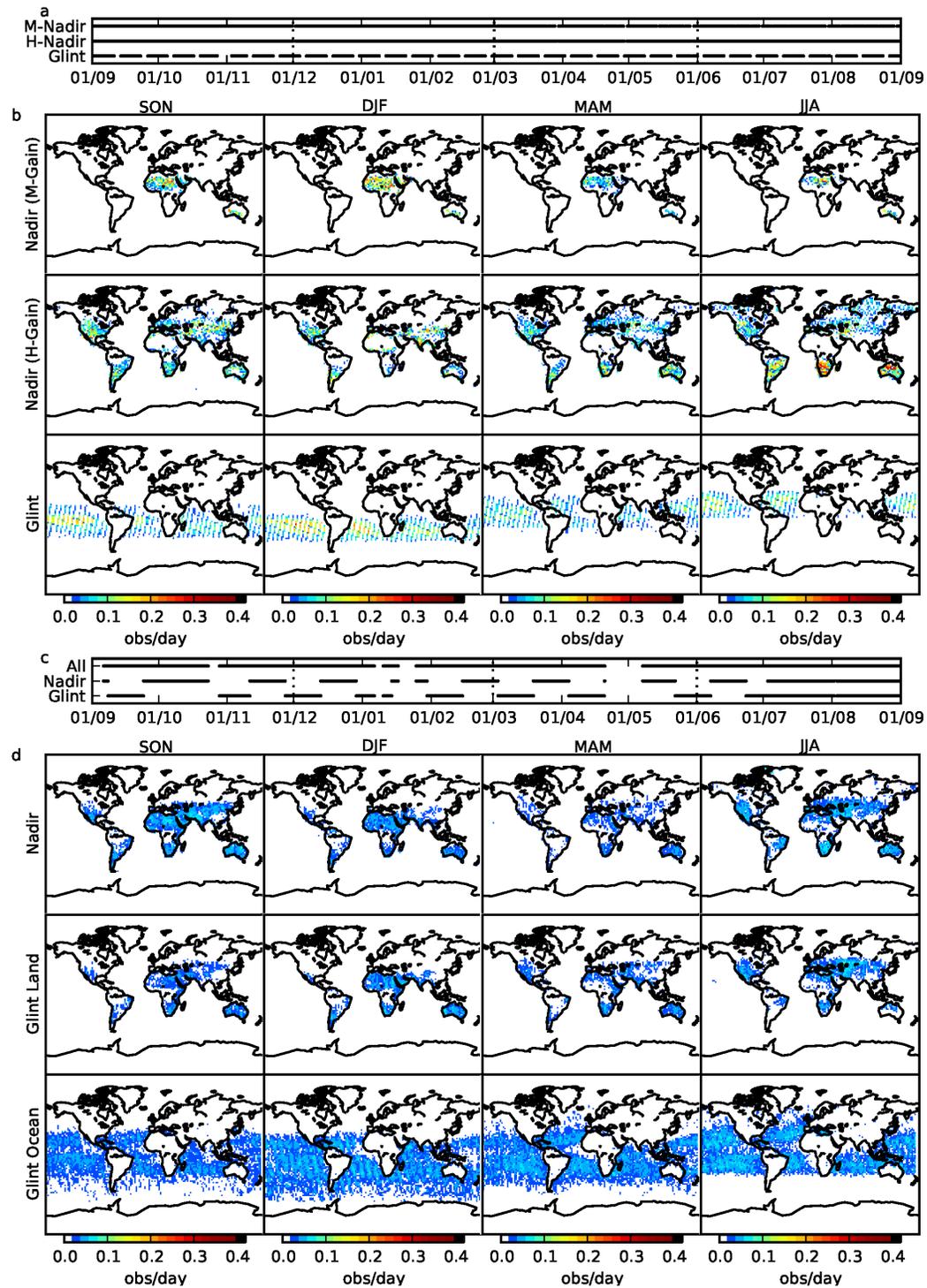


Figure 3.2: GOSAT and OCO-2 measurements for each viewing mode. The time series for each viewing mode are shown for (a) GOSAT and (c) OCO-2. The spatial distribution of measurement number density for each viewing mode and season are also shown for (b) GOSAT and (d) OCO-2.

sampled in glint mode. There are several notable gaps in measurements, two shorter gaps between January 7th-9th and 18th-24th, and a longer gap between April 20th and May 7th. It should also be noted that a significant number of measurements were removed due to interference by cosmic rays over the South Atlantic Anomaly, which is particularly important over the Amazon. A new version of the OCO-2 ACOS algorithm will significantly increase the number of measurements in this region by accounting for cosmic rays in the data screening (C. O'Dell, personal communication, 2016).

OCO-2 warn level cut-off

Low WL soundings generally occur over bright desert regions with low cloudiness on land and over the tropical ocean where there is favorable glint-viewing geometry with lower airmass and high signal. WLs generally increase in regions with persistent cloudiness or near the poorly lit poles. Above WL 12, ACOS retrieval errors well in excess of the retrieval a posteriori errors should be expected, and it is recommended by the OCO-2 team that data should not be used above WL 15, which occur in scenes not well modeled by the retrieval code (i.e., complex reflectance, high aerosols, etc.). Note that a good quality flag requires that the WL is ≤ 15 for OCO-2 v7 data. WLs are calculated separately for each measurement mode, as different retrieval parameters are used to develop WL filters for measurements over land and measurements over ocean.

Here we look at the number of OCO-2 measurements which pass the quality flag requirement that fall into several WL ranges for each meteorological season. Since a data user can filter the data based on the warn levels, we examine here the spatial and temporal distribution of different warn level cut-offs to show how the choice of warn level influences the sensitivity of the data to surface fluxes. As shown in Figure 3.3, we find that ocean glint measurements generally have low WLs, whereas land measurements are much more evenly spread over WLs. Measurements over desert regions (North Africa, Middle East, Kalahari Desert, and Australia) generally have WLs less than or equal to 8, whereas measurements throughout the tropics and sub-tropics generally have WLs less than or equal to 10. High-latitude measurements occur almost exclusively with WLs greater than 10, and thus are not included in our sensitivity calculations. These high WL measurements are particularly prevalent over boreal regions in JJA. It should be noted that data with WL greater than 10 may still be useful in flux inversions.

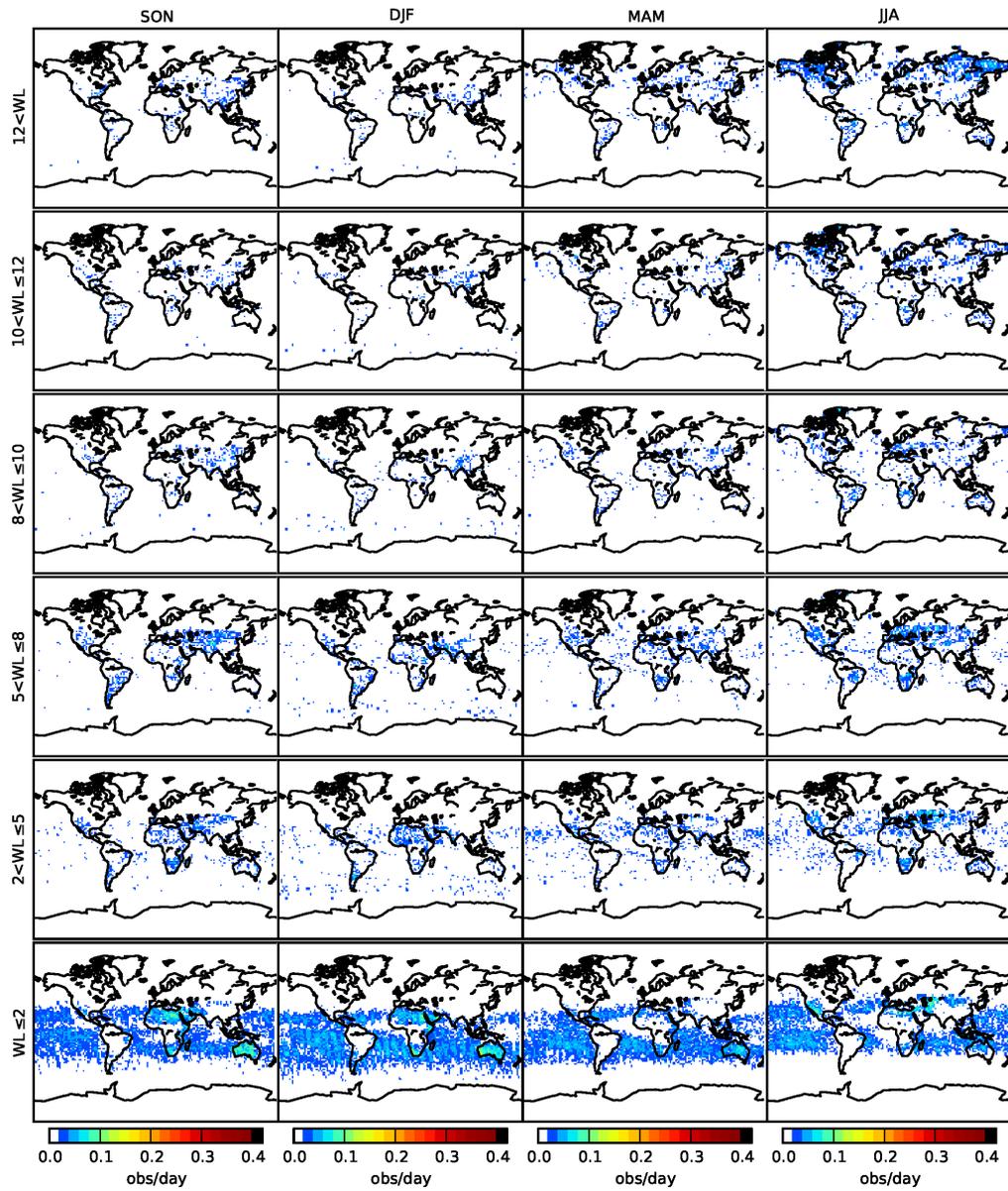


Figure 3.3: Number density of OCO-2 measurements for various WL ranges over four seasons. WLs decrease from the top to bottom.

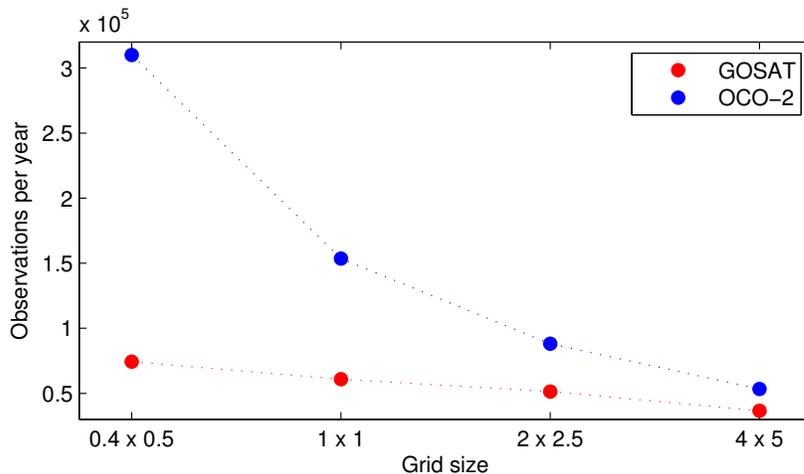


Figure 3.4: Number of GOSAT (red) and OCO-2 (blue) measurements per year after aggregation as a function of model grid size (degrees).

Grid size

It should be noted that the number and distribution of GOSAT and OCO-2 measurements used in the analysis are influenced by the resolution of the model to which they are aggregated. The larger the model grid cell, the more measurements that will occur in the same grid cell and will be aggregated to a single measurement. GOSAT measurements are separated by 263 km cross-track and 283 km along-track, which is on the order of $2^\circ \times 2.5^\circ$, whereas OCO-2 has much smaller separation between measurements, with roughly 400 soundings per degree latitude with a narrow swath (Crisp et al., 2012). Figure 3.4 shows the resulting number of GOSAT and OCO-2 measurements after aggregating the raw data to several different grid sizes. At smaller grid sizes, the number of gridded OCO-2 measurements is greater than those from GOSAT, which will inflate the OCO-2 sensitivities relative to the GOSAT sensitivities. The number of gridded measurements decreases with increased grid size much more rapidly for OCO-2 than for GOSAT because of the greater density of the OCO-2 measurements. It should be noted that, in flux inversions, measurement uncertainties are reduced with aggregation. However, because measurement uncertainties contain both systematic and random components, the reduction in measurement uncertainty can be quite small. The reduction in error with aggregation can be calculated using the expression $error^2 = a^2 + b^2/n$, where a represents systematic errors that do not decrease with averaging, b represents random errors that decrease with averaging, and n represents the number of satellite observations that are averaged (Kulawik et al., 2016). Kulawik et al. (2016) give $a = 0.8$ ppm and $b = 1.6$ ppm as mean Northern Hemisphere geometric (co-located) values for GOSAT. Over the range

of grid sizes shown in Figure 3.4, errors for OCO-2 remain relatively constant because $n > 50$ even at $0.4^\circ \times 0.5^\circ$ resolution, thus, the error on aggregated OCO-2 observations is approximately equal to a . For GOSAT, n ranges from about 2 (at $0.4^\circ \times 0.5^\circ$) to 4 (at $4^\circ \times 5^\circ$) so the random error component decreases significantly with increasing grid cell size.

3.3 Sensitivity experiments

We define two metrics, referred to as the “sensitivity” and “contribution” metrics, to quantify how the spatial and temporal distribution of the measurements can impact the sensitivity of the measurements to surface fluxes. We apply these metrics to the observing systems described above to examine how their sensitivity varies spatially and temporally.

3.3.1 Sensitivity metrics

The sensitivity of an air parcel’s CO₂ concentration to surface fluxes is a function of the history of that air parcel’s contact with the surface. We calculated this using the GEOS-Chem adjoint model (Henze et al., 2007), which calculates the derivative of the modeled CO₂ concentration with respect to a set of model parameters, f (such as the CO₂ surface fluxes). We define a response function J as a set of measurements with units of parts per million by volume (ppm) and use the adjoint model to calculate the gradient $\nabla_f J$. For surface measurements, J is defined as the sum of the CO₂ concentrations in the grid cells in which all of the measurements are located over a given season. For TCCON, GOSAT, and OCO-2, the response function is defined as the sum of all of the X_{CO₂} measurements scaled by the averaging kernels over the season:

$$J = \sum_{n=1}^N \left[\sum_{k=L_1}^{L_2} \frac{C_k \cdot A_k}{\left(\sum_{k=L_1}^{L_2} M_k \right)} \right] \cdot 10^6 \quad (3.1)$$

where C_k is the molar abundance of CO₂ at level k , M_k is the molar abundance of air, A_k is the column averaging kernel, and L_1 and L_2 are the lowest and highest atmospheric layers on which the retrieval occurs. N is the number of measurements, thus the summation over n indicates that the response function covers all measurements from the first day to the last day of the given period (which is a meteorological season in this experiment). Gas abundances (C_k and M_k) are obtained by sampling a forward model simulation at the locations and times of the observations.

The sensitivity of measurements to surface fluxes is obtained by calculating the gradient of J with respect to surface CO₂ fluxes, $f_{i,j,t}$, at longitude i , latitude j , and time t (following the notation of Liu et al. (2015)). Thus, the adjoint sensitivity of measurements to the surface fluxes is given by

$$\gamma_{i,j,t} = \frac{\partial J}{\partial f_{i,j,t}}. \quad (3.2)$$

A cumulative sensitivity to surface fluxes, β , can be attained by integration from the measurement time t_0 to an earlier time t_{-T} :

$$\beta_{i,j} = \sum_{t=t_0}^{t=t_{-T}} \gamma_{i,j,t}. \quad (3.3)$$

We call β the ‘‘sensitivity’’ metric. It gives an indication of how strongly the measurements are influenced by surface fluxes over the integration period. Therefore, for the same flux, a high sensitivity will imply that the flux can be better constrained by the measurements than a lower sensitivity. A detailed description of the adjoint calculation can be found in Appendix A of Deng et al. (2016).

A cumulative change in observed CO₂ concentration due to surface fluxes can be calculated if the surface fluxes are known:

$$\psi_{i,j} = \sum_{t=t_0}^{t=t_{-T}} \gamma_{i,j,t} \cdot f_{i,j,t}. \quad (3.4)$$

We call ψ the ‘‘contribution’’ metric. It gives the sensitivity scaled by the surface fluxes so that the contribution from each grid box to the observed signal (as aggregated in J) can be estimated. This is useful because if the sensitivity is high but the flux is low then the contribution to measurements can be similar to the contribution if the sensitivity is low but the flux is high.

3.3.2 Experiment set-up

We use the GEOS-Chem adjoint model (Henze et al., 2007) for all of the simulations. We employ the same configuration of the model as described in Nassar et al. (2011). We use 3-hourly Carnegie Ames Stanford Approach (CASA) annually balanced biospheric fluxes (Olsen and Randerson, 2004).

We calculate β and ψ for each observing system for each meteorological season. The response function (J) is generated using C_k and M_k abundances from the forward model simulation at the measurement times for a given observing systems. Measurements are

included every day of the season and adjoint sensitivities ($\gamma_{i,j,t}$) are summed from the last day (t_0) to the first day (t_{-T}) of the season. We calculate the metrics seasonally for the following two reasons. First, seasonal variations in the solar zenith angle and cloud cover cause the distribution of spaceborne measurements to vary with the seasons (Figure 3.1). Second, the dominant patterns of surface CO₂ exchange vary over the seasons. Thus, differences in the seasonal sensitivity of measurements to fluxes can lead to annual biases as found by Liu et al. (2014).

Because of the large scales on which our analysis is conducted, using the hourly surface in situ measurements could give unrealistically high β and ψ values since it is unlikely that these data will be independent. This would provide significantly greater weight to the surface in situ data compared to the GOSAT or OCO-2 data. For this reason, we aggregate surface measurements to 6-hour time steps. This is similar to the approach used for the spatial aggregation of satellite observations, which in effect de-weights the sensitivity provided by these observations (see Sec. 3.2.2). This only reduces the number of measurements significantly at the three sites in East Asia and at the one site in Scandinavia.

3.3.3 Results

Observing system comparison

Figures 3.5 and 3.6 show the spatial distributions of the sensitivity (β) and contribution (ψ) for each season. There are significant variations in β and ψ between seasons for all observing networks. Surface measurements are concentrated in North America, Western Europe and east Asia, resulting in β and ψ having the largest magnitude in these regions. The prevailing westerlies result in a relatively high sensitivity zonally to surface fluxes over the northern mid-latitudes. However, outside of this region β and ψ drop off rapidly, and the surface measurements have little sensitivity in the tropics and southern hemisphere. In contrast to all other observing systems, β values for the surface measurements are highest over North America and Western Europe in DJF and lowest in JJA, which is caused by seasonal differences in transport (see Section 3.3.3). All observing systems have the largest global magnitude of ψ value in JJA, which is driven by the strong biospheric uptake of CO₂ during the Northern Hemisphere growing season.

TCCON measurements are concentrated in North America, Europe, and Japan. However, TCCON sites make column measurements rather than boundary layer measurements, thus the spatial distribution of β and ψ for the TCCON measurements reflect a greater influence of the large-scale, free tropospheric transport (Keppel-Aleks et al.,

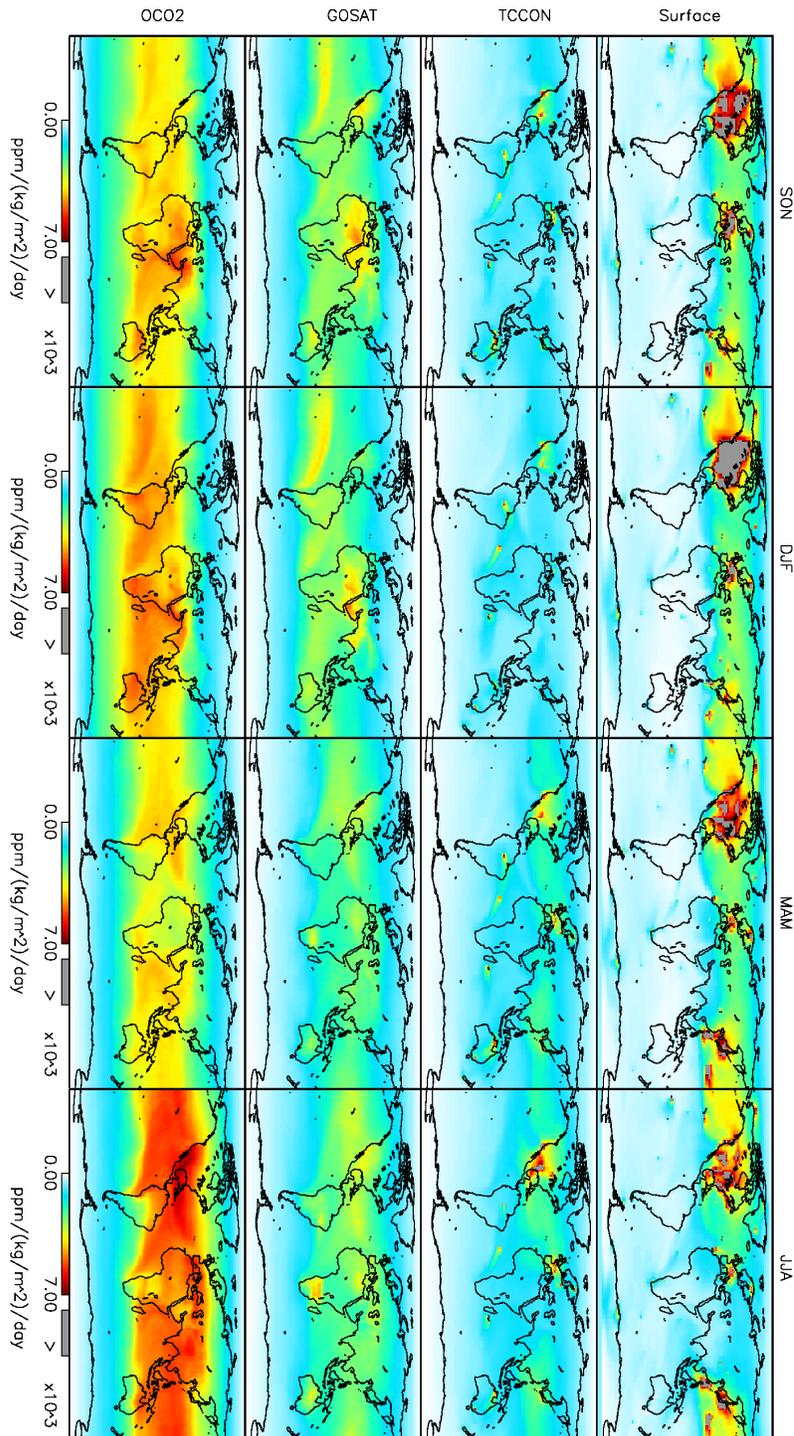


Figure 3.5: Spatial distribution of the sensitivity metric (β) for observing system (rows) and each season (columns).

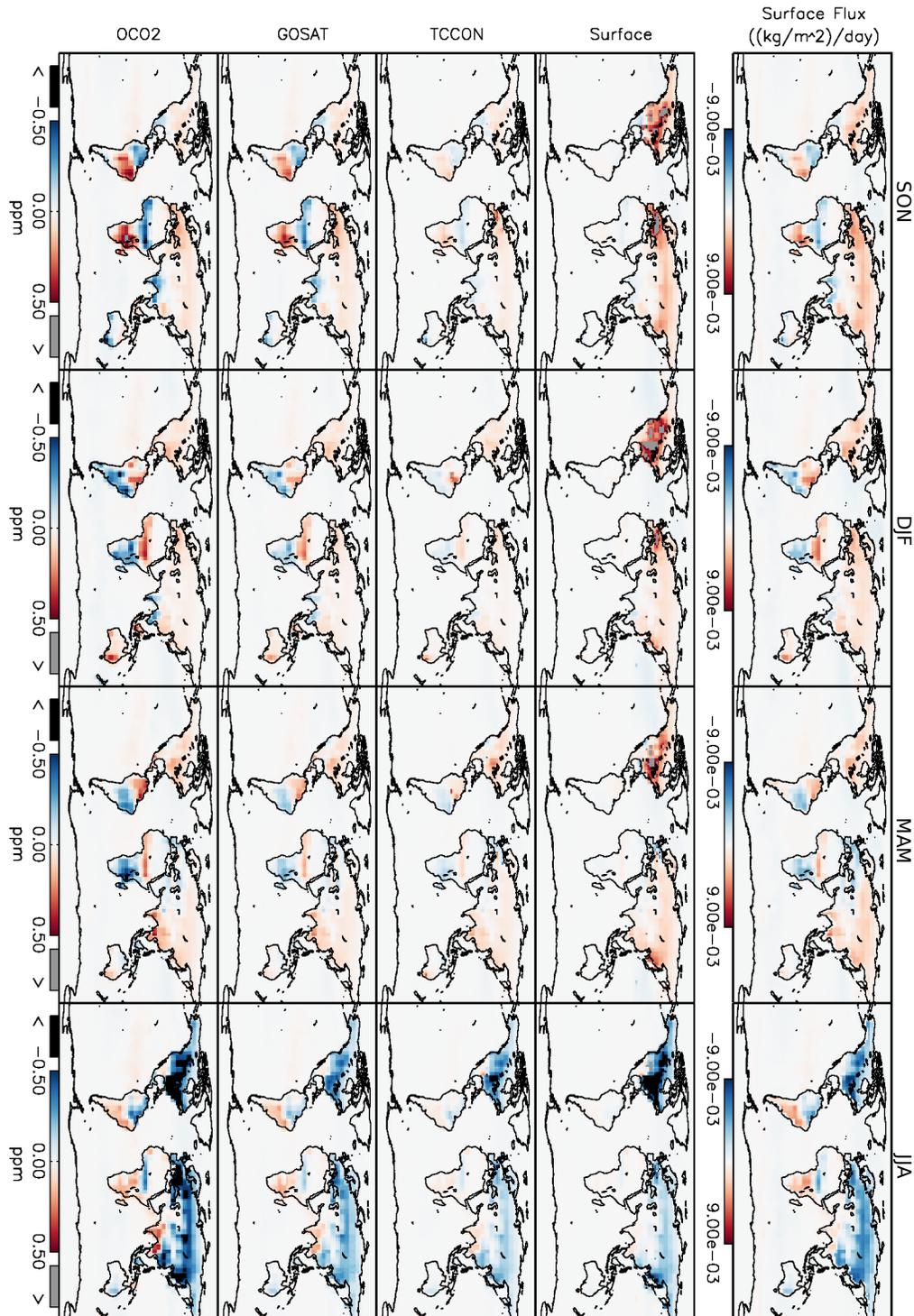


Figure 3.6: Spatial distribution of (top) seasonal surface fluxes and (bottom) the contribution metric (ψ) for each observing system (rows) and each season (columns).

2011) than for the surface measurements. The sensitivity is highest in JJA when there are the most daylight hours in the Northern Hemisphere and thus the most measurements. Significantly, we find that a single site in Manaus provides very high β and ψ locally. This is notable because there are few satellite observations over this region (primarily due to cloud cover). Thus, TCCON sites in tropical regions could complement satellite observations in constraining regional fluxes in the tropics. Alternatively, tower in situ instruments could be deployed.

GOSAT measurements have high β values in the tropics and sub-tropics throughout the year, with lower values at higher latitudes that vary seasonally. The GOSAT β and ψ are more smoothly varying spatially than those from the ground-based measurements; however, β has significant seasonal variability. This will result in large seasonal differences in the ability of GOSAT data to constrain regional fluxes, particularly in the extratropics (Liu et al., 2014). It is notable that GOSAT (and OCO-2) have high β and ψ over the Amazon and tropical Africa, where there are limited measurements. The sensitivity to these regions comes largely from downwind ocean glint measurements. Figure 3.7 shows β for various sub-sets of measurements for GOSAT. H-Gain nadir and ocean glint measurements have much higher β than M-Gain. In the northern extratropics, H-Gain measurements dominate β throughout the year. M-Gain nadir has the highest β in SON and DJF, particularly over North Africa. In the tropics, ocean glint is most sensitive to the oceans throughout the year. Ocean glint also has the highest β over tropical land in SON and DJF; however, in MAM and JJA H-gain nadir land data has the highest β values. Ocean glint measurements show dramatic temporal variations in β over the southern extratropics due to the changing spatial distribution of observations, with high β in SON and DJF and low β in MAM and JJA. It should be noted that glint and nadir observations generally have different random and systematic errors and may need to be weighted differently in inversion experiments.

OCO-2 measurements have similar spatial and seasonal distributions of β and ψ as GOSAT but with larger mean values because of the larger number of measurements (Fig 3.7). OCO-2 also shows a much larger increase in sensitivity in JJA relative to the other seasons in comparison to GOSAT. The reason for this is unclear, it may partially be due to the fact that OCO-2's sampling strategy was changed on July 2, 2015, increasing the number of measurements. Ocean glint measurements dominate β throughout the year, especially over the Tropics and Southern Hemisphere (Figure 3.7). Land nadir and land glint sensitivities have similar spatio-temporal distributions, with land nadir sensitivities being slightly larger. Land measurements provide the highest relative sensitivities in comparison to ocean glint to fluxes in North America and Eurasia.

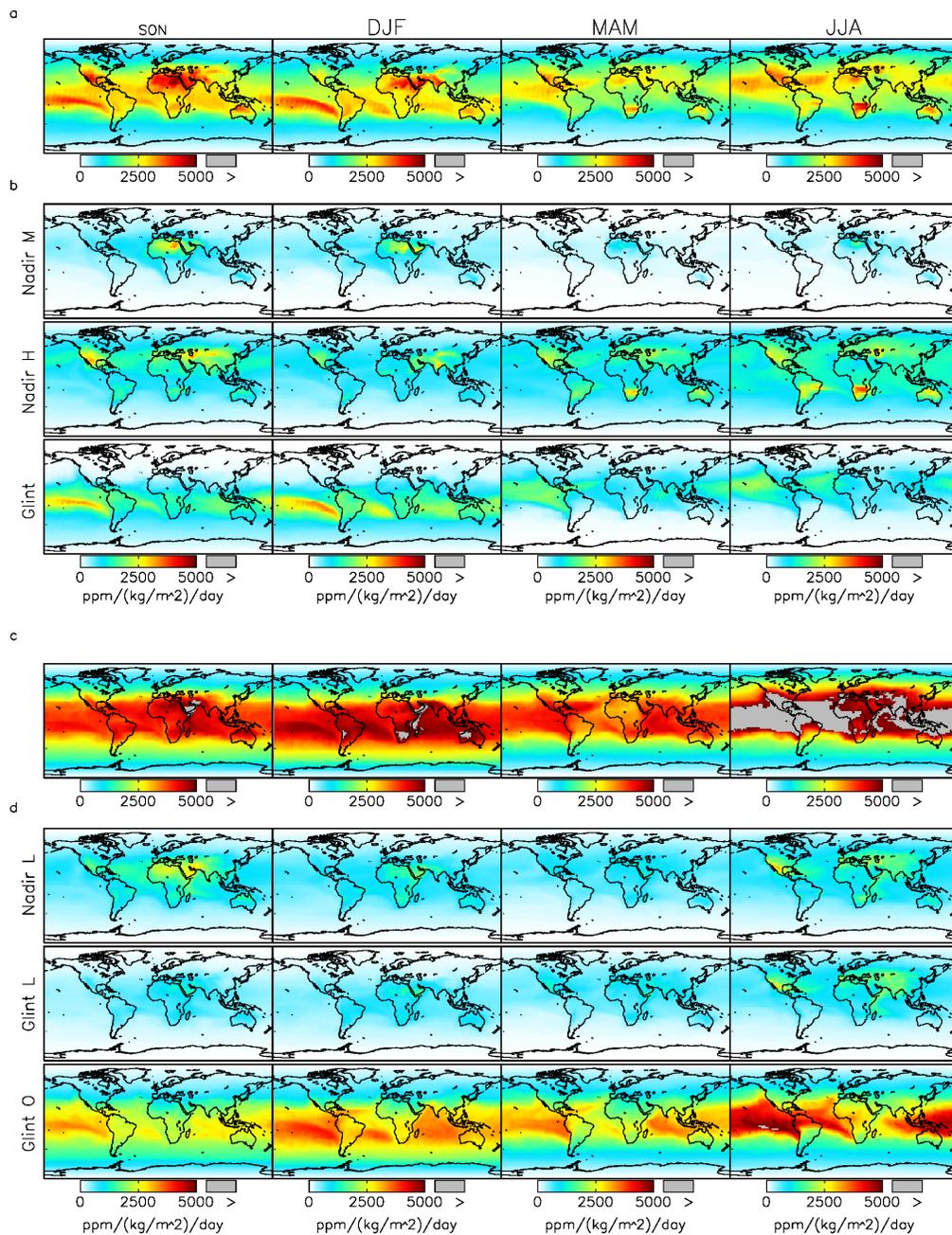


Figure 3.7: Sensitivity metric for OCO-2 and GOSAT viewing modes. Sensitivity of (a) all GOSAT viewing modes, and of (b) M-Gain nadir, H-Gain nadir, and glint measurements separately. Sensitivity of (c) all OCO-2 measurements, and of (d) land nadir, land glint, and ocean glint separately.

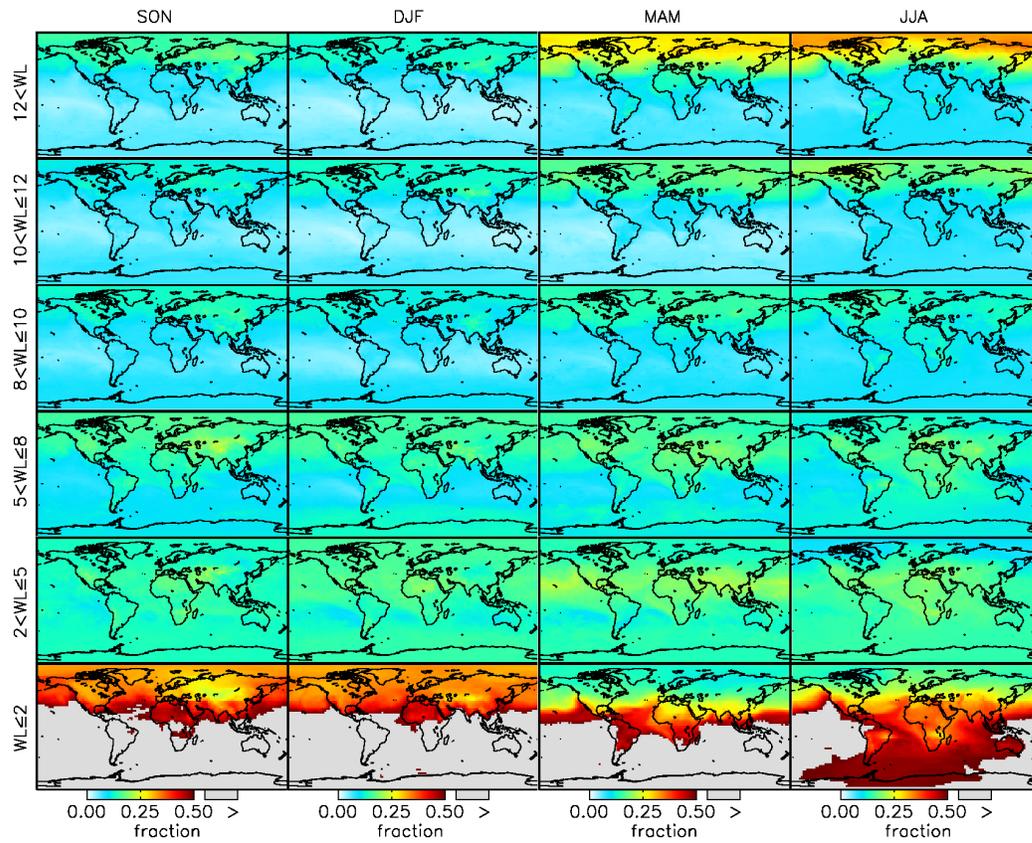


Figure 3.8: Spatial distribution of the relative contribution to the sensitivity metric from OCO-2 observations in various WL ranges for each season (columns).

Figure 3.8 shows the relative importance of various OCO-2 WL ranges to the cumulative β distribution for each season. WLs less than or equal to two provide the largest fractional contribution to β of all the WL ranges examined here, particularly for the tropics and Southern Hemisphere. WLs above 8 are important for β in the northern extratropics. There are pronounced seasonal differences in the fraction of β due to different WL ranges. For SON and DJF, WLs less than or equal to 10 provide the majority of β over the entire planet. In contrast, for MAM and JJA WLs above 10 are important for β in the northern extratropics, particularly in the Arctic.

Transport

Seasonal differences in the spatial structure of β are due to the combined impact of differences in observational coverage and seasonal transport. Here we examine how seasonal differences in transport impact β for each observing system. To isolate the impact of transport on β , we calculate β for each season using the SON observational coverage,

such that every season has identical measurements. Therefore, differences in β between the seasons are solely due to transport. Figure 3.9 shows the differences in sensitivity between each season and SON, calculated using the seasonally varying observational coverage and the fixed SON observational coverage.

For the surface measurements, we are already using the SON observational coverage for each season, thus all of the differences in β between seasons are due to transport. Seasonal differences in transport patterns have a much larger impact on β for surface measurements than for other observing systems. This is likely due to the fact that surface measurements are sensitive to vertical transport within the planetary boundary layer. β is highest near measurement sites during DJF, particularly over North America and Europe, and lowest near measurement sites in JJA. This is because there is strong surface heating in the Northern Hemisphere in the summer, which results in a deeper planetary boundary layer and more rapid vertical mixing of air parcels and, consequently, less sensitivity near the surface measurement sites. In contrast, in SON and DJF when there is less heating, and thus weaker vertical mixing, air parcels remain near the sites for longer periods, resulting in greater sensitivities to the fluxes.

For TCCON, GOSAT and OCO-2, differences in β due to changing observational coverage are significantly larger than differences due to transport, as expected. However, there are still some significant differences in β due to differences in transport. For TCCON, differences in β are most evident near TCCON sites, i.e., near Manaus for JJA-SON. For GOSAT and OCO-2, differences in β due to transport have similar spatio-temporal structure. These differences appear to be influenced by large scale atmospheric transport patterns and have large regional structures. For example, in DJF, there is enhanced sensitivity just south of the equator, in southern Africa and over the Indonesian region, whereas in JJA the sensitivity is greater in the northern tropics. The peak in sensitivity over southern Africa in DJF could be explained by strong ascent of air parcels near the surface due to the ITCZ, which are then transported to regions with enhanced observational coverage. Over South Asia, there is increased sensitivity in JJA and decreased sensitivity in DJF. This is consistent with the Asian monsoon in JJA, which drives ascent over South Asia and enhances the export of air parcels from the surface to nearby regions with more dense observational coverage. Seasonal differences in transport have a similar impact on ψ (not shown) as β . The absolute magnitude of ψ is enhanced (or diminished) with the same spatial patterns as β .

It should be noted that in flux inversion studies, the transport model has an impact on the spatial and temporal structure of a posteriori fluxes (Baker et al., 2006a). In particular, biases in the transport model can adversely influence the flux estimates. In

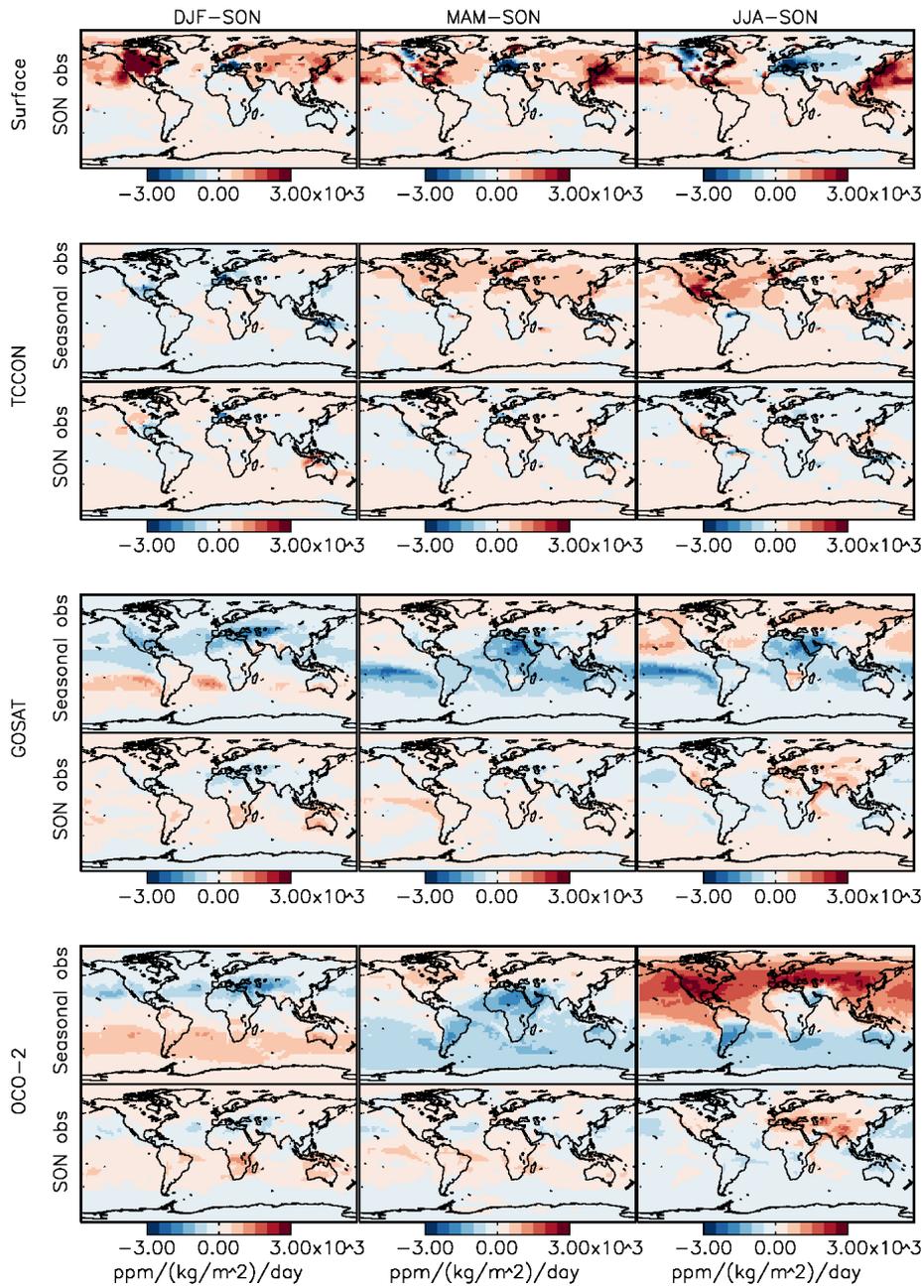


Figure 3.9: Seasonal differences (relative to SON) in sensitivity due to observational coverage and transport. For the surface data (top row), the observational coverage for SON is fixed over the seasonal cycle, so the seasonal differences are due only to transport. For TCCON, GOSAT, and OCO-2 data, the top panels show differences due to the changing observational coverage and to transport, whereas the bottom panels show differences due to transport (with fixed SON observational coverage).

our analysis we do not attempt to characterize the biases, therefore it is unclear how possible transport biases in GEOS-Chem might impact the sensitivity results. Nevertheless, given that the model is driven by assimilated meteorology, we expect that the large-scale sensitivity patterns found are likely to be robust.

3.4 Flux inversion OSSEs

To contextualize how β and ψ relate to flux inversions, we perform an idealized set of OSSEs with the same spatio-temporal distribution of measurements as before. The key benefit of conducting the OSSE is that we can define a set of "true" fluxes and use GEOS-Chem to generate pseudo-data based on these true fluxes. This will enable us to examine how differences in the spatio-temporal distribution of the pseudo-data impact our ability to recover the true fluxes, and in particular, how they can result in systematic differences in flux estimates between different observing systems. This experiment is highly idealized and is only intended to give insights into how β and ψ relate to a posteriori fluxes. We do not intend to extrapolate any quantitative conclusions about regional or annual biases; in a more realistic inversion these would be based on the relative weighting of measurements and the spatially and temporally heterogeneous mismatch between true and a priori fluxes.

3.4.1 Experiment set-up

As before, we use the same forward model configuration as Nassar et al. (2011). To obtain biased a priori fluxes for the OSSE, which are different from those used to generate the pseudo-data, we scale the CASA land fluxes by 0.6 so that the a priori and true seasonal cycles have the same phase, spatial distribution and balanced annual net flux, and leave ocean fluxes unchanged. The biospheric CO₂ fluxes are balanced in the standard configuration of GEOS-Chem and we leave that unchanged in our OSSE. We use the same error statistics as Deng et al. (2014), applying 16% error to fossil fuels, 38% error to biomass burning, 22% error to ocean fluxes, and 44% error to terrestrial ecosystems. To simplify the interpretation of the retrieved fluxes, we assume a uniform measurement error of 0.7 ppm for surface, TCCON and OCO-2 data, and assume diagonal covariance matrices. In a more realistic inversion, different uncertainties would be applied to the different observing systems. Because many more OCO-2 observations are aggregated to $2^\circ \times 2.5^\circ$ grid cells than GOSAT observations, we should expect random errors for OCO-2 to be much smaller than for GOSAT (since the random error scales with $1/\sqrt{n}$).

Thus, for GOSAT we assume larger measurement errors of 1.4 ppm, consistent with error estimates from Kulawik et al. (2016). To optimize surface fluxes, we minimize the 4D-Var data assimilation cost function as described in Deng et al. (2014) to retrieve scaling factors for the a priori fluxes in each grid cell.

To be consistent with the sensitivity calculations, we assimilate all observations over a given meteorological season to optimize flux scaling factors over the same period. A limitation of this short assimilation window is that there will be about three months of observations to constrain the fluxes in the earliest part of the season but less than one month of observations toward the end of the season. A short assimilation window will also reduce the impact of distant flux regions, as the window can be shorter than the transport timescales to these regions. This will act to enhance biases introduced by inhomogeneous spatial observational coverage, as optimized source fluxes will not benefit from observations later in the seasonal cycle. Using a longer assimilation window, following Deng et al. (2014), for example, will mitigate some of these biases. Thus, using a three-month assimilation window, as is done here, will give us the maximum impact that the seasonally-varying coverage could have on the flux estimates.

3.4.2 Inversion OSSE results

The OSSE experiments show a large degree of consistency with β and ψ results shown earlier. As shown in Figure 3.10, we find that surface measurements give consistently strong constraints on northern extratropical fluxes throughout the year. OCO-2 best provides consistently strong constraints on fluxes in the tropics and Southern Hemisphere throughout the year, but seasonally dependent constraints on the northern extratropics. It should be emphasized that the magnitude of the flux correction is dependent on the prior and observation error covariances which are not given realistic values in this experiment, thus, it is not possible to make conclusions about the magnitude of flux correction for the real world. Differences between the a posteriori and a priori fluxes (Figure 3.10c) generally have the same spatial characteristics as ψ for each observing system (see Figure 3.6). A notable difference between ψ and a posteriori fluxes for OCO-2 and GOSAT is that the a posteriori fluxes in the northern extratropics are as well constrained as the tropical fluxes even though ψ is generally larger in the tropics and Southern Hemisphere. This may be due to the fact that terrestrial surface fluxes in the northern extratropics are generally spatially coherent over any given season, whereas fluxes in the tropics and Southern Hemisphere are more heterogeneous, with some regions taking up carbon and other regions emitting carbon. Another possibility is that the signal is larger in the

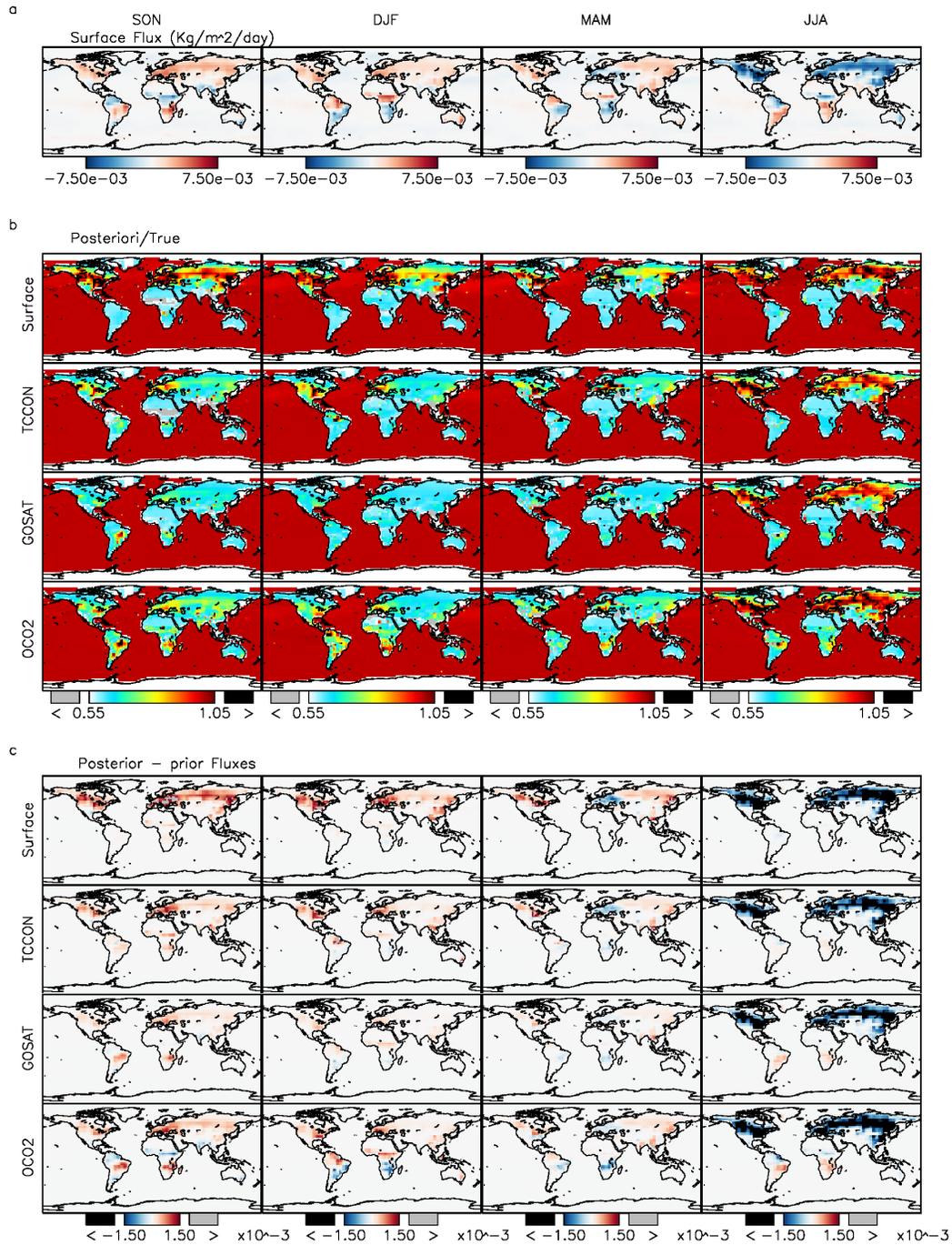


Figure 3.10: Results of idealized OSSEs. (a) Spatial distribution of the true surface fluxes (in kg CO₂ m⁻² s⁻¹) for each season. (b) The ratio of a posteriori to true fluxes for each observing system and season. (c) Difference between a posteriori and a priori surface fluxes for each observing system (in kg CO₂ m⁻² s⁻¹).

northern extra tropics because there ratio of land to water is higher.

Figure 3.11 shows the regional differences in the flux estimates, relative to the true fluxes, inferred for each observing system. To help interpret the results, the differences have been aggregated to the 11 TransCom land regions (Gurney et al., 2002). In this OSSE, surface measurements provide strong constraints on North American and European fluxes throughout the year. TCCON, GOSAT and OCO-2 provide temporally variable constraints in North America and Europe which result in significant biases in annual net fluxes. As suggested by the sensitivity experiments, much stronger estimates are available in JJA than the other seasons. OCO-2 is superior to GOSAT at retrieving fluxes in these regions due to both observational coverage and smaller measurement uncertainties. In the northern extratropics, GOSAT is unable to retrieve northern extratropical fluxes as well as OCO-2 due to the larger measurement uncertainties. This is most apparent in the northern extratropics in winter when small absolute fluxes confound GOSAT's ability to differentiate between the true and a priori fluxes. In the northern extratropical summer, when absolute fluxes are large, GOSAT performs nearly as well as OCO-2 in retrieving the true fluxes.

The net annual flux estimates are shown in Figure 3.11m. Both the a priori and true terrestrial biosphere fluxes are balanced annually, thus, nonzero net annual fluxes in the a posteriori are primarily due to uneven spatio-temporal sensitivity to the surface fluxes. The large net annual fluxes obtained here indicate that it is critical that the magnitude of the a priori seasonal cycle is realistic if annual net fluxes are to be inferred. OCO-2 and GOSAT retrieve negative net fluxes due to increased sampling of the summer hemisphere relative to the winter hemisphere. Regionally, all of the extratropical regions contribute a net negative flux for GOSAT and OCO-2.

3.4.3 Inversion OSSE discussion

The results of this study show that current observing systems have significant spatial and temporal variability in their sensitivities to surface fluxes, and that these variabilities are likely to lead to systematic biases in flux inversion studies. However, there are several major difficulties in relating the results of these OSSEs to flux inversions assimilating real observations. First, the magnitude of the flux correction between the a priori and a posteriori fluxes is dependent on the weighting of the observations by uncertainties. In this experiment, surface, TCCON, and OCO-2 observations were weighted by the same uncertainties; however, in reality different uncertainties would be applied to the different observing systems. Surface observations are particularly difficult to assimilate into trans-

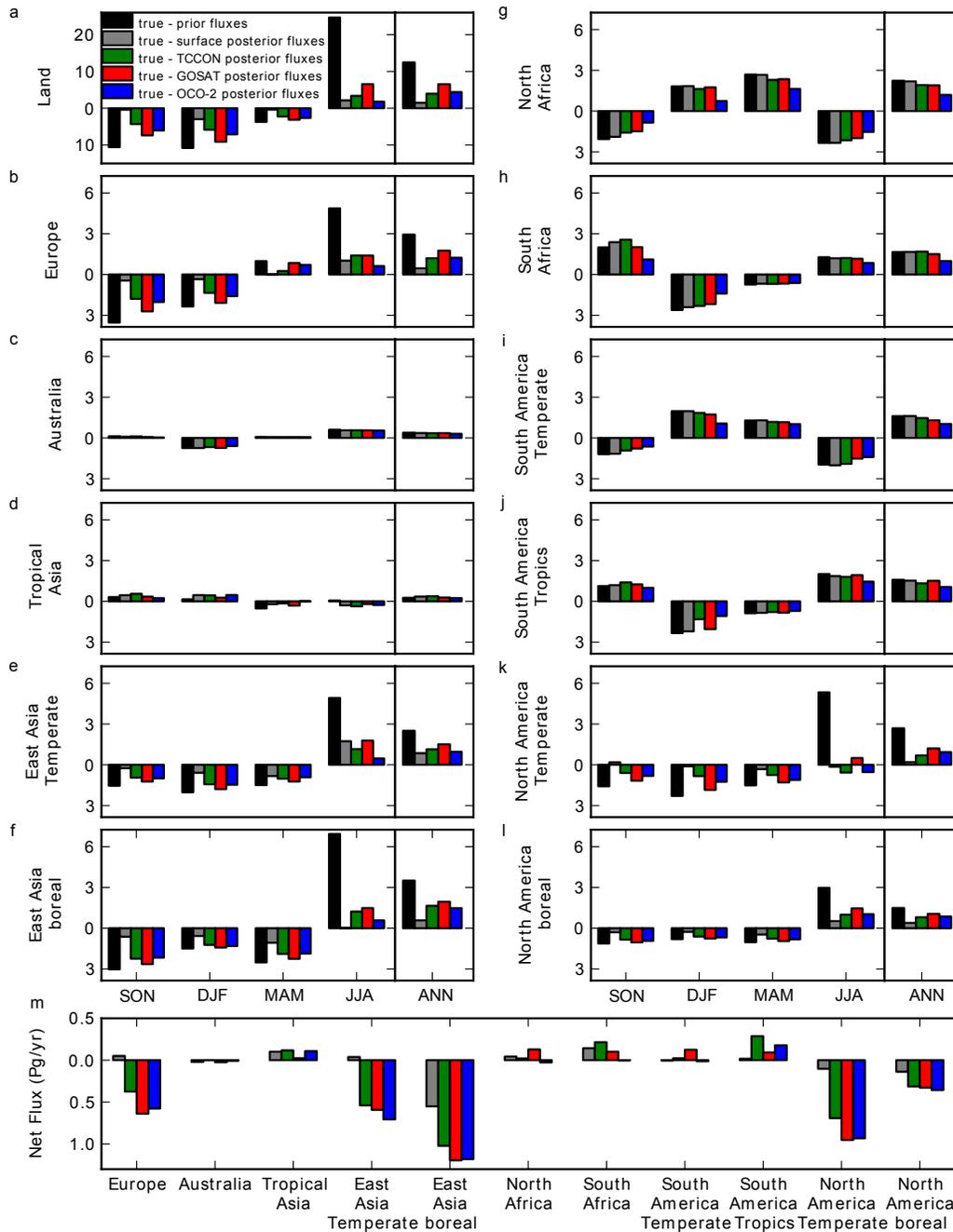


Figure 3.11: (left section of panels) Differences (in Pg/yr) between the true and a priori fluxes (black), true and surface a posteriori fluxes (gray), true and TCCON a posteriori fluxes (green), true and GOSAT a posteriori fluxes (red), and true and OCO-2 a posteriori fluxes (blue) for each season, and (right section of panels) the annual mean of the absolute seasonal differences shown in the left section for (a) all land, (b) Europe, (c) Australia, (d) tropical Asia, (e) temperate east Asia, (f) boreal east Asia, (g) north Africa, (h) south Africa, (i) temperate South America, (j) tropical South America, (k) temperate North America, and (l) boreal North America. Shown in (m) is the estimated annual net flux for each TransCom land region.

port models due to representativeness errors associated with linking the concentration observed at a point to the mean concentration of a grid cell. Quantifying these errors is quite difficult and must be done on a site-by-site basis. Generally, these errors are larger for mid-continental sites than remote marine sites (Gerbig et al., 2003). Errors applied to space-based observations are generally more spatially uniform.

Second, isolating biases due to the observational coverage in a posteriori fluxes is obscured by other major sources of systematic error in flux inversions, the most important of which are biases in the measurements and in the transport model. These additional factors obscure the origin of biases, which in most cases cannot be easily distinguished from one another. Furthermore, biases due to observational coverage will not be solely dependent on the sensitivity of the measurements, but are dependent on the innovation vector (difference between measurements and simulated measurements). Thus, a posteriori biases will depend on the choice of a priori fluxes, and the spatial and temporal differences between the a priori and true fluxes. For example, if we had chosen to increase the seasonal cycle in Section 3.4, rather than decrease it, all of the biases due to observational coverage would have had the opposite sign. Furthermore, the relative weighting of the measurements and a priori fluxes, through the error covariances, strongly affects the a posteriori fluxes. Thus any biases will be sensitive to the structure of the error covariances.

As shown by the OSSEs, inferred regional flux estimates do not agree closely when assimilating measurements from different observing systems due to spatial and temporal differences in sensitivities. As noted above, the short three-month assimilation window used in our inversion analyses enhances the biases in the seasonal flux estimates. But the results do demonstrate the potential impact of the varying observational coverage on the flux estimates. To some extent, flux inversions using GOSAT and OCO-2 should agree more closely, as biases should be similar given their similar spatio-temporal observational coverage, if the same a priori is used. This was seen in the regional net fluxes in Section 3.4. However, the differences between flux inversions using satellite and surface measurements should be much larger and more difficult to interpret because of the large differences in spatio-temporal observational coverage.

It should be easier to understand the differences in the a posteriori fluxes by looking at short periods (monthly to seasonal timescales) when sensitivities and fluxes have a significant amount of temporal and spatial coherence. In their assimilation of GOSAT data, Deng et al. (2014) found that the uncertainty reduction on the estimates of the northern high-latitude fluxes was negligible in winter and at a maximum in summer, which they attributed to the seasonal differences in the GOSAT observational coverage. In this case,

the a posteriori surface fluxes from the inversion should be in better agreement with true fluxes. For example, in Liu et al. (2014), the a posteriori monthly fluxes were in better agreement with the true fluxes, but the annual net fluxes were degraded.

3.5 Conclusions

The results of this work show that there are significant differences in the sensitivity of measurements to surface fluxes between observing systems and observing modes that vary spatially and temporally. We find that surface measurements are sensitive to surface fluxes over North America and Europe throughout the year, suggesting that these regions have temporally consistent constraints on fluxes throughout the year. However, we find that surface measurements have little sensitivity to the tropics and Southern Hemisphere (where there are few observations) and seasonally variable sensitivity in Asia, suggesting limited ability to constrain these regions. In contrast, we find that OCO-2 and GOSAT measurements have high sensitivity to tropical and sub-tropical fluxes throughout the year but have seasonally variable sensitivity in the extratropics. Our results suggest that space-based observing systems, particularly OCO-2, provide the most consistent constraints on regional surface fluxes in the tropics.

We found that, for GOSAT, the ocean glint data provide the greatest sensitivity to fluxes in the tropical and southern land regions in SON and DJF, whereas the nadir H-gain land data offer the greatest constraints on fluxes in these land regions in MAM and JJA, reflecting the seasonal latitudinal shift in the distribution of the GOSAT ocean glint data. Our results explain why combining land and ocean GOSAT data enabled Deng et al. (2016) to better disaggregate the regional fluxes in the tropics. For OCO-2 we found that the ocean glint data provided the dominant sensitivity to the surface fluxes in the northern subtropics, tropics, and Southern Hemisphere, mainly due to the greater OCO-2 data density and the greater latitudinal range of the OCO-2 ocean glint data coverage. We also found that more than 50% of the sensitivity of OCO-2 data to the fluxes in the tropics and Southern Hemisphere in SON, DJF, and MAM comes from data with WLS less than or equal to two. In the northern high latitudes, OCO-2 data with WLS greater than 10 provide the dominant sensitivity to surface fluxes in MAM and JJA. Our results suggests that it could be challenging for OCO-2 inversion analyses that omit data with WLS greater than 10 to reliably capture the uptake of CO₂ by the boreal forests in spring and summer. If OCO-2 data with WLS greater than 10 were to be included, more information might be obtained on the fluxes in the northern latitudes across more of the seasonal cycle.

Using a set of simple idealized OSSEs, we showed that differences in sensitivity to surface fluxes can result in significant differences in a posteriori fluxes between observing systems. In particular, because of their highly variable seasonal sensitivity to extratropical fluxes, space-based observing systems are susceptible to systematic biases in annual flux estimates in the extratropics. This suggests that obtaining high accuracy in a posteriori regional flux estimates will be a challenge with the current set of observing systems, and that we should not expect a posteriori regional flux estimates to be consistent between different observing systems. However, we find that surface and space-based measurements have complementary sensitivities to surface fluxes, which reinforces the idea that surface and space-based measurements can play complementary roles in quantifying regional sources and sinks of CO₂. Combining measurements from different observing systems in flux inversions may lead to a posteriori fluxes with improved accuracy. However, this will also come with a new set of challenges as these different data sets come with differences in spatial and temporal sampling as well as differences in random and systematic errors. Further study is required to find how these different observational data sets can be weighted in a combined inversion.

Chapter 4

Evaluating GPP and respiration estimates over northern mid-latitude ecosystems using solar induced fluorescence and atmospheric CO₂ measurements

This chapter is adapted from:

Byrne, B., Wunch, D., Jones, D. B. A., Strong, K., Deng, F., Baker, I., Köhler, P., Frankenberg, C., Joiner, J., Arora, V. K., Badawy, B., Harper A. B., Warneke, T., Perti, C., Kivi, R., and Roehl, C. M. (2018). Evaluating GPP and respiration estimates over northern midlatitude ecosystems using solar induced fluorescence and atmospheric CO₂ measurements. *J. Geophys. Res.-Biogeo.*, 123, 1–22, <https://doi.org/10.1029/2018JG004472>

4.1 Introduction

Observational constraints on GPP and R_e are required to identify errors within TBMs. In this study, we evaluate the seasonal cycle of GPP, R_e and NEE produced by four TBMs and FLUXCOM over northern mid-latitude ecosystems using observational constraints from atmospheric CO₂ and SIF measurements.

NEE can be constrained on large scales using atmospheric CO₂ observations. On seasonal timescales, variations of CO₂ are primarily driven by NEE fluxes. Therefore, modeled NEE fluxes can be evaluated by comparing simulated atmospheric CO₂ using

an atmospheric transport model with observed atmospheric CO₂. This method has previously been applied to evaluate TBM fluxes (Messerschmidt et al., 2013; Peng et al., 2015). Constraints on NEE can be related to GPP and R_e through Eq. 1.3. Therefore, if independent constraints on either GPP or R_e are used in combination with constraints on NEE, it would be possible to evaluate both GPP and R_e . Currently, there are no large-scale observational constraints on R_e , but recent advances in remote sensing have provided a new constraint on large-scale GPP.

Within the last few years, satellite observations of SIF have become possible (Frankenberg et al., 2011a; Joiner et al., 2011; Guanter et al., 2012). SIF is the emission of radiation by chlorophyll during photosynthesis and thus provides a proxy for GPP (Papageorgiou and Govindjee, 2007). Although challenges remain in quantifying GPP from satellite SIF observations, many studies have found linear relationships between satellite retrievals of SIF and GPP from the canopy to ecosystem scale on weekly to monthly time scales (Frankenberg et al., 2011b; Guanter et al., 2012; Yang et al., 2015; Damm et al., 2015; Zhang et al., 2016b,a; Wood et al., 2017; Sun et al., 2017). SIF is a more direct proxy for GPP than vegetation indices (Walther et al., 2016; Jeong et al., 2017; Luus et al., 2017), because other variables are required to estimate GPP from vegetation indices. For example, PAR and light-use efficiency are required to estimate GPP from NDVI (Field et al., 1995). In this study, we investigate the utility of using atmospheric CO₂ and SIF observations to evaluate fluxes of NEE, GPP and R_e . First, observations of atmospheric CO₂ and SIF are used to evaluate estimates of NEE and GPP, respectively. Then, constraints on NEE and GPP are combined to evaluate R_e estimates.

For atmospheric CO₂ observations, we use the retrieved X_{CO₂} from the TCCON (Wunch et al., 2011). Modeled X_{CO₂} is generated by simulating atmospheric CO₂ fields with the GEOS-Chem chemical transport model driven with imposed NEE as input surface fluxes. The simulated atmospheric CO₂ fields are then integrated in altitude and compared to TCCON. Using this method, the seasonal cycle of the model-NEE-based X_{CO₂} is compared with the seasonal cycle observed at several TCCON sites.

For SIF, the longest record of space-based observations is from the Global Ozone Monitoring Experiment-2 (GOME-2) instrument aboard the Meteorological Operational Satellite-A (MetOp-A), which was launched by the European Space Agency in 2006 (Joiner et al., 2013; Köhler et al., 2015). Eight years of GOME-2 SIF observations (from 2007–2014) are used to evaluate the mean seasonal behavior of GPP in the TBMs.

We evaluate the constraints that SIF and X_{CO₂} provide on four TBMs and FLUX-COM upscaled fluxes (Tramontana et al., 2016). Two of the TBMs examined here employ diagnostic phenology: the Carnegie-Ames Stanford Approach (CASA, Potter et al., 1993;

Randerson et al., 1996) and the Simple Biosphere model version 3 (SiB3, Baker et al., 2008). For these TBMs, satellite observations of vegetation indices are assimilated to prescribe phenology. These TBMs are widely used to generate prior fluxes for flux inversion analyses (Schuh et al., 2010, 2013; Gurney et al., 2004), and would likely be employed in flux inversions assimilating SIF and atmospheric CO₂. Thus, it is necessary to understand the level of a priori agreement to expect between the observations and these TBMs. We also evaluate two prognostic TBMs: the Canadian Terrestrial Ecosystem Model (CTEM, Melton and Arora, 2016), and JULES (Clark et al., 2011; Harper et al., 2018). In contrast to the diagnostic TBMs, CTEM and JULES model phenology only as a function of the driving meteorology. Prognostic TBMs are used in simulations of future climate, thus it is desirable to understand the agreement of these TBM fluxes with observational constraints. FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations of CO₂ fluxes, satellite remote sensing, and meteorological data (Tramontana et al., 2016; Jung et al., 2017). For this study, we examine upscaled fluxes generated using random forests (RF), multivariate regression splines (MARS) and artificial neural networks (ANN). FLUXCOM GPP and R_e are widely considered to be among the best estimates available, thus it is important to include these fluxes in our comparison. FLUXCOM NEE estimates are known to produce an unrealistic large annual net sink by the biosphere (18–28 Pg/yr) (Tramontana et al., 2016; Jung et al., 2017), thus we do not evaluate FLUXCOM NEE against TCCON.

After evaluating model GPP and NEE, we examine the possibility of combining GPP and NEE constraints from atmospheric CO₂ and SIF observations to evaluate model R_e . An “optimized” R_e seasonal cycle is calculated using NEE fluxes produced by two atmospheric CO₂ flux inversions and GPP fluxes produced by CASA, SiB3 and FLUXCOM, as these GPP fluxes give close agreement with the normalized seasonal cycle of SIF. We further examine the sensitivity of this estimate to uncertainties in GPP and NEE fluxes, discuss possible reasons for differences between our optimized R_e and TBM R_e , and discuss the current limitations of estimating optimized R_e with existing observational constraints. The area of study is limited to the northern extratropics (39°–65° N). These latitudinal limits were chosen because the seasonal variations in X_{CO₂} and SIF are largest over these latitudes, and thus provide the largest signal in the observations.

This paper is organized as follows. In Sec. 4.2, the data and our methods are described. In Sec. 4.3, we present the results of our experiments. We first describe the agreement between SIF and GPP, and then X_{CO₂} and NEE. Then the feasibility of evaluating R_e estimates by combining GPP and NEE constraints is examined. In Sec. 4.4, we discuss

the plausibility of our optimized R_e seasonal cycle and possible sources of error. Then the limitations of calculating optimized R_e given the observational constraints is discussed. In Sec. 4.5, we give our conclusions.

4.2 Data and methods

4.2.1 Terrestrial biosphere models

This study examines GPP, R_e and NEE from four TBMs that use a range of input parameters. The TBMs used are: CASA (Potter et al., 1993; Randerson et al., 1996), SiB3 (Baker et al., 2008), CTEM (Melton and Arora, 2016), and JULES (Clark et al., 2011; Harper et al., 2018). CASA and SiB3 assimilate satellite observations of vegetation indices to produce diagnostic phenology, while CTEM and JULES employ prognostic phenology in which the phenology is a function of the driving meteorology. We use two sets of CTEM fluxes that are driven by different meteorology to examine the impact of the driving meteorology on GPP, R_e , and NEE. CTEM-CRU is driven by NCEP-CRU (merged product of National Centers for Environmental Prediction (NCEP) reanalysis and Climate Research Unit (CRU) observations) (Wei et al., 2014) and CTEM-GEM is driven by Global Environmental Multi-scale - Modeling Air Quality and CHemistry - Greenhouse Gas (GEM-MACH-GHG) operational weather prediction model (Anselmo et al., 2010; Robichaud and Ménard, 2014; Makar et al., 2015; Polavarapu et al., 2016). Details of the TBM runs are given in Table 4.1, and additional description of the configuration of the TBMs is given in below.

Carnegie-Ames Stanford approach (CASA) model

The version of the model used here, CASA-GFED3, was modified from Potter et al. (1993) as described in Randerson et al. (1996) and van der Werf et al. (2006). It is driven by Modern Era Retrospective-analysis for Research and Applications (MERRA) reanalysis and satellite-observed NDVI to track plant phenology. We use the same fluxes as are used for the CarbonTracker 2016 (<http://carbontracker.noaa.gov>) prior. CASA outputs monthly fluxes of Net Primary Productivity (NPP) and heterotrophic respiration (R_H). From these fluxes, GPP and R_e are estimated to be $GPP = 2NPP$ and $R_e = R_H - NPP$. Temporal downscaling and smoothing was performed from monthly CASA fluxes to 90 minute fluxes using temperature and shortwave radiation from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-interim reanalysis (note this method differs from Olsen and Randerson (2004)). GFED_CMS is used for global fire

Table 4.1: Terrestrial biosphere models used in this study.

Model	Meteorology	Phenology	Respiration variables	Years
CASA	MERRA ^a	NDVI	carbon pool size (C), temperature (T) and soil moisture (M)	2007-2012
SiB3	MERRA ^b	MODIS FPAR and LAI ^e	T and M	2007-2012
CTEM-CRU	NCEP-CRU ^c	carbon-gain	C, T and M	2009-2010
CTEM-GEM	GEM-MACH-GHG ^d	carbon-gain	C, T and M	2009-2010
JULES	NCEP-CRU	temperature	C, T, and M	2005-2014

^aModern-Era Retrospective Analysis for Research and Applications (Rienecker et al., 2011)

^bPrecipitation scaled to Global Precipitation Climatology Project (GPCP: Adler et al., 2003) following Baker et al. (2010)

^cWei et al. (2014)

^dAnselmo et al. (2010); Robichaud and Ménard (2014); Makar et al. (2015)

^eStöckli et al. (2008)

emissions (<http://nacp-files.nacarbon.org/nacp-kawa-01/>). We use average model fluxes by averaging the fluxes for 2007-2012.

Simple biosphere model (SiB3)

SiB3 was originally designed as a lower boundary for General Circulation Models with explicit treatment of biophysical processes. The ability to ingest satellite phenology was later introduced (Sellers et al., 1996a,b), and further refinements included a prognostic canopy air space (Vidale and Stöckli, 2005), more realistic soil and snow (Baker et al., 2003) and modifications to calculations of root water uptake and soil water stress (Baker et al., 2008). The current version is called SiB3. Simulations used in this analysis use phenology (Leaf Area Index, LAI; fraction of Photosynthetically Active Radiation, fPAR) from the Moderate Resolution Imaging Spectroradiometer (MODIS). MERRA reanalysis is used as model inputs, with precipitation scaled to Global Precipitation Climatology Project (GPCP: Adler et al., 2003) following Baker et al. (2010).

Canadian terrestrial ecosystem model (CTEM)

CTEM is a dynamic vegetation model developed for inclusion in the Canadian Center for Climate Modeling and Analysis (CCCma) coupled general circulation model. Because CTEM is designed to model ecosystems under climate change, the phenology parametrization has to be independent of current climatic factors. Thus, a “carbon-gain” approach is used to determine phenology, which is based on local environmental

conditions. In this approach, leaf onset is initiated when it is beneficial for the plant, in carbon terms, to produce new leaves. Leaf offset is initiated by unfavorable environmental conditions that incur carbon losses and these include shorter day length, cooler temperatures, and dry soil moisture (Melton and Arora, 2016; Arora and Boer, 2005). We use two sets of CTEM fluxes which are driven by different meteorology. One set is generated using CRU-NCEP which we refer to as CTEM-CRU. Another set is generated using the GEM-MACH-GHG (run is refer to as CTEM-GEM). We use average model fluxes by averaging 2009-2010 fluxes, the only two years available.

Joint UK land environment simulator (JULES)

JULES is a community land surface model that has evolved from the UK Met Office Surface Exchange Scheme. Phenology in JULES affects leaf growth rates and timing of leaf growth/senescence based on temperature alone (Clark et al., 2011). Vegetation cover is predicted based on nine plant functional types that compete for space based on their relative productivity and height but are excluded from growing on agricultural land, based on a fraction of agriculture in each grid cell (Harper et al., 2018). CRU-NCEP was used as model forcing data.

4.2.2 Flux inversions

In addition to the NEE from the TBMs, posterior NEE fluxes from two flux inversion analyses are examined: one that assimilates boundary layer CO₂ observations (CT2016) and one that assimilates X_{CO₂} observations (GOSAT-Inv) from GOSAT. The motivation for employing two different inversion analyses is that flux estimates from these analyses are often divergent on regional scales. This is partially due to the sparsity of atmospheric CO₂ observations, which results in the observations strongly underconstraining NEE fluxes, although other factors, such as errors in model transport, further confound reliable NEE estimates. Thus, if features in optimized NEE are consistent between inversions, we have increased confidence in the robustness of the results.

For the first inversion analysis, we use optimized NEE from the National Oceanic and Atmospheric Administration’s CarbonTracker, version CT2016 (Peters et al., 2007, with updates documented at <http://carbontracker.noaa.gov>). CT2016 optimizes NEE by assimilating in situ observations of boundary layer atmospheric CO₂. It employs an ensemble Kalman filter approach to assimilate CO₂ with atmospheric chemical transport simulated by the TM5 offline atmospheric model (Krol et al., 2005). For CT2016, TM5 is driven by ERA-Interim assimilated meteorology from the ECMWF, with a horizontal

resolution of $3^\circ \times 2^\circ$ globally and $1^\circ \times 1^\circ$ in a nested grid over North America.

In a second inversion, referred to as GOSAT-Inv, NEE fluxes are optimized by assimilating GOSAT X_{CO_2} observations using the GEOS-Chem 4D-Var data assimilation, with version v35 of the GEOS-Chem adjoint model (Henze et al., 2007). To optimize surface fluxes, the 4D-Var cost function is minimized as described in Deng et al. (2014) to retrieve monthly scaling factors for prior ocean and terrestrial biosphere fluxes in each grid cell. We use an assimilation window of nine months and keep posterior fluxes from the first six months, then shift the inversion period forward by six months. Using this method, optimized NEE spanning 2010–2013 is generated. Prior NEE fluxes are based on the posterior fluxes from CT2016. We calculate a mean seasonal cycle using 3-hourly fluxes from the period 2010–2013. For error statistics, we assign 16% error to fossil fuels, 38% error to biomass burning, 44% error to ocean fluxes, and 44% error to terrestrial ecosystems, following Deng et al. (2014).

For the GOSAT retrievals, we use version v3.5 of the NASA ACOS GOSAT lite files. All bias-corrected measurements from the TANSO-FTS shortwave infrared channel are selected, including ocean glint, and high- and medium-gain nadir, which pass the quality flag requirement. We generate “super-obs” from the GOSAT retrievals by aggregating the observations to the grid size of our inversion. Error estimates are generated using the method described by Kulawik et al. (2016). The reduction in error with aggregation can be calculated using the expression $error^2 = a^2 + b^2/n$, where a represents systematic errors that do not decrease with averaging, b represents random errors that decrease with averaging, and n represents the number of satellite observations that are averaged (Kulawik et al., 2016). Kulawik et al. (2016) give $a = 0.8$ ppm and $b = 1.6$ ppm as mean Northern Hemisphere geometric (co-located) values for GOSAT, and these are the values that are used in our inversion analyses.

4.2.3 GOME-2 SIF

We use two different GOME-2 SIF products: NASA Level 2 GOME-2 version 26 (V26) 740 nm terrestrial chlorophyll fluorescence data (NASA-SIF, 2016; Joiner et al., 2013, 2016) and the GFZ Postdam product (GFZ-SIF, 2016; Köhler et al., 2015). We have selected observations spanning the period 2007–2014. A “daily correction” is performed to estimate daily average SIF from the instantaneous measurements (see supplementary material of Frankenberg et al. (2011b) for details of this calculation). The observations are then aggregated spatially to a $2^\circ \times 2.5^\circ$ grid and temporally to week of year by calculating the median value.

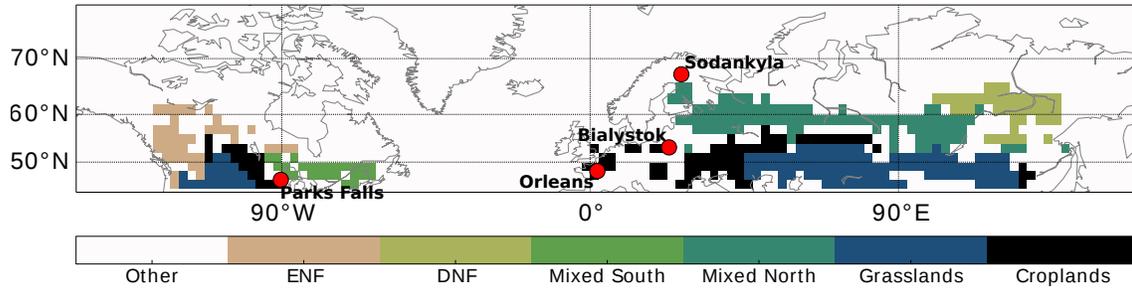


Figure 4.1: MODIS IGBP vegetation classification (at a horizontal resolution of $2^\circ \times 2.5^\circ$). Coloring indicates that the given vegetation type makes up more than 50% of the vegetation type in the gridcell. The vegetation regions are: ENF, DNF, southern mixed forests, northern mixed forests, grasslands and croplands. Red circles show the locations of the four TCCON sites examined in this study: Park Falls (45.9°N , 90.3°W), Orléans (48.0°N , 2.1°E), Białystok (53.2°N , 23.0°E), and Sodankylä (67.4°N , 26.6°E).

The relationship between SIF and GPP has been found to be dependent on vegetation type (Guanter et al., 2012), therefore, we examine different vegetation types separately. We examine GPP over the six northern vegetation regions shown in Figure 4.1: evergreen needle leaf forests (ENF), deciduous needleleaf forests (DNF), southern mixed forests, northern mixed forests, grasslands and croplands. The vegetation regions are based on the vegetation types from the MODIS International Geosphere-Biosphere Program (IGBP) land cover type classification product (Friedl et al., 2010; Channan et al., 2014). On a $2^\circ \times 2.5^\circ$ grid, each grid cell is assigned a given vegetation type if more than 50% of the grid is made up of a single vegetation type. Mixed forests occur in two distinct spatial regions in Eurasia and North America. For this reason, we split this category into two groups: "northern mixed forests" in Eurasia and "southern mixed forests" in North America.

To obtain a seasonal cycle for a given vegetation type, the spatial mean is calculated for each region. The seasonal cycle is then smoothed using a 3-week running mean. The mean offset of the SIF seasonal cycle is removed by subtracting the mean SIF value outside of the growing season. We ensure that the period is outside the growing season by checking that the time period is also outside of the growing season for FLUXCOM GPP. Finally, the seasonal cycle of SIF is normalized so that the integrated annual total SIF value equals 1. The normalization is required because the scaling between SIF and GPP is not well known and because TBMs produce a large spread in the magnitude of GPP (see Figure 4.2). Thus it is necessary to normalize annual GPP to compare the seasonality. The same scaling and averaging is applied to model GPP.

We construct an error estimate for SIF from the random noise in the seasonal cycle

and the error associated with removing the seasonal cycle offset. To estimate the random noise, we calculate the SIF seasonal cycle for 2007-2014 for each year individually, for both NASA and GFZ data products, and take the spread as the uncertainty. In reality, this spread is due to random error combined with interannual variability in SIF, thus it provides an upper bound on uncertainty. We estimate the error due to removing the offset as being the range of SIF values over the dates used to calculate the offset; again this provides an upper estimate of the errors. These errors are summed in quadrature to obtain the total error. See Appendix A.1 for more information on GOME-2 SIF error characterization. Note that these errors do not include systematic errors related to the retrieval or differences between SIF and GPP. Systematic errors are not well characterized and could impact the seasonality of SIF. During the winter months, snow cover and large air masses could introduce biases. Cloud cover is also a potential source of bias, as SIF is not observed under very cloudy conditions. However, Köhler et al. (2015) found that the cloud cover threshold did not have a large impact on the temporal patterns of SIF.

4.2.4 Atmospheric X_{CO_2}

TCCON X_{CO_2}

We examine X_{CO_2} from Sodankylä (Kivi et al., 2014; Kivi and Heikkinen, 2016), Bialystok (Deutscher et al., 2014), Orléans (Warneke et al., 2014) and Park Falls (Wennberg et al., 2014b). To simulate X_{CO_2} based on the NEE fluxes from the TBMs, we use the forward model component of the GEOS-Chem adjoint, at a horizontal resolution of $2^\circ \times 2.5^\circ$ with 47 vertical levels. The terrestrial biosphere fluxes are input at 3 hour resolution. We use annually repeating biospheric fluxes by averaging the fluxes over the years given in Table 4.1. All other fluxes are identical for each simulation, and are the same as those used in the inversion analyses described in Sec. 4.2.2. We simulate five years of CO_2 fields (2008-2012) with a one-year spin up period (2007) for each set of terrestrial biosphere fluxes. For comparison with TCCON, the simulated atmospheric CO_2 concentrations are sampled at the time (in 3 hour time intervals) and grid box in which TCCON measurements occur. The TCCON a priori information and averaging kernels are used to generate X_{CO_2} (using the method described in Wunch et al. (2011)).

Seasonal cycle fit

The X_{CO_2} time-series is detrended and an annual X_{CO_2} cycle is found by fitting a function of the form:

$$f(x) = a_0 + \sum_{k=1}^4 a_k \cos(2\pi kx) + b_k \sin(2\pi kx) \quad (4.1)$$

where x is the fraction of the year. Equation 4.1 is a truncated Fourier series and is similar to the NOAA seasonality fitting function that is commonly used to fit the mean annual cycle (Thoning et al., 1989). To estimate uncertainty in the fit, a Monte Carlo approach is used, in which a set of 50 fits is performed using randomly generated initial parameters and the standard deviation of the resulting curves is used to estimate the uncertainty. We then take the mean values of the curves as the best fit line. Note that this is only the uncertainty on the fit and does not include other sources of error, such as measurement error.

4.2.5 FLUXCOM

FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations, satellite remote sensing, and meteorological data (Tramontana et al., 2016; Jung et al., 2017). Jung et al. (2017) generate R_e products using several machine learning methods. For this study, we downloaded the products generated using random forests (RF), multivariate regression splines (MARS) and artificial neural networks (ANN) at daily resolution from the Data Portal of the Max Planck Institute for Biochemistry (<https://www.bgc-jena.mpg.de>). The mean seasonal cycle over 2008-2012 is calculated for each product.

4.2.6 Surface air and soil temperature

Surface air and soil temperatures are required for investigating the heterotrophic respiration (R_H) produced by TBMs (Sec. 4.3.4). Near-surface air temperature (T_{air}) is taken as the lowest atmospheric level of the assimilated meteorology from GEOS-5 after interpolation to $2^\circ \times 2.5^\circ$ with 47 vertical levels. This is the same meteorology used to run the GEOS-Chem model. Soil temperature (T_{soil}) is taken from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Gelaro et al., 2017). We take the soil temperature to be the maximum temperature of the top three layers of soil (TSOIL1, TSOIL2 and TSOIL3), which covers a depth down to 0.4 m (Koster et al., 2000).

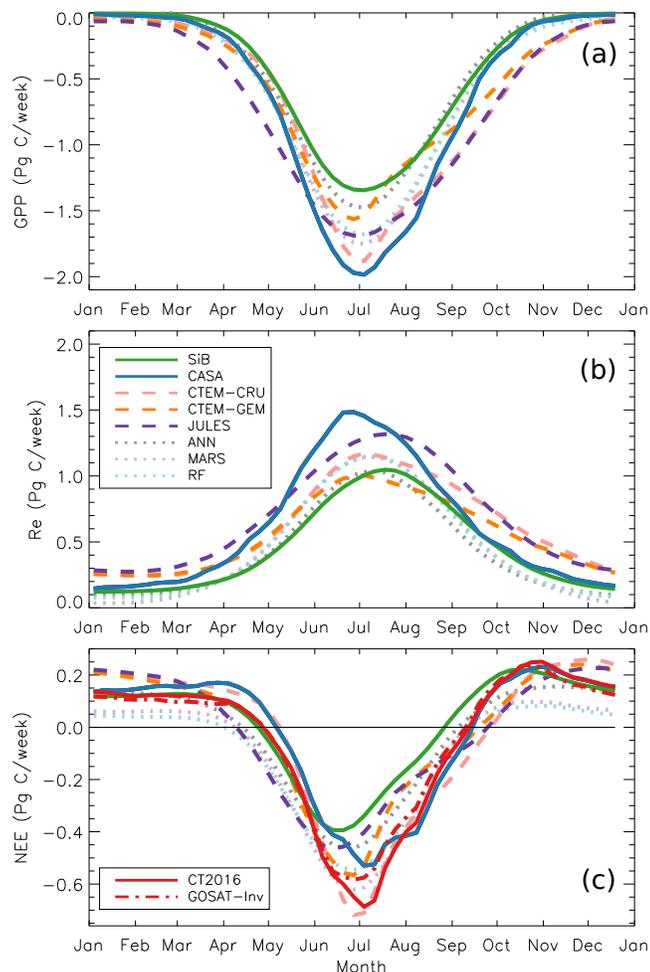


Figure 4.2: Mean seasonal cycles of (a) GPP, (b) R_e , and (c) NEE simulated by SiB3 (green), CASA (blue), CTEM-CRU (dashed salmon), CTEM-GEM (dashed orange), JULES (dashed purple), ANN (dotted gray), MARS (dotted purple-gray), and RF (dotted cyan) between 39°N and 65°N . CT2016 (solid red) and GOSAT-Inv NEE (dash-dotted red) are also plotted.

4.3 Results

4.3.1 Model fluxes

The mean GPP, R_e , and NEE fluxes for the TBMs, FLUXCOM, and inversions over all land between 39°N and 65°N are shown in Figure 4.2. This latitude range is used because, based on GEOS-Chem tagged tracer simulations, NEE over this region produces 85% of the amplitude of the seasonal cycle at Park Falls, 79% at Orléans, 87% at Białystok and 76% at Sodankylä. The mean seasonal cycle is calculated by averaging over the time periods listed in Table 4.1 to remove interannual variability. CT2016 is averaged over the period 2007–2012, and GOSAT-Inv is averaged over the period 2010–2013. Fluxes

are defined so that negative values represent removal of CO₂ from the atmosphere and positive values indicate emission to the atmosphere.

The TBMs have large differences in the magnitude of total annual GPP and R_e , consistent with the results from Huntzinger et al. (2012). The start of the growing season is quite variable between models, with JULES beginning earliest and SiB3 latest. For the timing of the end of the growing season, the TBMs are split into two groups: the TBMs with diagnostic phenology (CASA and SiB3) and FLUXCOM algorithms (ANN, MARS and RF) have growing seasons that end earlier, whereas the TBMs with prognostic phenology (CTEM and JULES) end two to four weeks later. For R_e , the start of the spring increase begins earliest for CASA and latest for SiB3. The timing of peak R_e is also quite variable between models, ranging from early June (CASA) to late July (SiB3). The diagnostic TBMs and FLUXCOM show similar timing in decreasing R_e throughout the fall, whereas the prognostic TBMs shows significantly higher R_e throughout the fall, mirroring what is seen in GPP. There are also significant differences between models in the magnitude of winter R_e . Prognostic TBMs give the highest winter R_e , followed by diagnostic TBMs, whereas FLUXCOM estimates are the lowest. The two sets of CTEM fluxes driven by different meteorology show some marked differences in fluxes, with CTEM-CRU GPP and R_e having larger magnitude than CTEM-GEM.

For NEE, all models produce a positive flux to the atmosphere in the winter months and net drawdown into the terrestrial biosphere during most of the growing season. However, there are differences between models in the timing of the period of net carbon uptake. The start of net uptake ranges between early April and May and the end of net uptake occurs between late August and early October. In contrast, there is close agreement in the seasonality of NEE between CT2016 and GOSAT-Inv. The most notable difference between the inversions is that GOSAT-Inv has somewhat weaker drawdown from June through September, and reaches peak drawdown earlier than CT2016. The inversion NEE fluxes generally fall in the middle of the modeled NEE fluxes from the TBMs and FLUXCOM during the spring and fall, but produce quite strong drawdown during the summer.

All three FLUXCOM NEE products are known to produce unrealistically large annual net sinks (Tramontana et al., 2016; Jung et al., 2017). However, both MARS and RF NEE show reasonable agreement with the flux inversion NEE through most of the growing season (June–September). This suggests that low R_e fluxes outside the growing season and enhanced drawdown during the early spring could cause the MARS and RF NEE annual bias. In contrast, ANN NEE shows reasonable agreement with the flux inversion NEE during the winter but weaker drawdown during the growing season, suggesting a

different source for the NEE annual bias.

4.3.2 Comparing model GPP and GOME-2 SIF

We compare the normalized seasonal cycle of SIF and model GPP over the six different vegetation regions (Fig. 4.3). The normalization is required because the scaling between SIF and GPP is not well constrained (see Sec. 4.2.3), so we are not able to evaluate the magnitude of GPP using SIF observations. The two GOME-2 SIF products (NASA and GFZ-Potsdam) are in close agreement throughout the year and for all vegetation regions. Differences between the SIF products are always less than the estimated uncertainties, and less than differences between SIF and modeled GPP. For the remainder of this study, comparisons will be performed with the NASA product.

Differences between the modeled GPP and NASA GOME-2 SIF are variable, although differences between model GPP and NASA GOME-2 SIF are generally consistent across the vegetation regions. For example, the growing season in NASA GOME-2 SIF ends earlier by several weeks than in the JULES GPP fluxes for all vegetation regions. The two TBMs with diagnostic phenology and FLUXCOM GPP deviate the least from the SIF seasonal cycle. The closer agreement for the diagnostic TBMs relative to the prognostic TBMs suggests that the assimilation of vegetation indices improves GPP fluxes, as expected, although differences in the driving meteorology could play a role. The normalized seasonal cycle of SiB3 GPP is always within the SIF uncertainty for all vegetation types and has a mean-root-mean square (RMS) difference of 0.0031. The normalized seasonal cycle of CASA GPP falls within the SIF uncertainties everywhere except in the fall for DNF and northern mixed forest regions, where CASA shows a more rapid decrease in GPP. CASA GPP has a mean RMS difference of 0.0037 across the vegetation regions. RMS differences for the FLUXCOM are 0.0038, 0.0038, 0.0044 for ANN, MARS and RF respectively. For all FLUXCOM algorithms, GPP is slightly phase shifted earlier in the year for all vegetation types, with the seasonal cycle starting and ending about one week earlier. The TBMs with prognostic phenology, CTEM and JULES, have larger differences between the seasonal cycle of modeled GPP and SIF. These prognostic models suggest growing seasons that are too long by several weeks compared with SIF across the vegetation regions studied here. See Appendix A.2 for additional details.

4.3.3 Comparing model NEE and TCCON X_{CO₂}

The mean seasonal cycles of observed and modeled X_{CO₂} at Sodankylä, Białystok, Orléans, and Park Falls are shown in Figure 4.4. All of the NEE fluxes reproduce the general shape

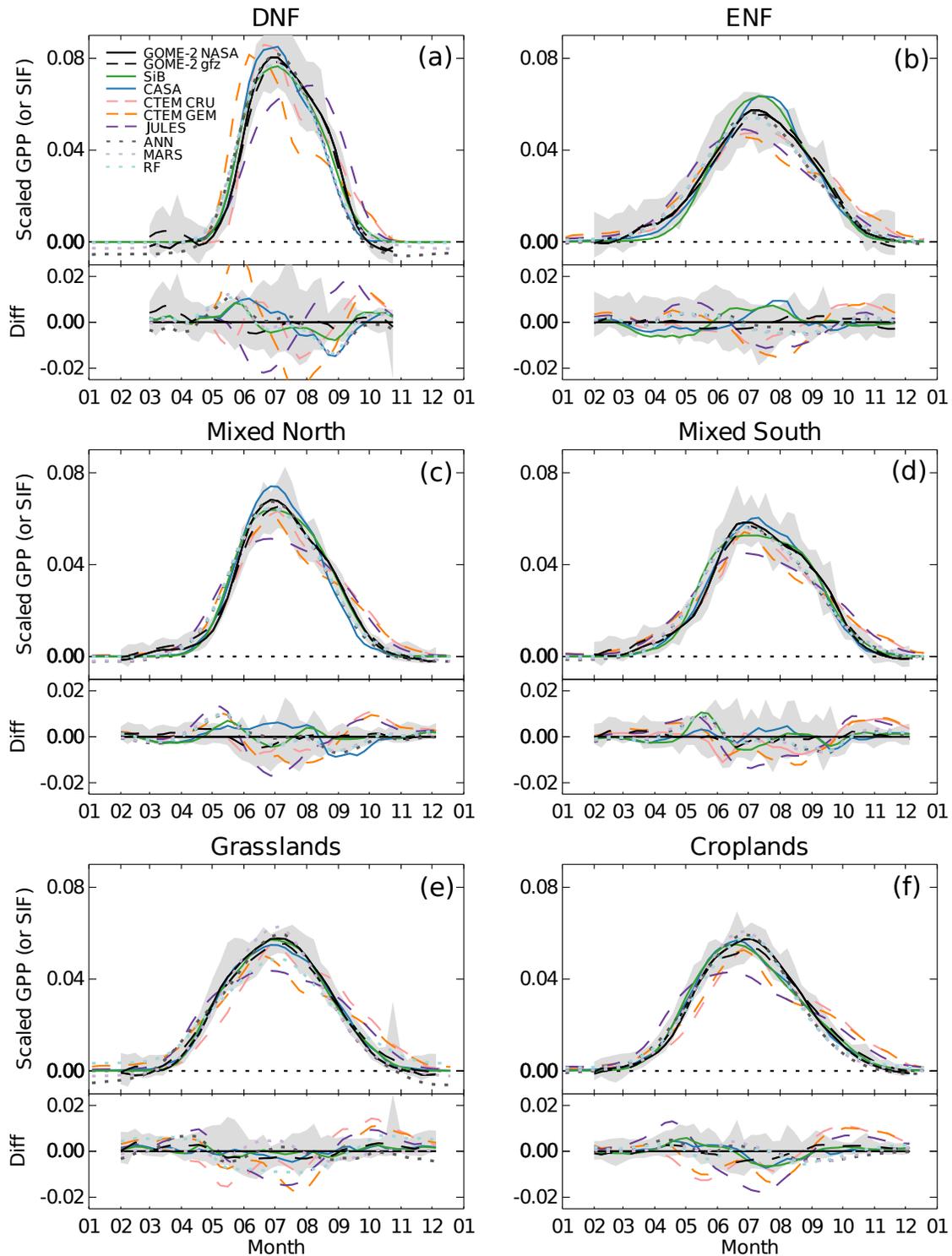


Figure 4.3: Normalized seasonal cycles of GOME-2 SIF (NASA and GFZ), and model GPP (SiB3, CASA, CTEM-CRU, CTEM-GEM, JULES and FLUXCOM ANN, FLUXCOM MARS, and FLUXCOM RF) for six vegetation regions. For each panel, the upper plot shows the seasonal cycle of GPP and SIF scaled so that the integral over the season equals one. The lower plot shows the difference between scaled TBM or FLUXCOM GPP and scaled NASA GOME-2 SIF. Grey shaded regions show the uncertainty estimate of the SIF seasonal cycle.

of the seasonal cycle, with maximum X_{CO_2} in the early spring and minimum in the late summer, although in some cases there are significant differences from the TCCON data in amplitude and phase. The posterior NEE fluxes from the two inversions produce the closest agreement with TCCON. This is expected because the inversions assimilate atmospheric CO_2 measurements (which are independent of TCCON observations). Of the two inversions, GOSAT-Inv has a smaller RMS difference of 0.21 ppm across the four sites, compared to 0.40 ppm for CT2016. However, the difference in RMS between GOSAT-Inv and CT2016 is not considered meaningful, because the difference could be due to transport differences between GEOS-Chem and TM5 (see Appendix A.3). Closer agreement for GOSAT-Inv is expected because the inversion was performed with the same chemical transport model used to simulate TCCON X_{CO_2} . The good agreement between the inversions and TCCON is reassuring because both inversions assimilate measurements that are independent of TCCON. These results show good agreement despite uneven observational coverage, errors in model transport, and biases in assimilated observations, all of which can strongly impact inversion analyses (Liu et al., 2014; Baker et al., 2006a; Miller et al., 2018). Of the TBMs, SiB3 and CASA give the best agreement with TCCON, with RMS differences of 0.50 ppm and 0.58 ppm, respectively. SiB3 gives a smaller seasonal cycle amplitude than TCCON, such that the annual minimum is about 1 ppm higher across the sites examined here. SiB3 also gives an earlier drawdown of X_{CO_2} than suggested by TCCON data at Sodankylä, Białystok and Park Falls. CASA gives a seasonal cycle that lags the TCCON data; the lag is largest at Orléans, where CASA is about two weeks later than the measurements. However, it is unclear whether these differences are significant, as they are approximately the same order of magnitude as transport errors. See Appendix A.3 for more details of transport error quantification, which is based on differences in X_{CO_2} simulated by GEOS-Chem and TM5 as well as previous studies (Houweling et al., 2010; Basu et al., 2011; Keppel-Aleks et al., 2011, 2012; Barnes et al., 2016).

Of the TBMs with prognostic phenology, CTEM-GEM has the lowest RMS difference relative to TCCON (0.73 ppm), whereas CTEM-CRU and JULES have RMS differences of 1.39 ppm and 1.18 ppm, respectively. Simulated X_{CO_2} using CTEM-CRU and CTEM-GEM NEE show quite different seasonal cycle shapes and phases. In comparison to CTEM-GEM, CTEM-CRU has a more rapid spring drawdown, which is delayed by ten days at all sites. The amplitude of the seasonal cycle is larger for CTEM-CRU than for CTEM-GEM by about 2 ppm at Orléans and Park Falls. In comparison to TCCON, CTEM-GEM produces better agreement in the shape and timing of the X_{CO_2} seasonal cycle than CTEM-CRU, and is within 2 ppm of the TCCON data at all sites. Simulated

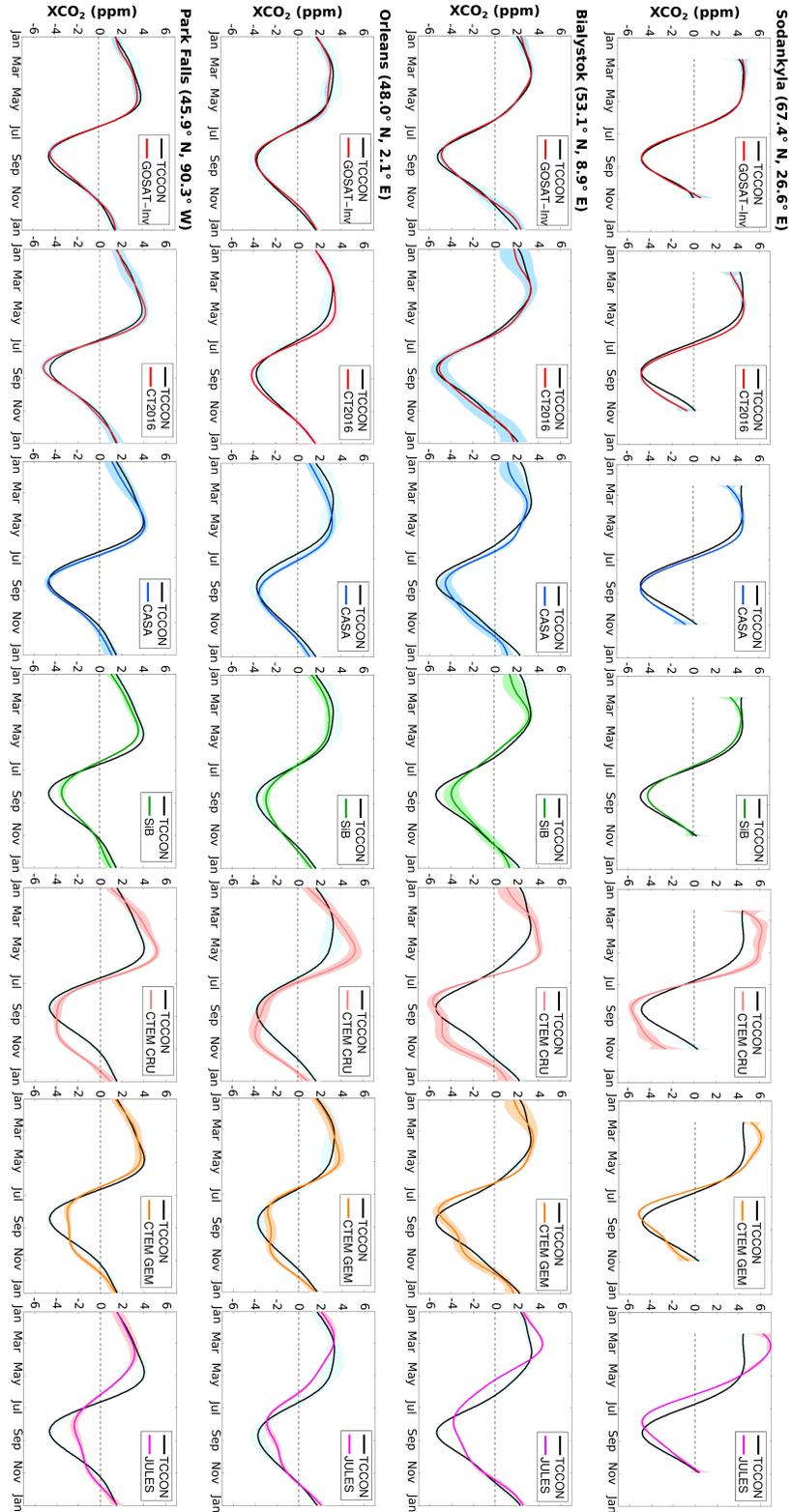


Figure 4.4: Five-year mean XCO_2 seasonal cycle at Sodankylä (top row), Białystok (second row), Orléans (third row) and Park Falls (bottom row). TCCON XCO_2 is shown in black. The columns show, from left to right: GOSAT-Inv, CT2016, CASA, SiB3, CTEM-CRU, CTEM-GEM, and JULES. Shaded regions indicate the uncertainty in the functional fit.

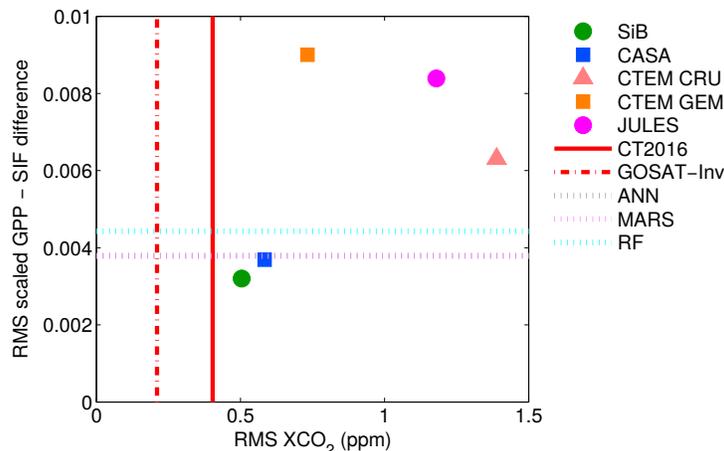


Figure 4.5: The RMS difference between normalized GPP and NASA GOME-2 SIF (averaged across all vegetation regions) versus the RMS difference between simulated X_{CO_2} and TCCON (averaged across the four TCCON sites).

X_{CO_2} using JULES NEE produces a seasonal cycle which is phase shifted early relative to TCCON at all sites by two to four weeks. JULES underestimates the amplitude of the seasonal cycle by 3 ppm at Park Falls but overestimate the amplitude by 2.5 ppm at Sodankylä.

4.3.4 Comparison of GPP and NEE

Figure 4.5 shows the mean RMS difference between model GPP and GOME-2 SIF versus the mean RMS difference between model-NEE-based X_{CO_2} and the TCCON X_{CO_2} . The TBMs with diagnostic phenology, SiB3 and CASA, have smaller SIF and TCCON RMS differences than the TBMs with prognostic phenology. This suggests that assimilating phenology improves both GPP and NEE fluxes. However, different driving meteorology was used for the diagnostic TBMs (driven by MERRA) and prognostic TBMs (driven by NCEP-CRU and GEM-MACH-GHG), thus it is unclear how much the differences in driving meteorology could have contributed to the differences between the prognostic and diagnostic TBM fluxes. It is possible that the driving meteorology is partially responsible for the better agreement with SIF and X_{CO_2} found with the diagnostic TBMs.

For the prognostic TBMs, the relative agreement between modeled GPP and GOME-2 SIF can be quite different than between model-NEE-based X_{CO_2} and TCCON X_{CO_2} . For example, X_{CO_2} simulated with CTEM-GEM NEE has a smaller RMS with respect to TCCON than CTEM-CRU, but CTEM-GEM GPP has a larger RMS with respect to SIF than CTEM-CRU. This shows two things. First, it indicates that a small GPP RMS dif-

ference does not necessarily predict small X_{CO_2} RMS difference. This suggests that there are compensating discrepancies in GPP and R_e that improve the NEE fluxes. Second, it shows that TBM fluxes are highly sensitive to the driving meteorology. High sensitivity to the driving meteorology has previously been reported for other TBMs (Poulter et al., 2011). Differences between CTEM-GEM and CTEM-CRU fluxes are primarily due to differences in moisture between the NCEP-CRU and GEM-MACH-GHG (Badawy et al., 2018). NCEP-CRU is wetter than GEM-MACH-GHG, which increases GPP and R_e fluxes in CTEM. These results highlight the difficulty in validating TBMs using only constraints on NEE. This remains a major difficulty in relating optimized NEE from flux inversions to TBM errors or deficiencies.

Comparing SiB3 and CASA

As discussed above, GPP fluxes and model-NEE-based X_{CO_2} from CASA and SiB3 show close agreement with SIF and TCCON data, respectively. Figure 4.6 compares the seasonality of GPP (times negative one), R_e and NEE between CASA and SiB3 over 39° - 65° N. The seasonality of GPP is similar for CASA and SiB3 but the seasonality of R_e and NEE shows significant differences. R_e peaks about a month earlier in CASA than in SiB3, causing CASA NEE to peak about a month later than in SiB3. Similar results are found for individual vegetation regions (see Appendix A.4). The differences between CASA and SiB3 found here are consistent with the results of Messerschmidt et al. (2013), who found that differences in seasonality of NEE between TBMs were primarily due to the differential phasing of R_e with respect to GPP.

The difference in timing of R_e between SiB3 and CASA could be explained by differences in the parameterizations of heterotrophic respiration (R_H) in the TBMs. SiB3 uses a “zero order” parameterization in which

$$R_H \propto f(T) \cdot f(M). \quad (4.2)$$

So that R_H is only dependent on soil temperature through a function,

$$f(T) = 2^{(T_{\text{soil}} - 298)/10}, \quad (4.3)$$

and soil moisture through a function $f(M)$ (Denning et al., 1996). This parameterization closely follows the seasonal cycle of soil temperature and peaks at the same time. Figure 4.6b shows SiB3 R_e and $f(T)$ scaled to have the same annual flux as R_e . Clearly, the seasonal cycle of SiB3 R_e and $f(T)$ align closely. In contrast, the CASA R_e curve is phase

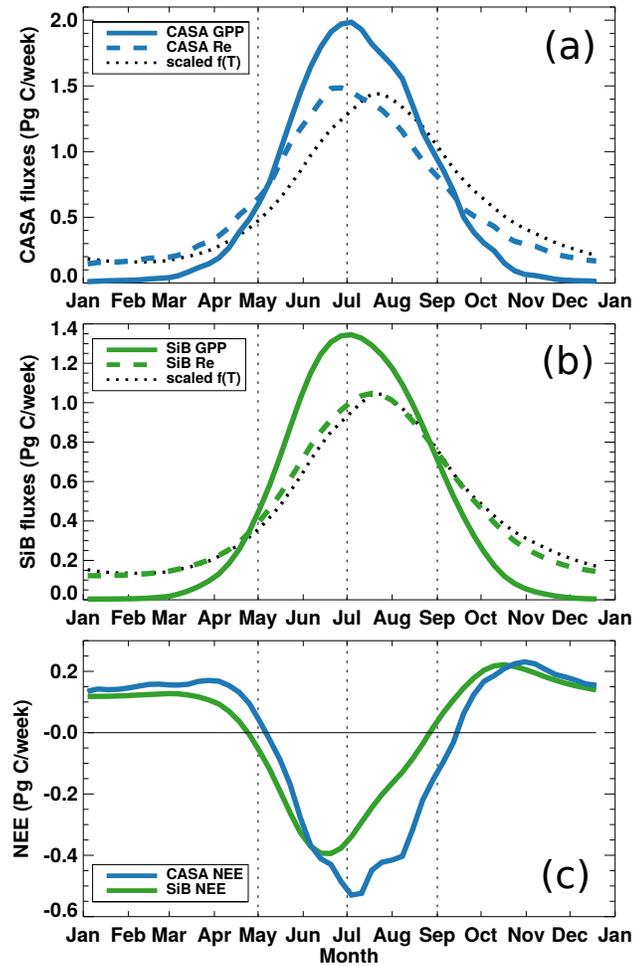


Figure 4.6: GPP $\cdot(-1)$, R_e , and NEE fluxes over 39° – 65° N. (a) CASA GPP $\cdot(-1)$, R_e and scaled $f(T)$ (Eq. 4.3). (b) SiB3 GPP $\cdot(-1)$, R_e , and scaled $f(T)$ (Eq. 4.3). (c) SiB3 and CASA NEE.

shifted earlier in the year relative to the $f(T)$ curve (Figure 4.6a). CASA uses a “first order” parameterization (Randerson et al., 1996) that has an additional dependence on the carbon pool size (C) available for R_H , resulting in a phase shift in R_e earlier in the season (Randerson et al., 1996). The reason is that leaf and fine root litter pool grow at the end of the growing season. Low temperatures throughout the winter prevent significant respiration, thus, this pool is near its maximum in the early spring. This results in a larger quantity of available substrate for respiration early in spring (Randerson et al., 1996).

4.3.5 Estimating R_e

Ideally, we would like to optimize R_e by performing a “flux inversion” that assimilates atmospheric CO₂ and SIF observations. This would involve first forward modeling SIF and X_{CO₂}, and then optimizing GPP and R_e by simultaneously assimilating SIF and X_{CO₂} observations. This is a complicated task, although tools to do this are under development (e.g., Schuh et al., 2016). Instead, we take a simpler approach that requires additional assumptions. We assume NEE and GPP are known, which allows us to simply calculate R_e . We take inferred NEE fluxes, which have been optimized using CO₂ observations and produce X_{CO₂} values that agreed closely with TCCON data, to be our “true” NEE. Modeled GPP from CASA, SiB3, and FLUXCOM are taken as the “true” GPP for our calculation. Given that the normalized seasonal cycles of SiB3, CASA and FLUXCOM GPP were in close agreement with SIF, it is reasonable to assume that the differences in NEE between the TBMs and inversions are primarily due to differences in the seasonal cycle of R_e and the magnitude of GPP. Note that it is not possible to use SIF in place of the modeled GPP because the scaling between SIF and GPP is not well known. The magnitude of GPP from SiB3 and CASA differ by about 45% at the peak of the growing season and FLUXCOM lies between the two TBMs (Figure 4.2). Any error in the magnitude of GPP will also be projected onto our optimized respiration estimate. Here we examine “optimized” respiration in several steps. First, we calculate an “optimized” R_e over the six vegetation types combined using the equation:

$$optR_{inv-mod} = NEE_{inv} - GPP_{mod}. \quad (4.4)$$

Therefore, $optR_{inv-mod}$ is the respiration, which, combined with CASA, SiB3 or FLUXCOM GPP, results in the inversion NEE. Second, we examine the sensitivity of $optR_{inv-mod}$ to the magnitude of GPP by calculating $optR_{inv-mod}$ after scaling GPP over a range of magnitudes, spanning the range of SiB3, CASA and FLUXCOM GPP products. Finally,

possible ecological implications for the seasonality of $optR_{inv-mod}$ are discussed.

Optimized R_e

GPP, R_e and $optR_{inv-mod}$ for SiB3, CASA and FLUXCOM are shown in Figure 4.7. The $optR_{inv-mod}$ curves show several features consistent across the ensemble. In particular, $optR_{inv-mod}$ generally gives a broader seasonal cycle with a less pronounced summer maximum than R_e produced by SiB3, CASA and FLUXCOM. However, comparing the actual fluxes is difficult due to the differences in magnitude and the fact that the annual net drawdown is different between TBMs. All FLUXCOM algorithms overestimate the net annual drawdown, such that the total annual $optR_{inv-FLUXCOM}$ is greater than that modeled by the algorithms. In contrast, SiB3 fluxes are generated by assuming the annual net NEE flux is approximately zero at each grid cell, and thus the magnitude of $optR_{inv-SiB3}$ is smaller than R_e produced by the TBM.

To simplify these comparisons, the seasonal cycle of GPP, R_e and $optR_{inv-mod}$ are normalized by the annual total flux. After normalization, more features become clear (Figure 4.7b,d,f,h). As with the comparisons with SIF, close agreement is found in the seasonality of normalized GPP between the models. In contrast, there are larger differences in normalized R_e between SiB3, CASA and FLUXCOM. In general, differences between $optR_{inv-mod}$ and model R_e are consistent with a broader seasonal cycle in $optR_{inv-mod}$. Normalized $optR_{inv-mod}$ is systematically lower in June–July and higher during October relative to modeled R_e . Optimized R_e for CASA ($optR_{inv-CASA}$) shows the largest differences from the mean. This behavior is due, primarily, to the larger magnitude of GPP in CASA relative to SiB3 and FLUXCOM.

Sensitivity to GPP magnitude

If Eq. 4.4 is re-written as:

$$optR_{inv-mod} = -GPP_{mod} \left(1 - \frac{NEE_{inv}}{GPP_{mod}} \right), \quad (4.5)$$

it is clear that $optR_{inv-mod}$ will become closer to the shape of GPP as the magnitude of GPP increases. Here we calculate $optR_{inv-mod}$ after scaling SiB3, CASA and FLUXCOM GPP over a range of values to examine the sensitivity of our results. Over the vegetation regions examined here, SiB3 GPP gives an uptake of 23 Pg yr⁻¹, ANN gives 23 Pg yr⁻¹, MARS gives 28 Pg yr⁻¹, RF gives 29 Pg yr⁻¹, and CASA gives 32 Pg yr⁻¹. Therefore, we scale GPP to vary over the range 23–32 Pg yr⁻¹ and recalculate $optR_{inv-mod}$. The

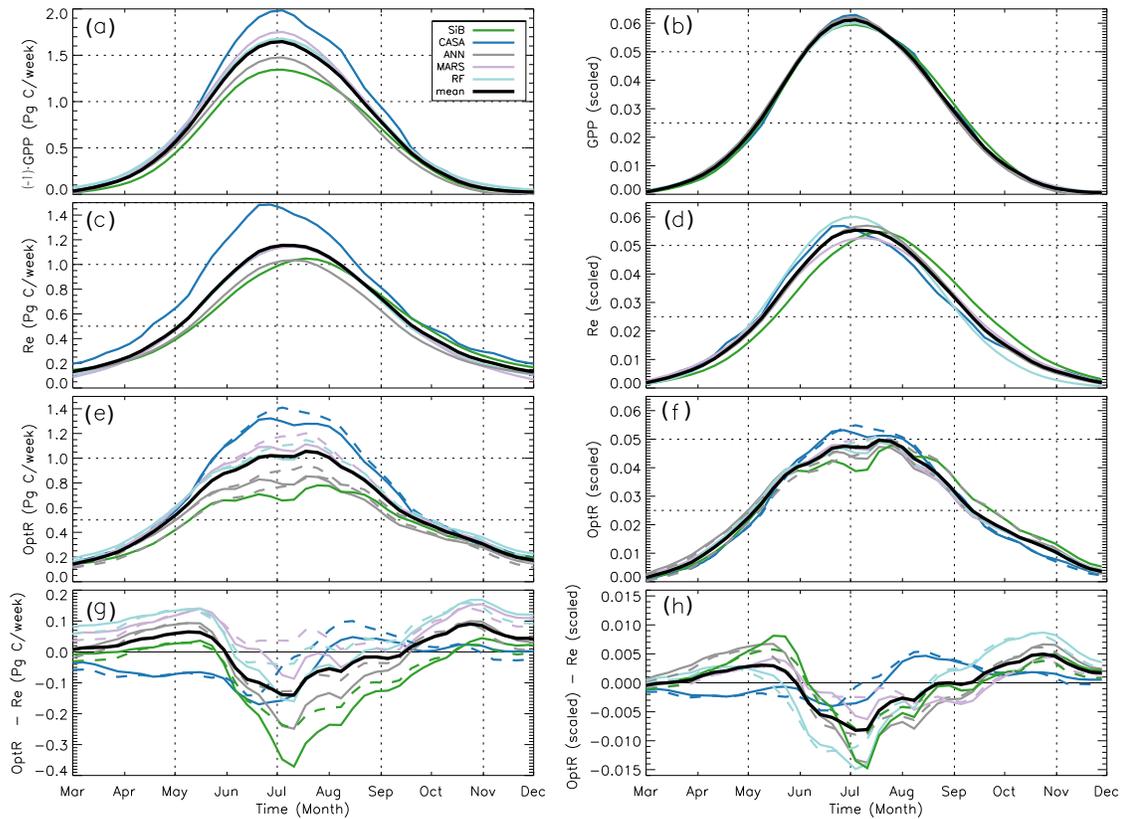


Figure 4.7: Left column shows (a) model $\text{GPP} \cdot (-1)$, (c) model R_e , (e) $\text{opt}R_{\text{inv-mod}}$, and (g) the difference between optimized and model R_e for SiB3, CASA, and FLUXCOM (ANN, MARS, and RF). $\text{opt}R_{\text{CT2016-mod}}$ (based on CT2016 NEE) is represented by solid lines, whereas $\text{opt}R_{\text{GOSATinv-mod}}$ (based on GOSAT-Inv) is indicated by dashed lines. Right column shows the normalized seasonal cycles of (b) model GPP, (d) model R_e , (f) $\text{opt}R_{\text{inv-mod}}$, and (h) the difference between optimized and model R_e . In all panels the solid, thick black line represent the mean of all the curves shown.

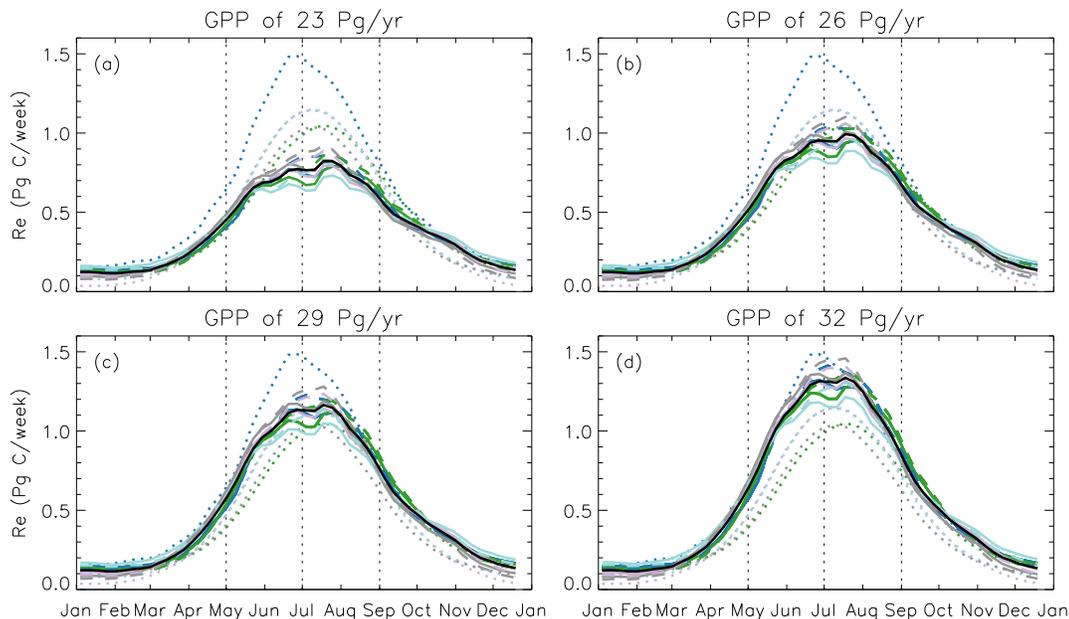


Figure 4.8: $optR_{inv-mod}$ (solid lines), after scaling GPP over a range of values, for SiB3, CASA and FLUXCOM. $optR_{inv-mod}$ are calculated after first scaling GPP to (a) 23 Pg yr^{-1} , (b) 26 Pg yr^{-1} , (c) 29 Pg yr^{-1} , and (d) 32 Pg yr^{-1} . R_e produced by SiB3, CASA, and FLUXCOM without scaling is indicated by the dotted lines. Colors are as in Figure 4.7.

resulting curves are shown in Figure 4.8. The $optR_{inv-mod}$ curves are similar when GPP is scaled to the same annual total for SiB3, CASA and FLUXCOM. For the range of annual total GPP examined here, $optR_{inv-mod}$ retains a broad summer maximum. However, a summer peak in respiration becomes more defined as GPP is increased.

Implications of $optR_{inv-mod}$

The $optR_{inv-mod}$ curves give systematically broader summer peaks in R_e than are modeled by the diagnostic TBMs: CASA and SiB3. Here we examine how R_e fluxes in these TBMs could be changed to bring the TBMs in agreement with $optR_{inv-mod}$. The objective is to determine whether realistic changes in TBM parameters could produce the seasonality found in $optR_{inv-mod}$, or whether changes to model equations are required. For this

analysis, we consider the model equations for SiB3 and CASA,

$$R_e = R_a + R_H, \quad (4.6)$$

$$R_a = -c_0 \cdot GPP, \quad (4.7)$$

$$R_H = C \cdot f(T) \cdot f(M), \quad (4.8)$$

$$f(T) = Q_{10}^{(T-c_1)/10}, \quad (4.9)$$

where R_a is autotrophic respiration, R_H is heterotrophic respiration, and C is the leaf and fine root litter carbon pool size. For CASA, the constants given in Eq. 4.6–4.9 are: $c_0 = 0.5$, $Q_{10} = 1.5$, $c_1 = 273.15$, and $T = T_{air}$ (Randerson et al., 1996). For SiB3, the constants given in Eq. 4.6–4.9 are: $c_0 = 0.6$, $Q_{10} = 2$, $c_1 = 298$, and $T = T_{soil}$ (Denning et al., 1996). Note that R_H in SiB3 does not have leaf and fine root litter carbon pool dependence, we have artificially introduced this dependence for this analysis.

In Sec. 4.3.4, it was found that the seasonality of heterotrophic respiration (R_H) has a high sensitivity to the leaf and fine root litter carbon pool size. In the following discussion, we show that adding realistic changes to the leaf and fine root litter carbon pool can bring R_e fluxes into closer agreement with $optR_{inv-mod}$, but does not completely resolve the differences. First, R_H is calculated using Eq. 4.6 and Eq. 4.7, and assuming $R_e = optR_{inv-mod}$, then, $R_H/f(T)$ is calculated. $R_H/f(T)$ is proportional to the leaf and fine root litter carbon pool, neglecting soil moisture dependence. Figure 4.9 shows $optR_{inv-mod}$, R_H , and $R_H/f(T)$ based on CASA and SiB3 model parameterizations. The calculated R_H increases in the spring (Mar–May). At the start of June, there is a rapid decrease in R_H which remains steady until mid-July when R_H increases. R_H remains high until late November when R_H decreases into winter. $R_H/f(T)$ indicates that the carbon pool generally increases from July through October and decreases through the rest of the year. Notably, there is a very rapid decrease in $R_H/f(T)$ during the spring. This rapid decrease corresponds to the decrease in R_H . These results suggest that the leaf and fine root litter carbon pool is largely depleted in early June. This lack of substrate then results in a significant decrease in R_H at this time. R_H then remains low until this carbon pool begins to increase in mid July.

Overall, the calculated seasonal cycle of R_H and $R_H/f(T)$ suggests that there is an abundance of substrate for heterotrophs to consume early in the growing season, which is largely depleted by June, resulting in reduced R_H over the summer. In general, the seasonal cycle of R_H and $R_H/f(T)$ calculated here seems plausible, however, it is unlikely that the carbon pool would begin to increase in July and is more likely this carbon pool would not increase until the fall (Randerson et al., 1996). Thus, plausible changes in

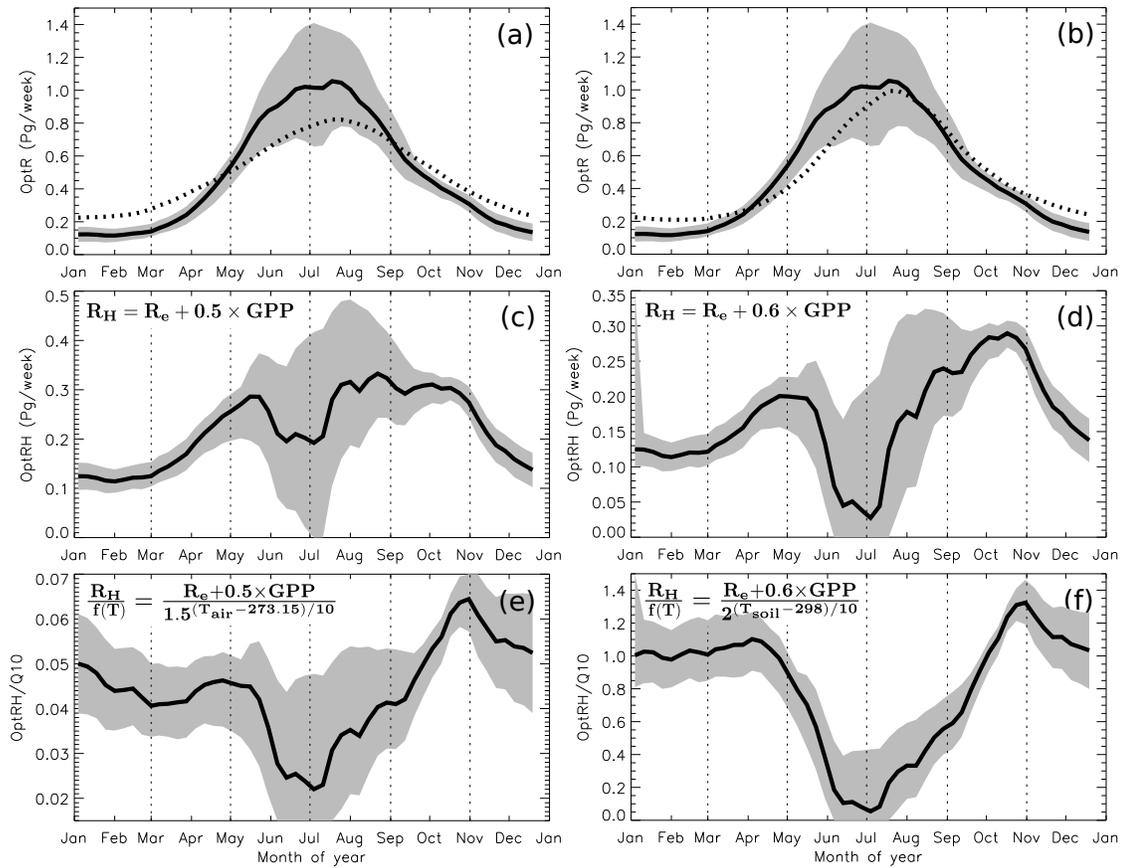


Figure 4.9: (a, b) $\text{opt}R_{\text{inv-mod}}$, (c, d) optimized R_H , and (e, f) optimized $R_H/f(T)$ using parameterizations from Randerson et al. (1996) (left column) and from (Denning et al., 1996) (right column). The solid black line shows the mean seasonal cycle and shaded gray region shows the range of optimized R_e for all NEE and GPP. The dotted black line shows $f(T)$ for CASA (a) and for SiB3 (b) scaled to fit plot area.

the seasonal cycle of the leaf and fine root litter carbon pool could partially account for the differences found between R_e and $optR_{inv-mod}$, but it seems unlikely that this could be the sole cause of the difference, as some unphysical changes to the carbon pool are required. It is possible that neglecting the soil moisture term impacts the result here, as soil moisture depletion during July could suppress R_e until the fall, resulting in more realistic seasonality.

The analysis presented here assumes that the model equations (Eqs 4.6–4.9) are correct. However, in some cases the true biosphere is known to deviate from these expressions. In many TBMs, including CASA and SiB3, R_a is assumed to be a constant fraction of GPP implying a constant CUE. However, observations suggest substantial variability in the CUE over the year (Arneeth et al., 1998; DeLucia et al., 2007; Heskell et al., 2013; Wehr et al., 2016; Tcherkez et al., 2017). Errors in model parameterizations, such as a constant CUE, will result in systematic errors in modeled fluxes (discussed in Sec. 4.4). Therefore, errors in model parameterization could also explain differences between R_e and $optR_{inv-mod}$.

Continental scales

So far, the analysis for $optR_{inv-mod}$ has focused on all land between 39°–65° N. In this section, $optR_{inv-mod}$ is calculated separately over North America (51°–167° W), Europe (12° W–41° E), and Asia (41°–180° E) between 39°–65° N to evaluate the continental differences in $optR_{inv-mod}$. Figure 4.10 shows the normalized seasonal cycle of GPP, R_e , and $optR_{inv-mod}$ for each continent (absolute fluxes are given in Appendix A.5), revealing significant differences in the seasonality of model GPP between the continents (Fig. 4.10a–c). Europe has the longest growing season: normalized model GPP increases earlier in spring than for the other continents, peaks in June and then slowly decreases from July to November. For North America and Asia, modeled GPP peaks in early July. Asia has the shortest growing season of the three continents.

Normalized model R_e indicates that Europe has the broadest seasonal cycle of the three continents, while North America and Asia have narrower summer peaks in modeled R_e (Fig. 4.10d–f). Similarly, normalized $optR_{inv-mod}$ indicates a broader seasonal cycle for Europe relative to North America and Asia (Fig. 4.10g–i). Comparing $optR_{inv-mod}$ and modeled R_e , we find that there is consistency between modeled R_e and $optR_{inv-mod}$ throughout the spring and summer for Europe relative to North America and Asia, but, $optR_{inv-mod}$ suggests enhanced R_e in the late fall (Fig. 4.10k). For North America and Asia, there are large differences between $optR_{inv-mod}$ and modeled R_e (Fig. 4.10j,l). For both continents, differences between $optR_{inv-mod}$ and modeled R_e are similar to those seen

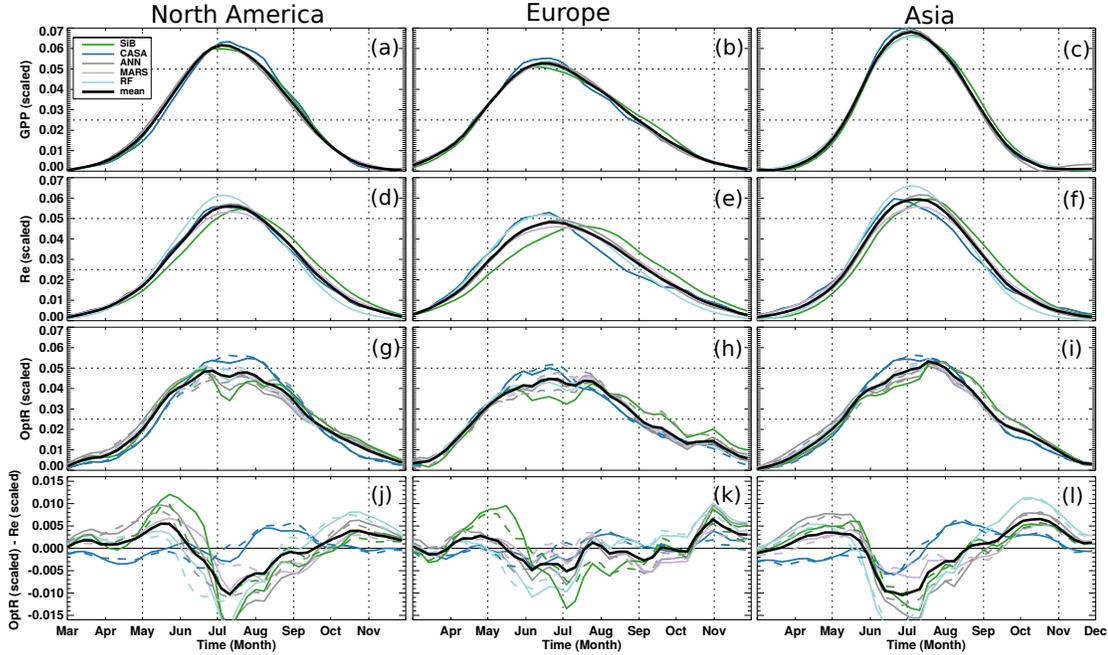


Figure 4.10: Normalized seasonal cycles of (a–c) model GPP, (d–f) model R_e , (g–i) $optR_{inv-mod}$, and (j–l) the difference between $optR_{inv-mod}$ and model R_e for (left column) North America, (middle column) Europe, and (right column) Asia. For subplots g–l, $optR_{CT2016-mod}$ is represented by solid lines, whereas $optR_{GOSATinv-mod}$ is indicated by dashed lines. In all panels the solid, heavy black line represent the mean of all the curves shown.

for the entire northern extratropics, with $optR_{inv-mod}$ suggesting reduced R_e in the early summer and enhanced R_e in the fall. Systematic differences between $optR_{inv-mod}$ and modeled R_e are largest for Asia, for which $optR_{inv-mod}$ suggests systematically lower R_e throughout June and into early July, and systematically higher R_e from mid-September to early November.

Why do $optR_{inv-mod}$ and model R_e generally show consistency for Europe, but large systematic differences for Asia? The answer may be linked to the differences in vegetation and climate between the continents. Europe has a milder climate than Asia due to the influence of the Gulf Stream. Furthermore, Europe has a large fraction of croplands while Asia has a high fraction of evergreen and deciduous needleleaf forests. The simplest explanation for the difference is that Europe has a milder climate and thus the TBMs and FLUXCOM suggest a broader season in R_e than for Asia. Since $optR_{inv-mod}$ generally indicates a broader seasonal cycle in R_e across the northern extratropics, this will result in a smaller difference for Europe since the seasonality of R_e is already broad.

4.4 Discussion

The differences between $optR_{inv-mod}$ and FLUXCOM R_e may indicate biases in the methods used to model GPP and R_e at FLUXNET sites. At these sites, partitioning methods are required to decompose observed NEE fluxes into GPP and R_e components. Standard methods perform the partitioning using hypothesized responses of GPP and/or R_e to light, water, and/or temperature (Wehr et al., 2016). These methods are also applied to generate FLUXCOM products (Tramontana et al., 2016). One possible source of bias in partitioning methods is that there is less turbulence at night which makes flux detection more challenging. Another possibility is that current partitioning techniques are missing ecological processes. Recently, Wehr et al. (2016) used isotopic measurements to determine daytime R_e in a temperate deciduous forest (Harvard forest) and found that daytime R_e was only about half as large as night-time NEE during June–July but roughly equal to night-time NEE during August–September. Standard partitioning methods do not account for this variability in daytime R_e , thus Wehr et al. (2016) suggest that FLUXNET R_e fluxes are over-estimated in June–July relative to August–September. This suggests that reduced daytime R_e fluxes during June–July could be present across much of the northern mid-latitudes, particularly in North America and Asia. The ecological explanation for reduced June–July daytime R_e suggested by Wehr et al. (2016) is the “Kok effect”, wherein leaf respiration is inhibited by light (Heskel et al., 2013). Therefore, this effect is largest during June–July when insolation is greatest. It is plausible that the reduced $optR_{inv-mod}$ fluxes obtained in the work presented here could be due to the Kok effect. However, it should be noted that Wehr et al. (2016) also found reduced GPP during June–July (to balance NEE). In contrast, we find strong agreement in the seasonality of FLUXCOM GPP and SIF. Therefore, it is possible that another mechanism is responsible. Croft et al. (2015) argue that neglecting variations in leaf chlorophyll abundances may bias LUE GPP estimates.

The continental-scale differences between $optR_{inv-mod}$ and model R_e found in this study are consistent with those expected from the Kok effect. The magnitude of the Kok effect has been found to depend on the plant species and ecosystem (Heskel et al., 2013; Tcherkez et al., 2017). As pointed out by Tcherkez et al. (2017), the inhibition of daytime respiration was found to be small for a multi-site study of European grasslands (Gilmanov et al., 2007), but significant in a North American forest (Jassal et al., 2007). Furthermore, the impact of the Kok effect has been suggested to be larger for evergreen vegetation than for deciduous vegetation (Wohlfahrt et al., 2005; Heskel et al., 2013). Therefore, one could expect that the Kok effect would be larger in temperate Asia than Europe,

which has more evergreen forests but less cropland. Consistent with this, $optR_{inv-mod}$ suggests greater reductions in R_e in June–July in Asia than in Europe (Sec. 4.3.5).

The inhibition of daytime R_a due to the Kok effect also has implications for TBMs, as it implies variability in the CUE throughout the year. In a recent review, Tcherkez et al. (2017) concluded that leaf day respiration should be regarded as a central actor of plant carbon-use efficiency. Furthermore, He et al. (2018) argue that understanding the mechanisms behind spatiotemporal changes in R_a is critical for better quantifying global CUE. However, as discussed in Sec. 4.3.5, many TBMs assume constant CUE throughout the year. These results suggest that this assumption may introduce a systematic bias of high R_a in June–July relative to August–September in TBMs, and may explain why differences in the leaf and fine root litter carbon pool between SiB3 and CASA could not fully explain the differences between model R_e and $optR_{inv-mod}$.

Enhanced fall R_e recovered in $optR_{inv-mod}$ provides an interesting parallel with a recent study of the Alaskan carbon cycle by Commane et al. (2017), who used aircraft and tower atmospheric CO₂ observations and GOME-2 SIF observations to show that R_e fluxes from Alaskan tundra are significant during October–December. We obtain similar enhanced fall R_e over our much larger study region, with the largest enhancement of R_e in Asia from mid-September to mid-November. These results suggest that fall R_e fluxes are larger across boreal and northern regions than has previously been appreciated. Precisely why we obtain enhanced fall R_e relative to SiB3, CASA and FLUXCOM is unclear. Commane et al. (2017) showed that TBMs do not represent fall respiration well, especially when soil temperatures are near 0 °C. They suggested that during the zero curtain period, when the active layer is freezing from above and below, microbes continue to metabolize in the subsurface as long as liquid water is present (Zona et al., 2016), and that this process can persist for months after the surface is frozen and snow covered. It is unclear if a similar process could explain the enhanced fall R_e in our more southern domain (39°–65° N). Monson et al. (2006) showed that snow cover during the fall can insulate the soils, producing enhanced R_e . This mechanism could provide enhanced R_e fluxes in the northern parts of our domain. A second possibility is that the size of the leaf and fine root litter carbon pool has an impact on the fall R_e . The comparison presented in Sec. 4.3.5 suggests that this carbon pool may increase rapidly over the fall and peak in mid-October. If this is the case, it would provide a large quantity of substrate for heterotrophic respiration in the mid to late fall, and provide a possible explanation for the enhanced rates of fall R_e found in this study.

4.4.1 Remaining challenges

One challenge with exploiting SIF and X_{CO_2} data is the differences in the scales on which the two types of observations provide information on CO₂ surface fluxes. The footprints of SIF observations are highly localized; the observed SIF is representative of the footprint of the satellite. In contrast, an X_{CO_2} observation has a large surface NEE footprint. On seasonal timescales, variations in X_{CO_2} are driven by the meridional flux distribution (Keppel-Aleks et al., 2011). Therefore, the scales that can be examined by combining SIF and X_{CO_2} observations are limited by the scales on which X_{CO_2} observations can inform surface fluxes.

Differences between inversions in regional net annual fluxes have been well documented, as annual net fluxes have been the primary focus of the majority of CO₂ flux inversion studies. To a lesser extent, regional-scale differences in the seasonal cycle of NEE between inversions have also been documented in the literature, particularly when comparing inversions which assimilate in-situ versus GOSAT observations (Chevallier et al., 2014; Ishizawa et al., 2016). As a demonstration of these differences, we compare the NEE fluxes from CT2016 and GOSAT-Inv at $2^\circ \times 2.5^\circ$ resolution. Figure 4.11 shows the maximum rate of NEE drawdown during the growing season for CT2016 and GOSAT-Inv for each grid-cell. There are substantial differences between the inversions which have structure on the scales of the biomes examined in this study. Thus, it is unlikely that reliable NEE seasonal cycle estimates are possible on these scales, however, more research is needed to quantify the scales that can be constrained.

As demonstrated in Sect 4.3.5, the magnitude of GPP has a large influence on the $optR_{\text{inv-mod}}$ seasonal cycle. Thus, the fact that there is no consensus on the global total GPP (Anav et al., 2015) remains a major limitation on inferring R_e . Furthermore, we do not have a forward model that can relate GPP to SIF observed from space. Previous studies, which assimilated SIF to optimize GPP, have relied on independent GPP estimates (Parazoo et al., 2014; Bowman et al., 2017; Liu et al., 2017). Parazoo et al. (2014) prescribed the annual magnitude of GPP (using MPI-BGC), but optimized the temporal-spatial structure redistributed by the assimilation of SIF. Therefore, the development of a forward model relating GPP to SIF would greatly improve our ability to isolate GPP and R_e fluxes. These estimates could then be directly compared to FLUXCOM and TBM GPP and R_e .

Furthermore, most previous studies that assimilated SIF to optimize GPP have relied on independent GPP estimates (Parazoo et al., 2014; Bowman et al., 2017; Liu et al., 2017). Parazoo et al. (2014) prescribed the annual magnitude of GPP (using MPI-BGC) but optimized the temporal-spatial structure redistributed by the assimilation of SIF.

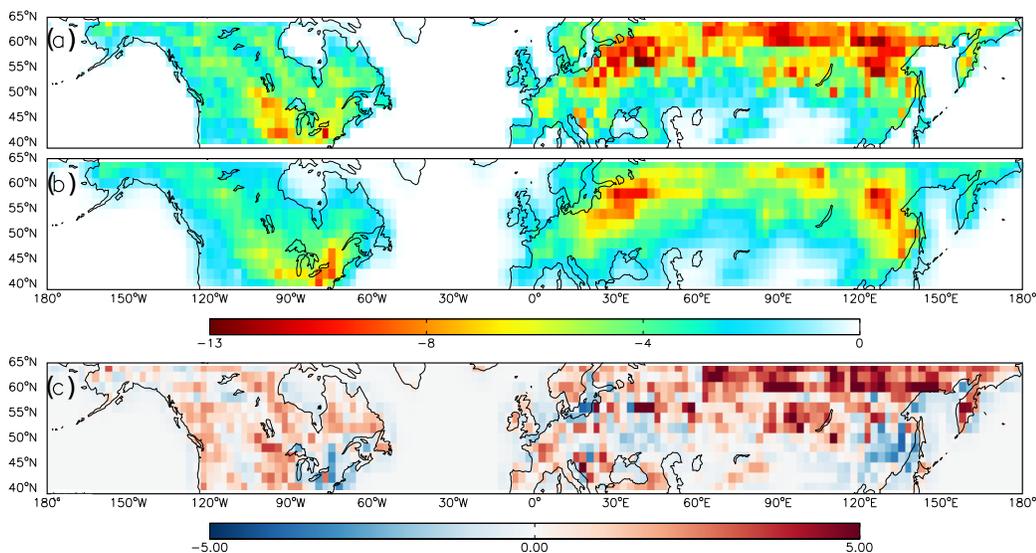


Figure 4.11: Maximum rate of drawdown ($\text{gC m}^{-2} \text{day}^{-1}$) during the growing season for (a) CT2016, (b) GOSAT-Inv, and (c) GOSAT-Inv minus CT2016.

Recently, forward models relating GPP to observed SIF have been developed (Van der Tol et al., 2014; Lee et al., 2015). Employing these models in inverse calculations to optimize GPP could improve our ability to isolate R_e fluxes. These estimates could then be directly compared to FLUXCOM and TBM GPP and R_e .

4.5 Conclusions

In the first part of this study (Sec. 4.3.1–4.3.4), GOME-2 SIF and TCCON X_{CO_2} data were employed to evaluate carbon fluxes produced by three FLUXCOM products (ANN, MARS, and RF) and four TBMs (CTEM, JULES, CASA, and SiB3). In general, the normalized seasonal cycle of GPP for the TBMs with diagnostic phenology (SiB3 and CASA) and FLUXCOM were in close agreement with SIF (with RMS differences less than 0.0045). TBMs with prognostic phenology (CTEM and JULES) showed comparably worse agreement with SIF (with RMS differences greater than 0.006). The closer agreement between the seasonality of GPP and SIF for the diagnostic TBMs relative to prognostic TBMs suggests that the assimilation of vegetation indices improves GPP fluxes. However, we did not control for the driving meteorology, which could be partially responsible for the differences. Comparisons of simulated X_{CO_2} with TCCON showed close agreement for the diagnostic TBMs (with RMS differences less than 0.6 ppm) and worse agreement for the prognostic TBMs (with RMS differences greater than 0.7 ppm). Differences in the driving meteorology for CTEM resulted in large differences in simu-

lated fluxes and, consequently, in the agreement with observations.

In the second part of this study (Sec. 4.3.5), a simple method for estimating R_e fluxes using constraints on GPP and NEE was described and tested. In this method, $optR_{inv-mod}$ is calculated by making idealized assumptions about NEE and GPP. Strong agreement between the seasonality of the normalized seasonal cycle of SiB3, CASA, and FLUXCOM GPP with SIF suggested that differences between model-NEE-based X_{CO_2} and TCCON X_{CO_2} seasonality were driven by differences in R_e . Thus, we assumed that GPP from SiB3, CASA, and FLUXCOM were correct. To generate constraints on NEE, we used optimized NEE from two flux inversions, which produced a posteriori CO₂ fields that were in close agreement with TCCON data. Assuming that GPP and NEE were known, $optR_{inv-mod}$ was calculated as the difference between the optimized NEE and the TBM and FLUXCOM GPP (Eq. 4.4).

Using this approach, we calculated $optR_{inv-mod}$ for all possible combinations of GPP (ANN, MARS, RF, SiB3, and CASA) and NEE (CT2016, GOSAT-Inv). This ensemble of GPP and NEE was found to produce $optR_{inv-mod}$ with reasonable precision. The largest differences in the seasonality of $optR_{inv-mod}$ curves were due to the magnitude of GPP, which is variable among models but not well constrained by observations (Anav et al., 2015). $optR_{inv-mod}$ exhibited a broader summer peak than R_e modeled by SiB3, CASA, and FLUXCOM. The seasonality of $optR_{inv-mod}$ suggested reduced R_e in the summer but enhanced R_e in the spring and fall. Differences were systematic from FLUXCOM and the TBMs during June–July, when $optR_{inv-mod}$ was reduced, and during October, when $optR_{inv-mod}$ was enhanced. Reduced R_e during the early summer is consistent with the results of Wehr et al. (2016), and could be explained by the Kok effect (inhibition of leaf respiration by light). Enhanced fall R_e is consistent with Commane et al. (2017), who found significant fall R_e in Alaska, and suggests that fall R_e may be greater than previously appreciated.

The seasonality of $optR_{inv-mod}$ has significant implications on R_e calculations in TBMs. We demonstrated that carbon pool dependence for R_H is important for recovering R_e consistent with $optR_{inv-mod}$, however, it was also shown that this carbon pool dependence could not solely explain systematic differences in R_e (Sec. 4.3.5). Instead, the results suggest that using a constant CUE throughout the year introduces biases into R_a fluxes. Overall, the results suggest that the inclusion of variable CUE for R_a and carbon pool dependence for R_H are important for accurately simulating R_e .

Our results highlight the utility of the SIF data for informing CO₂ flux inversions. The significant differences found between bottom-up and top-down estimates of R_e motivate further development of inversion methods to assimilate both atmospheric CO₂ and SIF

observations. Based on this analysis, both CASA and SiB3 produce realistic prior GPP and NEE fluxes and can therefore provide useful prior fluxes for future analysis. A current limitation is that only large-scale $optR_{inv-mod}$ was investigated due to the fact that the accuracy of the seasonal cycle of NEE from flux inversions on smaller scales is uncertain. The scales over which the mean seasonal cycle of NEE is consistent between inversions is not well documented in the literature, but needs to be further investigated to provide GPP and R_e estimates on smaller scales.

Chapter 5

On what scales can GOSAT flux inversions constrain anomalies in terrestrial ecosystems?

This chapter is adapted from:

Byrne, B., Jones, D. B. A., Strong, K., Polavarapu, S. M., Harper, A. B., Baker, D. F., and Maksyutov, S. (2018). On what scales can GOSAT flux inversions constrain anomalies in terrestrial ecosystems? *submitted*, Atmos. Chem. Phys.

5.1 Introduction

Many studies have investigated IAV in the carbon cycle using observations from the global network of in situ CO₂ measurements (Bousquet et al., 2000; Rödenbeck et al., 2003; Bruhwiler et al., 2011; Peylin et al., 2013; Marcolla et al., 2017). This network provides by far the longest direct record of atmospheric CO₂ measurements, with many sites functioning for decades. However, the spatial distribution of sites is inhomogeneous, with sites most densely located in North America and Europe and comparatively few elsewhere. Therefore, in situ observations from the global observation network are relatively insensitive to CO₂ fluxes over much of Asia and in the tropics (Chapter 3), where IAV is the largest. Recently, space-based observations of atmospheric CO₂ have provided expanded observational coverage for atmospheric CO₂. One of the satellites, GOSAT, has been providing measurements of atmospheric CO₂ since 2009. With multiple years of measurements, it is now possible to investigate IAV in the carbon cycle with GOSAT data.

In this study, we investigate interannual flux anomalies estimated from GOSAT measurements using the “flux inversion” method, wherein surface fluxes are estimated from atmospheric CO₂ measurements using a tracer transport model and Bayesian inverse methods. A series of flux inversions using the GEOS-Chem 4D-Var data assimilation system (Henze et al., 2007) are performed with different spatial resolutions, prior fluxes and prior error covariances. We also examine the posterior fluxes from two publicly available flux inversion estimates, the GOSAT Level 4 product (Maksyutov et al., 2013) and CarbonTracker, version CT2016 (Peters et al., 2007, with updates documented at <http://carbontracker.noaa.gov>), which is a flux inversion that assimilates CO₂ observations from the surface network.

Posterior anomalies in NEE from the inversions are compared with “proxies”: variables that are known to be closely associated with IAV in the carbon cycle. Agreement between the anomalies in the inversions and proxies provides corroborating evidence that the inversions are correctly recovering anomalies in NEE (Deng et al., 2016; Liu et al., 2017). Three proxies are examined: soil temperature (T_{soil}) anomalies from the MERRA-2 reanalysis (Reichle et al., 2011, 2017), the Monthly Self-calibrated Palmer Drought Severity Index (scPDSI) (Dai, 2017), and SIF observed by GOME-2 (Joiner et al., 2016). We also use flux data from FLUXCOM, which provides data-driven NEE anomaly estimates (Tramontana et al., 2016; Jung et al., 2017).

Anomalies in temperature and water availability are closely linked to anomalies in terrestrial ecosystems. On the local scales of FLUXNET sites (Baldocchi et al., 2001), temperature and precipitation have both been shown to be major controls on NEE (see Baldocchi et al. (2018) for a review). On regional and larger scales, stronger correlations have been found with temperature anomalies than with precipitation anomalies (Wang et al., 2013; Jung et al., 2017), particularly in the tropics. Jung et al. (2017) suggest that this is partially due to sub-regional-scale spatial variability in water availability anomalies that compensate, thereby reducing the influence of these anomalies on larger scales, while temperature anomalies are generally more spatially coherent. Nevertheless, both temperature and water availability anomalies strongly influence NEE anomalies over a wide range of scales. Temperature anomalies and drought conditions in the tropics are largely driven by ENSO variability. During the warm phase of ENSO (El Niño) large-areas of tropical land become dryer and warmer, leading to a net emission of CO₂ from the land to the atmosphere, which amplifies the atmospheric CO₂ growth rate. During the cold phase of ENSO (La Niña), much of the tropical land is cooler and wetter than average, leading to anomalously low CO₂ growth rates (Jones and Cox, 2005). In the extratropics, there is also significant variability in the carbon balance of terrestrial ecosystems related

to temperature and moisture anomalies (Conway et al., 1994; Bousquet et al., 2000). Wunch et al. (2013) show that the summer minima in the column-averaged dry-air mole fraction of CO₂ (X_{CO_2}) observed at northern mid-latitude TCCON sites is correlated with surface temperature, indicating that mid-latitude positive temperature anomalies correspond to reduced uptake by the northern extratropical biosphere during the growing season. Many studies have examined extreme heatwaves or droughts in the extratropics, such as the 2003 European heatwave (Ciais et al., 2005) and 2010 Russian heat wave and wildfires (Guerlet et al., 2013). In these cases, positive temperature anomalies and drought conditions result in a release of CO₂ from terrestrial ecosystems to the atmosphere. Zscheischler et al. (2014) show that relatively few extreme events dominate anomalies in GPP, and likely NEE. Due to the large seasonal cycle of temperature, precipitation and insolation in the extratropics, the relationship between anomalies in NEE and the proxies is likely a function of time of year. We focus our study of the northern extratropics to the Northern Hemisphere summer (JJA).

SIF is the emission of radiation by chlorophyll during photosynthesis and thus provides a measure of GPP (Papageorgiou and Govindjee, 2007; Frankenberg et al., 2011b; Guanter et al., 2012; Yang et al., 2015; Damm et al., 2015; Zhang et al., 2016a,b; Wood et al., 2017). Therefore, reduced GPP is associated with reduced SIF, and vice-versa. The relationship to anomalies in NEE is less direct because GPP and R_e anomalies are highly correlated (Baldocchi et al., 2018). Therefore, the extent to which SIF anomalies and NEE anomalies should be correlated is not well understood. One study, Shiga et al. (2018), shows that SIF can be used to inform the spatiotemporal distribution of NEE over North America.

Upscaled NEE estimates from eddy-covariance measurements at flux towers can be used to generate an observation-based estimate of NEE anomalies. Kondo et al. (2015) compared the GOSAT L4 product and empirical eddy flux upscaling and found similar responses to climate anomalies in temperate and boreal regions, while poorer agreement was found in the tropics. Here, we use upscaled NEE estimates from FLUXCOM that are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations of CO₂ fluxes, satellite remote sensing, and meteorological data (Tramontana et al., 2016; Jung et al., 2017). For this study, upscaled fluxes generated using MARS are used. Similar results were found for other upscaling algorithms.

It is important to acknowledge that none of these proxies (or FLUXCOM) should be expected to be perfectly correlated with the true NEE anomalies. Therefore, when there is disagreement between the inversions and proxies, it unclear whether this should

be attributed to the inversion NEE or the proxy. Comparisons of flux inversions with the proxies are most useful for identifying “positive” results for which the assimilation of atmospheric CO₂ observations has introduced a strong correlation with the proxies. However, these comparisons are less useful for identifying the limits of the inversions with “negative” results, in which the null hypothesis (no correlation) cannot be rejected.

In addition to comparing our flux inversions with the proxies and the FLUXCOM data, we also compare several TBMs with the proxies and the FLUXCOM data. TBMs simulate GPP and R_e , and therefore provide estimates of NEE. TBMs are widely applied to simulate projections of the future carbon cycle, however, different models show large disagreements on the relative importance of different processes driving the uptake (Huntzinger et al., 2017). One of the primary goals of atmospheric flux inversions is to provide better constraints on NEE to evaluate these models. Therefore, it is useful to determine whether the agreement between flux inversions and the proxies is closer than the agreement between TBMs and the proxies.

This study has three main objectives. The first is to quantify the agreement between GOSAT flux inversions and the flux proxies. This will be useful for identifying the utility of using proxies to corroborate flux inversions results. The second is to determine the spatial scales over which the GOSAT inversion constrains flux anomalies. GOSAT observations are expected to best constrain fluxes on large scales, such as the entire tropics. As scales decrease, finer scale structures in the atmospheric CO₂ fields are required to constrain fluxes, the smallest scales at which GOSAT observations provide useful constraints on NEE anomalies is currently unclear. We quantify the ability of GOSAT flux inversions to quantify NEE anomalies over a range of spatial scales by, first, examining the agreement between the inversions and proxies over a range of spatial scales and, second, examining the ability of GOSAT inversions to recover true flux anomalies by performing a series of OSSEs. Monthly anomalies in the tropics are examined throughout the year while anomalies in the northern extratropics are examined during the summer (JJA). The third objective is to quantify the sensitivity of the results for the first two objectives to the inversion set up. This is investigated with a series of GOSAT flux inversions with different model resolution, prior fluxes, and prior error covariances.

This chapter is structured as follows. In Sec. 5.2, we describe the datasets used, flux inversions performed, and how anomalies are calculated. Sec. 5.3, presents the results of our analysis. Flux inversion NEE anomalies are compared with the proxies in the tropics and northern extratropics individually. We then present an OSSE to examine the smallest spatial scales for which NEE anomalies can be recovered from GOSAT observations. Sec. 5.4 discusses the agreement in anomalies between the GOSAT flux inversions and

proxies, the scales constrained by GOSAT flux inversions, and the sensitivity of these results to the inversion set up. Finally, conclusions are given in Sec. 5.5.

5.2 Data and methods

5.2.1 FLUXCOM NEE data

FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations, satellite remote sensing, and meteorological data (Tramontana et al., 2016; Jung et al., 2017). Explanatory variables from remote sensing measurements are averaged to produce a mean seasonal cycle (Tramontana et al., 2016), such that all IAV is introduced by the driving reanalysis (NCEP CRU). In particular, IAV is driven by air temperature, incoming global radiation combined with the mean seasonal cycle of NDVI, and model based water availability index. Jung et al. (2017) generate NEE products using several machine learning methods. We downloaded these products from the Data Portal of the Max Planck Institute for Biochemistry (<https://www.bgc-jena.mpg.de>). We find that the different algorithms generally give similar results, therefore we only present results using the multivariate regression spline (MARS) NEE in this study.

5.2.2 Proxies

Dai Global Palmer Drought Severity Index

The scPDSI (Dai, 2017) provides a measure of drought severity on a $2.5^\circ \times 2.5^\circ$ grid. The scPDSI is computed using observed monthly surface air temperature and precipitation and provides a measure of surface aridity anomalies and changes on seasonal to longer time scales (Dai et al., 2004; Dai, 2011). We note that scPDSI may not be a good proxy of soil moisture content over the high latitudes ($>50^\circ$).

SIF

We use the monthly gridded “SIF daily average” product from the NASA Level 3 GOME-2 version 27 (V27) terrestrial chlorophyll fluorescence data (NASA-SIF, 2016; Joiner et al., 2013, 2016). SIF anomalies are multiplied by negative one to change the sign of the anomalies so that reduced SIF will be correlated with positive NEE anomalies (emission of CO_2 to the atmosphere).

Soil temperature

For the soil temperature proxy, we use soil temperatures from the MERRA-2 (Reichle et al., 2011, 2017) reanalysis. Specifically, we use the average soil temperature over levels 1–3 (TSOIL1, TSOIL2, and TSOIL3), which reaches a depth of 0.73 m.

NINO 3.4 index

For the phase of ENSO, we use the sea surface temperature (SST) anomaly in the NINO 3.4 region (5°S – 5°N , 120°S – 170°N) of the tropical Pacific Ocean. This region has been widely used to diagnose ENSO activity. The SST values are calculated from the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST) dataset. The SST anomalies were downloaded from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) website (<https://www.esrl.noaa.gov>).

5.2.3 Inversion analyses

CarbonTracker

We use optimized NEE from the NOAA’s CarbonTracker, version CT2016 (Peters et al., 2007, with updates documented at <http://carbontracker.noaa.gov>). CT2016 optimizes NEE by assimilating in situ observations of boundary layer atmospheric CO_2 . It employs the ensemble Kalman filter approach to assimilate CO_2 with atmospheric chemical transport simulated by the TM5 offline atmospheric model (Krol et al., 2005). For CT2016, TM5 is driven by ERA-Interim assimilated meteorology from the ECMWF, with a horizontal resolution of $3^{\circ} \times 2^{\circ}$ globally and $1^{\circ} \times 1^{\circ}$ in a nested grid over North America. CT2016 also has IAV in biomass burning. Therefore, when analyzing posterior IAV in CT2016 we examine the IAV in NEE alone (referred to as CT2016) and IAV due to NEE and biomass burning combined (referred to as CT2016_{w/BB}).

GOSAT level 4 data

We use the GOSAT level 4 data product (Maksyutov et al., 2013) produced by the National Institute for Environmental Studies (NIES). This product is produced by assimilating NIES Level 2 retrievals of XCO_2 into the NIES global atmospheric tracer transport model (NIES-TM) to optimize monthly CO_2 fluxes for 64 sub-continental regions. The Vegetation Integrative Simulator for Trace gases (VISIT), a prognostic biosphere model (Ito, 2010; Saito et al., 2011), is used to generate prior biospheric fluxes for the inversion

Table 5.1: Set-up of GEOS-Chem flux inversions. Differences are in model transport resolution, prior fluxes, and prior errors.

Name	Resolution	Prior flux error	Prior flux IAV
GC _{2×2.5-200%}	2° × 2.5°	200%	No (mean 2010–2013)
GC _{2×2.5-66%}	2° × 2.5°	66%	No (mean 2010–2013)
GC _{4×5-100%}	4° × 5°	100%	No (mean 2010–2013)
GC _{4×5-44%}	4° × 5°	44%	No (mean 2010–2013)
GC _{4×5-100%-IAV}	4° × 5°	100%	Yes
GC _{4×5-44%-IAV}	4° × 5°	44%	Yes

analyses. The GOSAT L4 product also has IAV in biomass burning. Therefore, when analyzing posterior IAV, we examine IAV in NEE alone (referred to as GOSAT L4) and IAV due to NEE and biomass burning combined (referred to as GOSAT L4_{w/BB}).

GEOS-Chem

The GEOS-Chem flux inversions performed in this study are shown in Table 5.1. The flux inversions are performed with different model configurations to examine the sensitivity of the results to the inversion set-up. We perform inversions at two spatial resolutions, 2° × 2.5° and 4° × 5°. The spatial resolution is varied to examine whether changes in model transport significantly impact our results. It has previously been shown that there are significant differences in tracer transport as model resolution is decreased in GEOS-Chem (Yu et al., 2018; Stanevich, 2018a,b). In particular, Stanevich (2018a) show that resolution-induced biases of up to 30% can be introduced on the scale of TransCom regions for 4° × 5° relative to 2° × 2.5° for atmospheric methane (CH₄) inversions.

The prior error statistics are varied between inversions. The prior error covariance provides a metric of the uncertainty in the prior fluxes. If prior fluxes are well known then small errors are applied. If they are poorly known then large prior errors are applied and the observations will have a larger impact on the posterior fluxes. However, in general, atmospheric CO₂ observations strongly underconstrain the fluxes and additional regularization considerations are required. To prevent overfitting of assimilated observations, prior flux errors are typically tighter than the true uncertainty in NEE fluxes. Therefore, the motivation for varying prior errors in this study is to examine the sensitivity of the posterior flux anomalies to these prior constraints.

Finally, the prior NEE fluxes are varied between flux inversions. For all GEOS-Chem inversions, the prior NEE fluxes are based on the posterior fluxes from CT2016. CT2016 fluxes are used because they are informed by atmospheric CO₂ observations, and thus provide a seasonal cycle of NEE which is closer to the truth than a TBM forward run (see

Chapter 4). Using prior fluxes which are closer to the truth then justifies tighter prior flux error covariances. We use two different set-ups of the CT2016 prior fluxes in the inversions. For four inversions we remove the IAV from the CT2016 fluxes. To do this, the fluxes are averaged over the four years (2010–2013) to generate a mean seasonal cycle. We then repeat this climatology NEE fluxes for each year of the inversion. The reason for removing prior IAV is so that all posterior NEE anomalies will be introduced through the assimilation of GOSAT observations. This set-up is different to many previous flux inversion studies which have included IAV in the prior fluxes. Therefore, to examine the sensitivity of the posterior IAV to prior IAV, we also perform two inversions that employ three-hourly CT2016 NEE fluxes over 2010–2013 unchanged from those available at <http://carbontracker.noaa.gov>, other than spatial interpolation to fit our grid, so that IAV is present on the prior NEE for these inversions. The inversions are given names with a subscript following the convention “model resolution – percentage error applied to prior fluxes – presence of prior IAV”, such that, an inversion analysis at $4^\circ \times 5^\circ$ resolution with 100% uncertainty applied to prior fluxes and with prior IAV is named “GC_{4×5–100%–IAV}.”

For GOSAT observations, we use version v3.5 of the NASA ACOS GOSAT lite files. We selected all measurements from the TANSO-FTS SWIR channel, including ocean glint, high gain and medium gain nadir, which pass the quality flag requirement and have warn levels less than or equal to 15. We generate “super-obs” from the GOSAT retrievals by aggregating the observations to the grid size of our inversion. We generate error estimates using the method described by Kulawik et al. (2016). The reduction in error with aggregation can be calculated using the expression $error^2 = a^2 + b^2/n$, where a represents systematic errors that do not decrease with averaging, b represents random errors that decrease with averaging, and n represents the number of satellite observations that are averaged (Kulawik et al., 2016). Kulawik et al. (2016) give $a = 0.8$ ppm and $b = 1.6$ ppm as mean Northern Hemisphere geometric (co-located) values for GOSAT, and these are the values that we use.

Observing system simulation experiments

Five OSSEs are performed, for which pseudo-data are generated by simulating atmospheric CO₂ with GEOS-Chem at $4^\circ \times 5^\circ$ spatial resolution and with year specific NEE from JULES. The GEOS-Chem CO₂ distribution is sampled according to the GOSAT observational coverage. We generate pseudo XCO₂ using the GOSAT averaging kernel weighting and apply random errors to the XCO₂ pseudo-obs consistent with the error estimates described in Sec. 5.2.3. The inversion configuration for three of the OSSEs is identical to GC_{4×5–44%}, GC_{4×5–200%}, and GC_{4x5–100%–IAV}, which use the pos-

terior CT2016 fluxes as their prior NEE (see Table 5.1). These OSSEs are referred to as $\text{OSSE}_{\text{CT2016-44\%}}$, $\text{OSSE}_{\text{CT2016-100\%}}$, and $\text{OSSE}_{4\times5-100\%-\text{IAV}}$, respectively. Two more OSSEs use the same set up as $\text{GC}_{4\times5-44\%}$ and $\text{GC}_{4\times5-200\%}$, except that for these we use the 2010-2013 mean NEE fluxes from JULES as the prior fluxes. These two OSSEs are referred to as $\text{OSSE}_{\text{JULES-44\%}}$ and $\text{OSSE}_{\text{JULES-100\%}}$.

5.2.4 Terrestrial biosphere models

JULES

JULES is a community land surface model that has evolved from the UK Met Office Surface Exchange Scheme. Phenology in JULES affects leaf growth rates and timing of leaf growth/senescence based on temperature alone (Clark et al., 2011). Vegetation cover is predicted based on nine plant functional types that compete for space based on their relative productivity and height but are excluded from growing on agricultural land, based on a fraction of agriculture in each grid cell (Harper et al., 2018). CRU-NCEP was used as model forcing data.

VISIT

VISIT is a prognostic biosphere model (Ito, 2010; Saito et al., 2011) that simulates carbon exchanges between the atmosphere and biosphere and among the carbon pools within terrestrial ecosystems at a daily time step. Modeling of plant CO_2 assimilation in VISIT is based on a model of light extinction in the canopy, following the formulation of Monsi and Saeki (1953). Autotrophic respiration is formulated as the sum of growth respiration and maintenance respiration. Growth respiration is simulated as the cost to produce new biomass, while maintenance respiration is represented as a function of ground surface temperature. Heterotrophic respiration is the sum of respiration from two layers, litter and humus, which is regulated by soil temperature and soil moisture at each depth. VISIT is driven by reanalysis/assimilation climate datasets provided by the Japan Meteorological Agency (JMA): the Japan 25-year reanalysis (JRA-25)/JMA Climate Data Assimilation System JCDAS) (Onogi et al., 2007) for the period 1979–present.

Carnegie-Ames-Stanford Approach (CASA) Global Fire Emissions Database (GFED) Carbon Monitoring System (CMS) model

The version of the model used here, referred to as CASA CMS, was modified from Potter et al. (1993) as described in Randerson et al. (1996) and van der Werf et al. (2006). It is driven by MERRA reanalysis and satellite-observed Normalized Difference Vegetation Index (NDVI) to track plant phenology. These flux estimates were computed at monthly time steps with 0.5° spatial resolution.

CASA GFED 4.1

The version of the model used here, CASA GFED 4.1, was modified from Potter et al. (1993) as described in van der Werf et al. (2017). It is driven by ECMWF reanalysis and satellite-observed NDVI to track plant phenology. These flux estimates were computed at monthly time steps with 0.25° spatial resolution.

5.2.5 Anomalies and correlations

Monthly anomalies are calculated by subtracting the mean 2010-2013 value for a given month from the monthly value for a specific year. For example, the NEE anomaly for a given month and year is calculated using:

$$ANOM[year, month] = NEE[year, month] - \frac{1}{4} \sum_{i=2010}^{2013} NEE[i, month]. \quad (5.1)$$

Anomalies are calculated over a range of spatial scales. In each case, the quantity of interest is first averaged into a spatial mean for each month, then anomalies are calculated. The same procedure is followed for JJA anomalies except that the anomaly is calculated over the entire three-month period instead of for a single month.

In the tropics, temporal correlations are performed to quantify agreement between anomalies between NEE anomalies and proxy/FLUXCOM anomalies. We choose a null hypothesis in which the correlation is zero. This is the correct null hypothesis for flux inversions for which the prior NEE fluxes have no IAV. In flux inversions for which there is IAV in the prior, the correlation between the proxies and prior IAV should be used as the null hypothesis. However, this would be a significantly more difficult null hypothesis to test, so for simplicity we choose a null hypothesis of zero correlation for all cases. The threshold for rejection of the null hypothesis (α) is chosen to be 0.05, such that the null hypothesis is rejected if the P-value (P) is less than 0.05. We acknowledge that this α

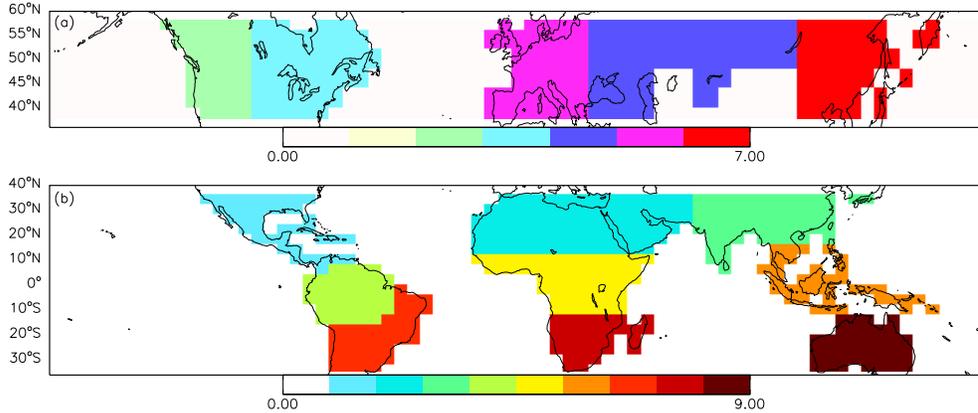


Figure 5.1: Sub-continental regions in (a) the extratropics and (b) the tropics. In the tropics, we generate three continents by combining the regions in the Americas, Africa and the Middle East, and the Asia-Pacific and Indian sub-continent.

threshold is largely arbitrary but is widely used in the literature (Benjamin et al., 2018; Lakens et al., 2018). Throughout the manuscript, correlations that satisfy this criterion are called “strong”. In most cases a second test is performed, in which we test if the correlation between the flux inversion IAV and the proxy is greater than that between the NINO 3.4 index and the proxy, and conclude that the inversion and proxy only show good agreement if both of these thresholds are met. The coefficient of correlation is referred to as R .

We also perform a series of linear regressions. In the tropics, linear regressions are performed after aggregating over all tropical land, such that the regression is performed on a single 48 point time series. In the northern extratropics, linear regressions are performed for the set of four JJA anomalies across five sub-continental regions resulting in a 20 point dataset. For all regressions the y-intercept is close to zero, and thus is not reported. The slope of the regressions and coefficient of determination (R^2) are reported.

5.3 Results

5.3.1 Tropics

Monthly anomalies in the tropics are examined over a range of spatial scales. The anomalies are aggregated to $4^\circ \times 5^\circ$, $8^\circ \times 10^\circ$, sub-continental regions (shown in Fig. 5.1), continents, and the entire tropics. Figure 5.2 shows the mean correlation coefficient between the inversions/TBMs and proxies/FLUXCOM on scales ranging from $4^\circ \times 5^\circ$ grid cells to the entire tropics. Correlations between the NINO 3.4 index and flux proxies are also

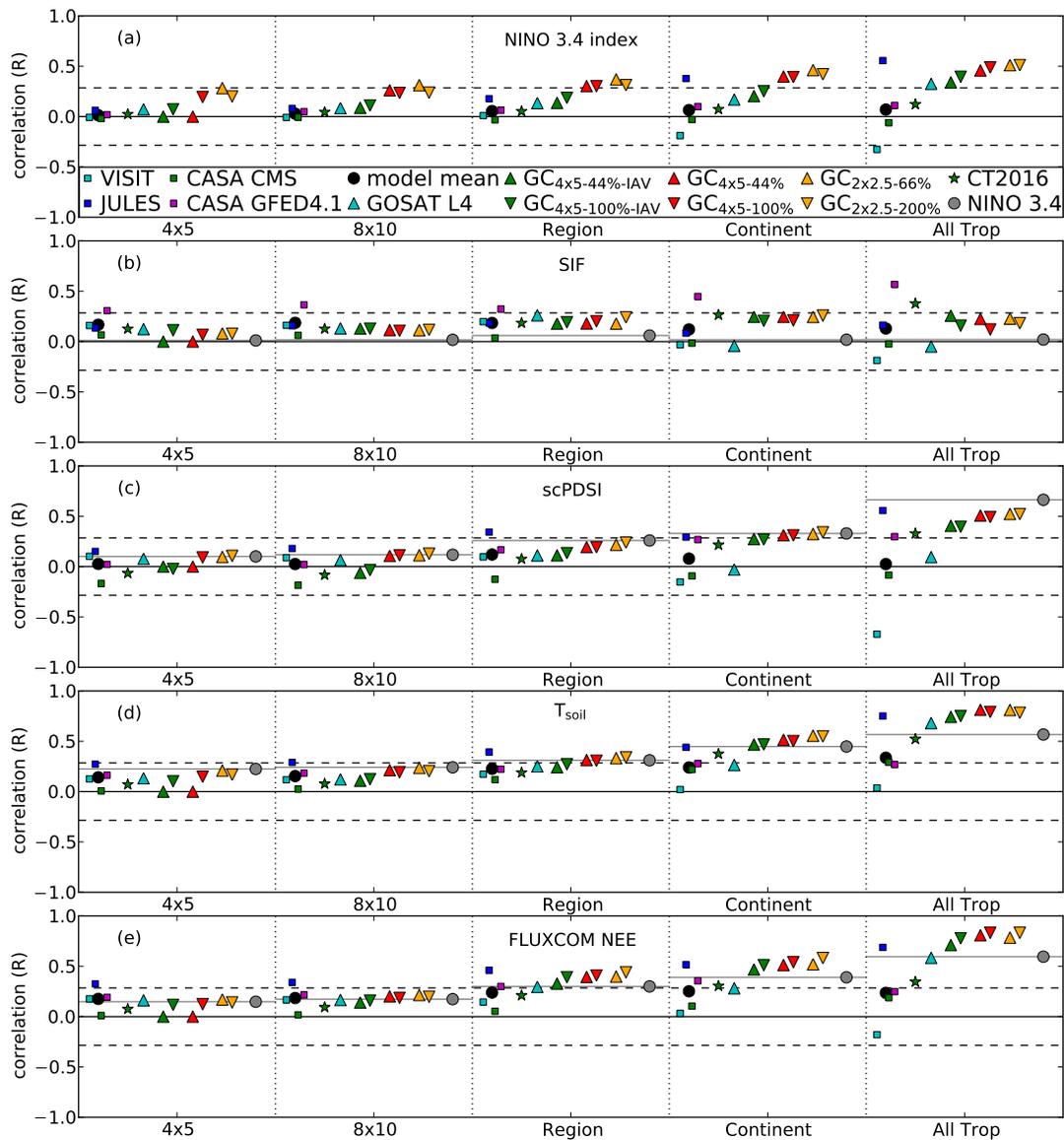


Figure 5.2: Correlation over a range of scales for models and inversions with (top) NINO 3.4 index, (second) $(-1) \times \text{SIF}$, (third) scPDSI, (fourth) T_{soil} , and (bottom) FLUXCOM NEE in the tropics. Squares represent terrestrial ecosystem models: VISIT (cyan), JULES (blue), CASA GFED CMS (green), CASA GFED 4.1 (magenta) and the black circle shows the mean correlation of the models. GOSAT flux inversions are represented by: GOSAT L4 (cyan up-triangle), GC_{4x5-44%-IAV} (green up-triangle), GC_{4x5-100%-IAV} (green down-triangle), GC_{4x5-44%} (red up-triangle), GC_{4x5-100%} (red down-triangle), GC_{2x2.5-66%} (orange up-triangle), and GC_{2x2.5-200%} (orange down-triangle). The green star shows CT2016. The gray circle and line show the correlation with the NINO 3.4 index. Dashed black lines indicate the correlation required for an α of 0.05, therefore, all correlations greater than the dashed black line indicate $P < 0.05$.

shown over the range of spatial scales. It is important to consider correlations between the inversions/TBMs and proxies/FLUXCOM with the influence of ENSO variability in mind, as ENSO is the primary driver of large-scale IAV in the tropics. Therefore, to understand how well the flux inversions are capturing IAV it is useful to contrast correlations between the inversion and proxies to correlations between the NINO 3.4 index and proxies.

The correlation between posterior NEE anomalies and proxy/FLUXCOM anomalies increase with aggregation (Fig. 5.2). This is expected as atmospheric CO₂ observations are expected to best constrain fluxes on large scales, such as the entire tropics. As scales decrease, the signal from variations in the fluxes become weaker and more difficult to constrain with the atmospheric CO₂ observations. Correlations between the proxies and the NINO 3.4 index also increase with aggregation. This is because the NINO 3.4 index reflects the large-scale ENSO-driven variability in the tropics. Therefore, increasing correlation with aggregation for the NINO 3.4 index is a reflection of the large-scale variability having a larger impact.

To categorize the agreement between the flux inversions and the proxies/FLUXCOM, we state that the flux inversions agrees with a proxy on a given scale only if the correlation is strong ($P < 0.05$) and greater than the correlation of the proxy/FLUXCOM with the NINO 3.4 index ($R > R_{\text{NINO3.4}}$). For the GEOS-Chem inversions, this occurs at regional and larger scales for correlations with FLUXCOM NEE and at continental and larger scales for T_{soil} . For the GOSAT L4 inversion, the correlation only reaches this threshold for T_{soil} at the largest aggregation scale. These results suggest that GOSAT observations provide flux information on continental and larger scales, while regional-scale constraints may be possible. The fact that the correlation coefficient is variable between GOSAT inversions indicates that the agreement between posterior fluxes and the proxies//FLUXCOM is sensitive to the inversion configuration.

We investigate the influence of the inversion configuration by comparing the correlations for the six GEOS-Chem inversions. The $2^\circ \times 2.5^\circ$ inversions generally show slightly better agreement with the proxies/FLUXCOM than the $4^\circ \times 5^\circ$ inversions at regional and continental scales. This could be due to improved transport with higher spatial resolution, however, other aspects of the inversion were changed such as the aggregation of assimilated observations and prior error covariances, which may have also influenced the results. The influence of prior IAV can be evaluated by comparing the $4^\circ \times 5^\circ$ inversions with and without prior IAV. Correlations are stronger for the inversions without IAV at regional and continental scales. This suggests that prior IAV degrades the inversion and is discussed in more detail in Sec. 5.4.3. The influence of prior error covariances can be

evaluated by comparing the inversions with small (44% for $4^\circ \times 5^\circ$ and 66% for $2^\circ \times 2.5^\circ$) and large (100% for $4^\circ \times 5^\circ$ and 200% for $2^\circ \times 2.5^\circ$) prior error. Larger prior errors generally result in larger correlations on regional and larger scales. Large prior errors means that more movement away from the prior during the inversion is allowed, therefore, better agreement with larger prior errors suggests that the GOSAT data information content is sufficiently large that loose prior errors can be applied without degrading the posterior results by over fitting the observations.

For CT2016, strong correlations that are greater than those for the NINO 3.4 index are only obtained for SIF and only on the scale of the entire tropics. However, given the poor correlation between SIF and the NINO 3.4 index, it is possible that this is a spurious correlation. CT2016 is also the only inversion that does not show strong correlations with the NINO 3.4 index when aggregated to the entire tropics. The poorer agreement between CT2016 and the proxies/FLUXCOM than for GOSAT inversions suggests that the network of surface observations does not provide sufficient information to constrain tropical fluxes. However, it is also possible that the inversion set-up could play a role.

For the TBMs, correlations are highly model dependent. Of the models, JULES shows the best agreement with the proxies/FLUXCOM. JULES shows strong correlations greater than for the NINO 3.4 index at all scales for FLUXCOM NEE, regionally and over the entire tropics for T_{soil} , and regionally for scPDSI. These results suggest that JULES predicts NEE anomalies in the tropics as well as the GOSAT inversion on continental and larger scales, and may be better at regional and smaller scales. This suggests that it may be challenging to use GOSAT flux inversions to evaluate IAV in JULES NEE. For the other models, less agreement is seen with the proxies/FLUXCOM. The one exception is CASA GFED 4.1 which shows strong correlations with SIF at all scales. This may be due to the fact that this model assimilates greenness indices to estimate GPP fluxes. Anomalies in the greenness indices are likely well correlated with SIF anomalies, therefore, if anomalies in CASA NEE are driven by anomalies in GPP, it may explain the strong correlation.

We now investigate the magnitude of tropical IAV in the inversions and the TBMs. The magnitude of IAV relative to the proxies/FLUXCOM can be obtained by performing linear regressions of the anomalies. We limit our analysis to the scale of the entire tropics, where the inversions and proxies/FLUXCOM agreed best. Table 5.2 shows the slope and coefficient of determination (R^2) for the regressions. There is a large amount of variability in the slopes between inversions/TBMs for each proxy/FLUXCOM. The GOSAT inversions are quite consistent with each other relative to CT2016 and the TBMs. The GOSAT inversions give slopes of 1.03–2.10 for FLUXCOM and 0.061–0.12 for T_{soil} .

Table 5.2: Slope and coefficient of determination (R^2) for linear regressions for anomalies across the entire tropics.

Model/Inversion gC m ⁻² day ⁻¹	FLUXCOM NEE gC m ⁻² day ⁻¹		SIF mW m ⁻² nm ⁻¹ sr ⁻¹		scPDSI		T_{soil} (K)	
	slope	R^2	slope	R^2	slope	R^2	slope	R^2
GC _{2×2.5-200%}	1.87	0.69	0.90	0.03	0.078	0.27	0.100	0.61
GC _{2×2.5-66%}	1.03	0.62	0.65	0.05	0.045	0.27	0.061	0.66
GC _{4×5-100%}	1.70	0.69	0.54	0.01	0.067	0.24	0.093	0.63
GC _{4×5-44%}	1.06	0.65	0.65	0.05	0.044	0.26	0.061	0.66
GC _{4×5-100%-IAV}	2.10	0.61	0.94	0.03	0.071	0.16	0.12	0.56
GC _{4×5-44%-IAV}	1.57	0.51	0.03	0.00	0.06	0.16	0.087	0.55
GOSAT L4	1.59	0.34	-0.30	0.00	0.017	0.01	0.106	0.46
GOSAT L4 _{w/BB}	1.69	0.33	-0.02	0.00	0.007	0.00	0.107	0.40
CT2016	0.66	0.12	1.58	0.14	0.042	0.11	0.057	0.27
CT2016 _{w/BB}	0.79	0.14	1.73	0.14	0.027	0.04	0.059	0.24
VISIT	-0.50	0.03	-1.15	0.04	-0.13	0.45	0.006	0.00
CASA 4.1	0.38	0.06	1.88	0.32	0.030	0.09	0.023	0.07
CASA CMS	0.33	0.04	-0.09	0.00	-0.010	0.01	0.029	0.08
JULES	1.85	0.47	0.96	0.027	0.10	0.31	0.116	0.56

Comparing the GEOS-Chem inversions, the largest differences in the slopes are related to the magnitude of the prior error covariances. Looser prior constraints result in slopes that are 30–80% larger. There are also large differences in the magnitude of posterior IAV between the inversions with and without prior IAV. For example, the slopes for the regression between FLUXCOM and the 4° × 5° GEOS-Chem inversions with prior anomalies are 25–50% larger than for GEOS-Chem inversions without prior IAV. The GOSAT L4 product gives slopes which are consistent with the GEOS-Chem inversions. Furthermore, the agreement between the GOSAT L4 product and proxies (or FLUXCOM) is not sensitive to the inclusion of biomass burning. For CT2016, the best agreement is found with T_{soil} ($0.24 \leq R^2 \leq 0.27$), for which CT2016 gives a smaller slope than the GOSAT inversions. The agreement between CT2016 and proxies/FLUXCOM is not sensitive to the inclusion of biomass burning. For the TBMs, JULES gives good fits with T_{soil} ($R^2 = 0.56$) and FLUXCOM ($R^2 = 0.47$) and gives slopes that are similar in magnitude to the flux inversions. The rest of the TBMs have R^2 that are too small to make meaningful comparisons.

Detailed analysis of GC_{2×2.5-200%}

We examine the agreement between the GC_{2×2.5-200%} inversion and the proxies/FLUXCOM in the tropics in more detail. Figure 5.3 (left column) shows the correlation coefficient for

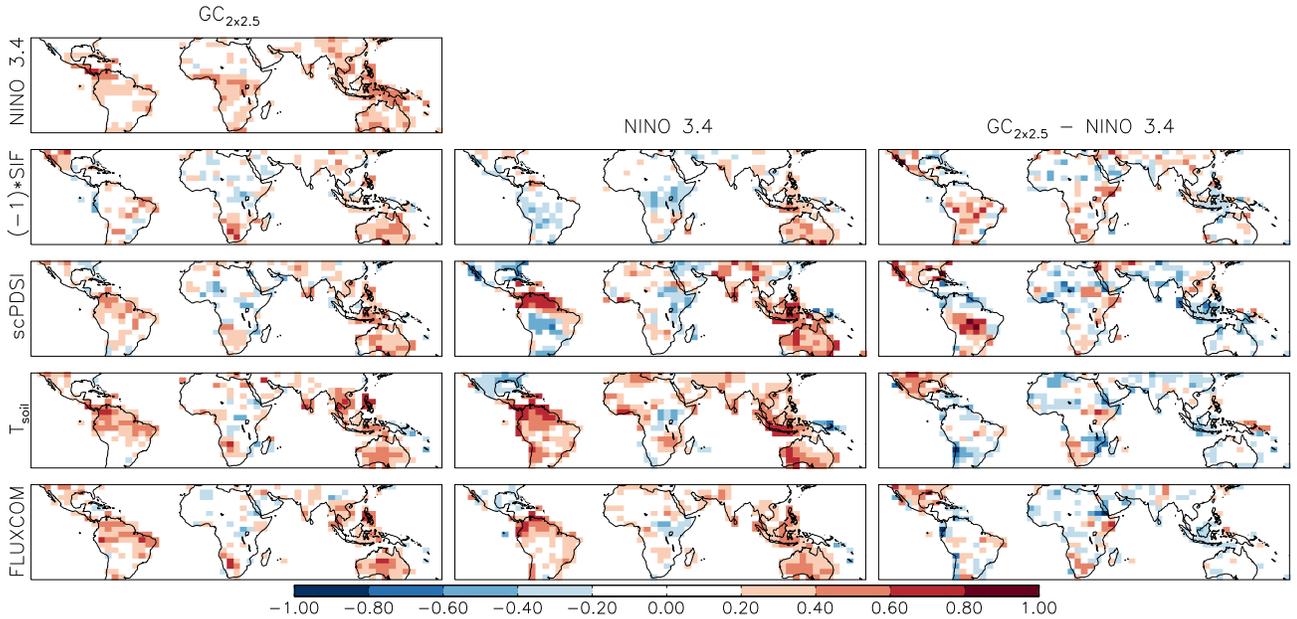


Figure 5.3: Correlations of monthly anomalies over tropical land at $4^\circ \times 5^\circ$ spatial resolution. Columns show coefficient of correlation (R) of (left) $GC_{2 \times 2.5-200\%}$, (center) NINO 3.4 index, and (right) the difference between the two with (top row) the NINO 3.4 index, (second row) $(-1) \times SIF$, (third row) $scPDSI$, (fourth row) T_{soil} , and (bottom row) FLUXCOM NEE.

each grid cell between the $GC_{2 \times 2.5-200\%}$ anomalies and the proxy/FLUXCOM anomalies. There are broad positive correlations with the NINO 3.4 index across Central and South America, tropical and southern Africa, and much of the Asia-Pacific. Generally, positive correlations are present between $GC_{2 \times 2.5-200\%}$ and SIF, $scPDSI$, T_{soil} , and FLUXCOM NEE in the Americas, southern Africa, and the Asia-Pacific. Figure 5.3 (center column) shows the correlation coefficient between the NINO 3.4 index and the proxies over the tropics. Generally, the proxies show strong correlations with the NINO 3.4 index in many of the same regions for which these proxies show strong correlations with $GC_{2 \times 2.5-200\%}$. This suggests that grid-scale correlations between $GC_{2 \times 2.5-200\%}$ and the proxies may be a reflection of the large-scale anomalies across the tropics and do not necessarily imply that the inversion is able to isolate the spatial footprint of ENSO-driven flux anomalies on smaller scales.

We examine whether $GC_{2 \times 2.5-200\%}$ is able to isolate flux anomalies that are separate from the large-scale tropical signal by comparing NEE anomalies for FLUXCOM NEE and $GC_{2 \times 2.5-200\%}$ as a function of time. First, we aggregate $GC_{2 \times 2.5-200\%}$ and FLUXCOM NEE to the entire tropics and the following continental-scale regions: the Americas, Africa plus the Middle East, and the Asia-Pacific plus the Indian sub-continent

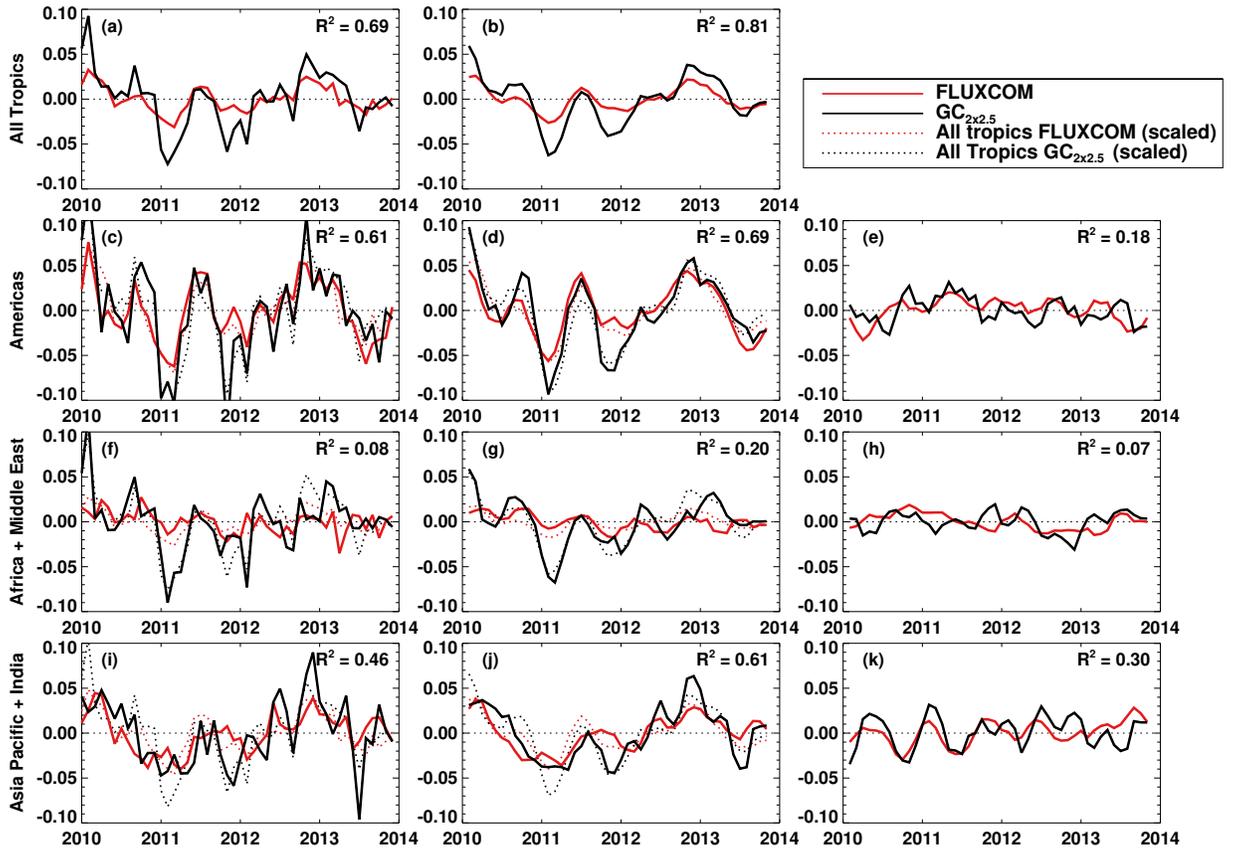


Figure 5.4: NEE anomalies ($\text{gC m}^{-2} \text{ day}^{-1}$) for FLUXCOM and $GC_{2 \times 2.5-200\%}$ in the tropics. (left column) Monthly anomalies, (center column) smoothed (3-month running mean) monthly anomalies, and (right column) continental anomalies minus the scaled mean tropical anomalies for (a–b) the entire tropics, (c–e) the Americas, (f–h) Africa and the Middle East, and (i–k) the Asia Pacific and Indian sub-continent. For each panel, R^2 shows the coefficient of determination between $GC_{2 \times 2.5-200\%}$ and FLUXCOM NEE anomalies within the panel.

(Fig. 5.1). Figure 5.4 shows $GC_{2 \times 2.5-200\%}$ and FLUXCOM NEE anomalies as a function of time over the entire tropics and the continental-scale regions. We show raw and smoothed (3-month running mean) monthly NEE anomalies as a function of time. Over the entire tropics, FLUXCOM and $GC_{2 \times 2.5-200\%}$ are highly correlated ($R^2 = 0.69$) (which is shown in Fig. 5.2). On continental scales, the agreement between FLUXCOM and $GC_{2 \times 2.5-200\%}$ is variable, ranging from $R^2 = 0.08$ for Africa plus the Middle East to $R^2 = 0.61$ for the Americas. All correlations improve after smoothing, suggesting that monthly scale variations are not correctly represented in $GC_{2 \times 2.5-200\%}$, FLUXCOM NEE, or both. We attempt to isolate anomalies specific to each continent by removing the large-scale anomaly across the entire tropics. This is done by subtracting a mean tropical anomaly (scaled to have the same variance as the continental anomaly) from the continental anomaly using the following equation:

$$\text{DIFF}_{\text{continent-tropics}} = \text{ANOM}_{\text{continent}} - \text{ANOM}_{\text{tropics}} \times \frac{\text{STD}(\text{ANOM}_{\text{continent}})}{\text{STD}(\text{ANOM}_{\text{tropics}})}, \quad (5.2)$$

where $\text{STD}()$ represents standard deviation and $\text{DIFF}_{\text{continent-tropics}}$ provides an estimate of anomalies in NEE for a given continent that are not associated with the large-scale ENSO-driven anomalies across the tropics. $\text{DIFF}_{\text{continent-tropics}}$ is shown for each continent in Fig. 5.4e,h,k. The magnitude of the anomalies are reduced after removing the tropical mean anomalies. Positive correlations are obtained for the Americas ($R^2 = 0.18$), Africa plus the Middle East ($R^2 = 0.07$), and the Asia Pacific and India ($R^2 = 0.30$). These results suggest that $GC_{2 \times 2.5-200\%}$ is partially able to isolate NEE anomalies on continental scales that are separate from the large-scale ENSO-induced variability, and suggests that GOSAT flux inversions can be used to examine continental-scale flux anomalies in the tropics. We note, however, that the the agreement in IAV between $GC_{2 \times 2.5-200\%}$ and FLUXCOM is not as strong in Africa and the Middle East.

5.3.2 Northern extratropics

In the northern extratropics, the observational coverage of GOSAT is highly seasonal and so we limit our analysis of anomalies in the northern extratropics to the summer (JJA), when observational coverage is the best (Chapter 3). Fig. 5.5 shows the anomalies for the proxies, FLUXCOM, and $GC_{2 \times 2.5-200\%}$ across the northern hemisphere for JJA 2010–2013. The proxies and FLUXCOM generally show high coherence in anomalies. Events for which FLUXCOM NEE gives enhanced emission to the atmosphere also show reduced SIF, increased scPDSI, and increased T_{soil} . We have highlighted (with

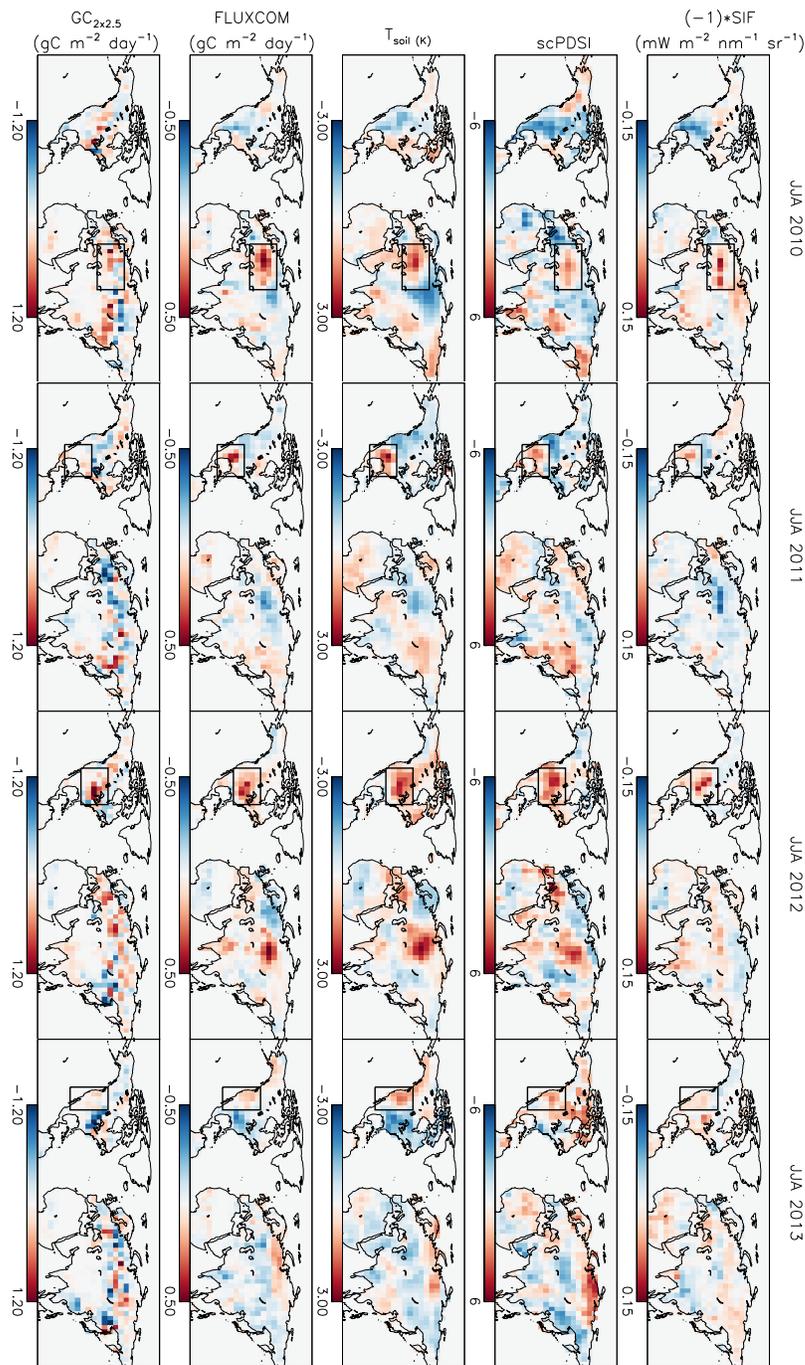


Figure 5.5: Northern extratropical anomalies during JJA. Anomalies for (top row) $(-1) \times \text{SIF}$, (second) scPDSI , (third) T_{soil} , (fourth) FLUXCOM NEE , and (bottom) $\text{GC}_{2x2.5}$ over JJA for (left to right) 2010–2013. Black boxes highlight major climate anomalies: the 2010 Russian heat wave, 2011 drought in Mexico and southern USA, the 2012 North American drought, and the 2013 California drought.

Table 5.3: Slope and coefficient of determination (R^2) for linear regressions for regional anomalies during JJA in the northern extratropics.

Model/Inversion gC m ⁻² day ⁻¹	FLUXCOM NEE gC m ⁻² day ⁻¹		SIF mW m ⁻² nm ⁻¹ sr ⁻¹		scPDSI		T_{soil} (K)	
	slope	R^2	slope	R^2	slope	R^2	slope	R^2
GC _{2×2.5–200%}	1.56	0.54	4.07	0.14	0.052	0.21	0.17	0.56
GC _{2×2.5–66%}	1.28	0.65	3.32	0.16	0.041	0.24	0.13	0.57
GC _{4×5–100%}	1.36	0.49	4.13	0.17	0.054	0.28	0.16	0.62
GC _{4×5–44%}	1.29	0.64	3.36	0.17	0.045	0.29	0.14	0.65
GC _{4×5–100%–IAV}	1.28	0.26	6.8	0.27	0.05	0.16	0.16	0.36
GC _{4×5–44%–IAV}	0.79	0.15	4.66	0.20	0.026	0.06	0.10	0.21
GOSAT L4	1.59	0.33	5.86	0.17	0.086	0.35	0.19	0.43
GOSAT L4 _{w/BB}	1.59	0.34	6.52	0.21	0.090	0.39	0.18	0.39
CT2016	0.21	0.01	4.03	0.13	0.000	0.00	0.04	0.03
CT2016 _{w/BB}	0.18	0.006	4.59	0.16	0.002	0.00	0.03	0.01
VISIT	0.93	0.47	3.25	0.21	0.059	0.67	0.10	0.50
CASA 4.1	0.37	0.12	3.96	0.48	0.020	0.11	0.05	0.20
CASA CMS	0.16	0.01	4.13	0.34	0.00	0.00	0.02	0.02
JULES	1.58	0.29	7.26	0.23	0.075	0.23	0.23	0.52

boxes) major climate anomalies over this time period: the 2010 Russian heat wave, the 2011 drought in Mexico and southern USA, the 2012 North American drought, and the 2013 California drought. GC_{2×2.5–200%} shows some ability to recover all of these major anomalies, suggesting that the inversion can recover sub-continental IAV. However, there are also instances where the inversion seems unable to recover anomalies. For example, in 2010, GC_{2×2.5–200%} indicates a positive anomaly in North America, whereas the proxies indicate near neutral or negative anomalies.

To examine agreement with the proxies on regional scales, we have aggregated the inversions, the TBMs, proxies, and FLUXCOM into the five extratropical subcontinental regions shown in Fig. 5.1. The JJA anomalies in these regions over 2010–2013 provide 20 data points. We performed a linear regression of these anomalies against the proxies and FLUXCOM. Table 5.3 shows the slope and R^2 values of the regressions. For the GOSAT inversions, the 2° × 2.5° and 4° × 5° with no prior IAV show the closest agreement with FLUXCOM NEE and T_{soil} ($0.49 \leq R^2 \leq 0.65$), while the inversions with prior IAV show substantially poorer agreement ($0.15 \leq R^2 \leq 0.36$). This is a larger difference between the inversions with and without prior IAV than was found for the tropics (see Sec. 5.4.3). The inversions with IAV also give a smaller slope indicating a smaller magnitude of IAV, which is the opposite of what was found in the tropics. Comparing the inversions without prior IAV, tight prior errors give $0.57 \leq R^2 \leq 0.65$, whereas loose prior constraints give

$0.49 \leq R^2 \leq 0.62$. This suggests that the inversions with loose prior constraints may be over-fitting the data and degrading the agreement with proxies. As with the tropics, the inversions with looser prior constraints give larger slopes, suggesting larger IAV.

Comparing the other inversions, the GOSAT L4 product shows agreement with FLUXCOM NEE ($R^2 = 0.33$) and T_{soil} ($R^2 = 0.43$). CT2016 shows poor agreement with all proxies, indicating that this inversion is unable to isolate zonally asymmetric fluxes in the northern extratropics, which is surprising given the high sensitivity of the surface CO_2 network to northern extratropical surface fluxes (Chapter 3). However, consistent with this result, Polavarapu et al. (2018) show that flux inversions assimilating observations from the surface network are largely unable to recover zonally asymmetric flux signals. CT2016 also includes prior IAV in the inversion which may negatively impact the posterior IAV, based on the GEOS-Chem inversion results.

For the TBMs, VISIT shows close agreement with FLUXCOM NEE, scPDSI, and T_{soil} anomalies and to a lesser extent SIF anomalies. This is notable as VISIT generally showed poor agreement with the proxies in the tropics. JULES shows close agreement with T_{soil} anomalies and some agreement with the other proxies. CASA GFED 4.1 shows good agreement with SIF anomalies, but comparatively poorer agreement with the other proxies. CASA GFED CMS shows some agreement with SIF anomalies, but little agreement with the other proxies.

5.3.3 Observing system simulation experiments

Tropics

Figure 5.6 shows the mean correlation coefficient between the posterior and true anomalies in the tropics over a range of scales. The results are highly reminiscent of the results between the GOSAT inversion and the proxies. The mean correlation between the posterior and true NEE anomalies increases with aggregation for all OSSEs. Strong correlations are obtained for all OSSEs on regional and larger scales. The inversion set up also has an impact on the correlations between the posterior and true IAV. The largest differences between OSSEs are obtained on regional and continental scales. On these scales, $\text{OSSE}_{\text{JULES-100\%}}$ has the largest correlation. This suggests that having a climatological seasonal cycle close to the truth is important for recovering IAV in the tropics. The inclusion of prior IAV ($\text{OSSE}_{4 \times 5-100\%-IAV}$) does not appear to significantly degrade the correlation relative to a prior without IAV ($\text{OSSE}_{\text{CT2016-100\%}}$), in contrast to what was found with the real data GOSAT inversions. The prior error constraints generally have a large influence on the correlation with the truth. Loose prior constraints give

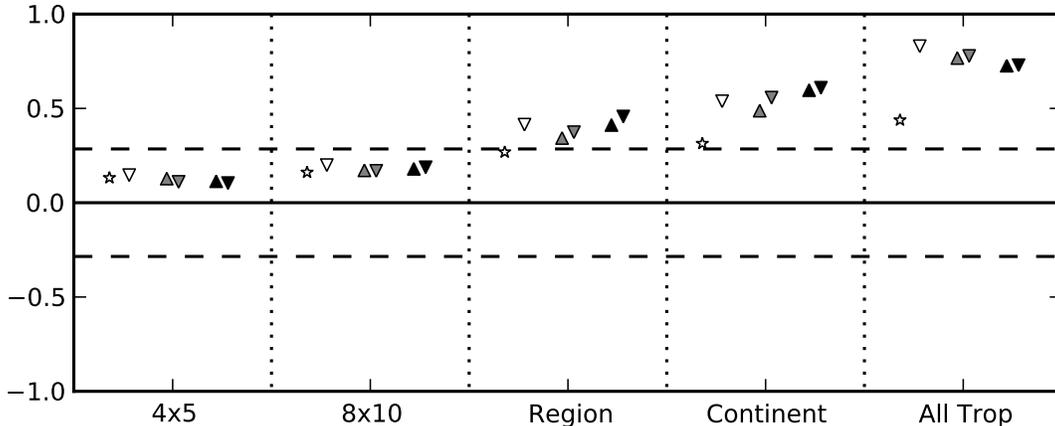


Figure 5.6: Mean correlation coefficient (R) with the truth over a range of spatial scales for CT2016 (white star), $\text{OSSE}_{4\times 5-100\%-\text{IAV}}$ (white down-triangle), $\text{OSSE}_{\text{CT2016}-44\%}$ (grey up-triangle), $\text{OSSE}_{\text{CT2016}-100\%}$ (grey down-triangle), $\text{OSSE}_{\text{JULES}-44\%}$ (black up-triangle), and $\text{OSSE}_{\text{JULES}-100\%}$ (black down-triangle).

better agreement for all OSSEs, consistent with the GOSAT inversions.

On the scale of the entire tropics, we performed linear regressions between the posterior and true anomalies, which are shown in Table 5.4. For all regressions, the magnitude of IAV in the posterior fluxes is less than the truth (slope of 0.42–0.75). This suggests that the inversions do not recover the full magnitude of NEE IAV. In addition to comparing posterior and true anomalies, we examine the similarities in posterior anomalies between OSSEs. The right column of Table 5.4 shows the results of linear regressions between posterior and $\text{OSSE}_{\text{JULES}-100\%}$ NEE anomalies. The OSSEs without prior IAV show better agreement with $\text{OSSE}_{\text{JULES}-100\%}$ posterior anomalies than the true anomalies. This suggests that the assimilation of pseudo-data is introducing NEE anomalies in a similar way for all OSSEs and recovering the true IAV is primarily limited by the observational coverage rather than the inversion set up. However, differences between the OSSEs and truth may also be due to systematic biases introduced due to factors such as uneven observational coverage (Chapter 3).

We examine the continental-scale anomalies in detail for $\text{OSSE}_{\text{JULES}-100\%}$, $\text{OSSE}_{\text{CT2016}-100\%}$, and $\text{OSSE}_{\text{IAV}-100\%}$ in Figure 5.7, which shows the timeseries of continental-scale flux anomalies in the tropics for the OSSEs. The correlation between the OSSEs and true anomalies improves after performing a three-month running mean, consistent with the GOSAT inversion results. Strong correlations between the OSSEs and truth are obtained after removing the mean tropical signal (using Eq. 5.2). These results provide further evidence that GOSAT inversions can largely recover continental-scale flux anomalies in the tropics.

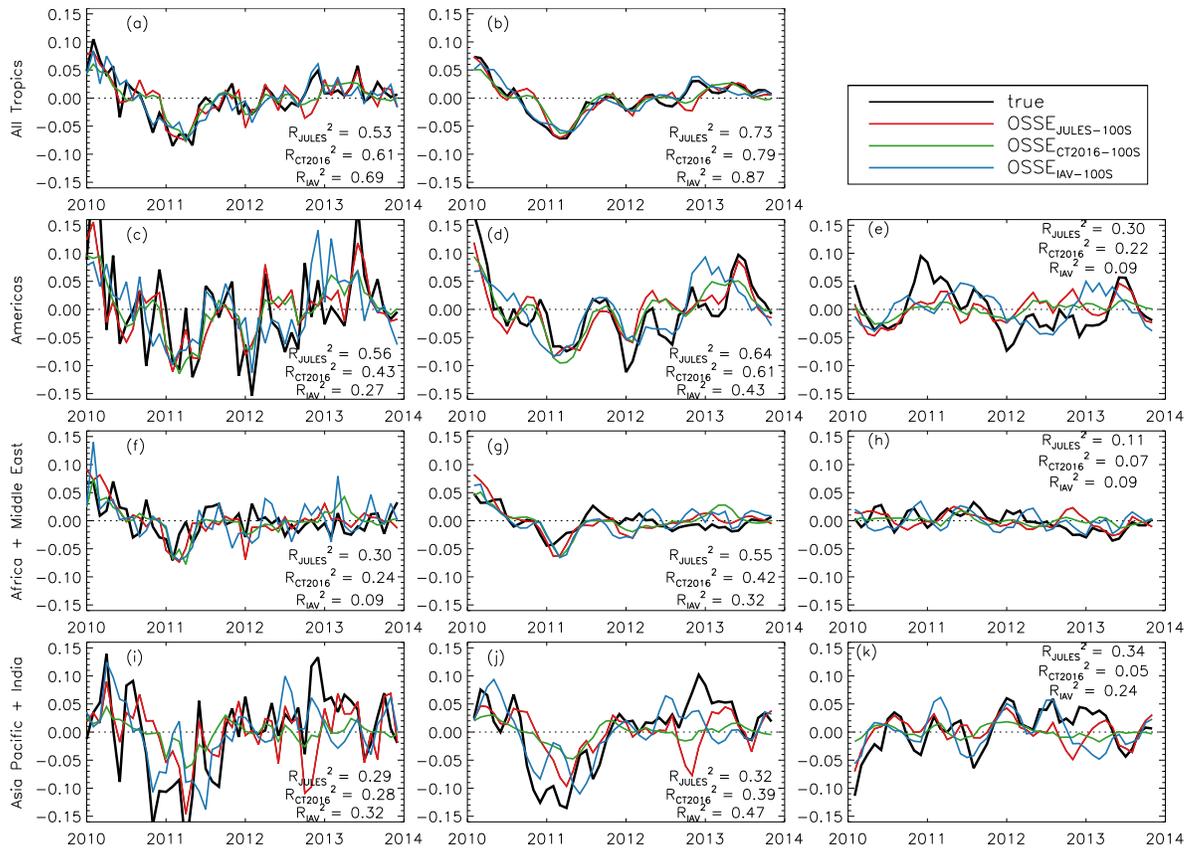


Figure 5.7: Monthly NEE anomalies (gC m⁻² day⁻¹) for OSSE_{JULES-100%} (red), OSSE_{CT2016-100%} (green), OSSE_{I4V-100%} (blue) and truth (black) in the tropics. (left column) Monthly anomalies, (center column) smoothed (3-month running mean) monthly anomalies, and (right column) continental anomalies minus the scaled mean tropical anomalies for (a–b) the entire tropics, (c–e) the Americas, (f–h) Africa and the Middle East, and (i–k) the Asia Pacific and Indian sub-continent.

Table 5.4: Slope and coefficient of determination (R^2) for linear regressions for OSSE experiments.

Tropics				
Inversion	Truth		OSSE _{JULES-100%}	
	slope	R^2	slope	R^2
OSSE _{JULES-100%}	0.67	0.53		
OSSE _{JULES-44%}	0.58	0.53	0.91	0.91
OSSE _{CT2016-100%}	0.55	0.61	0.84	0.84
OSSE _{CT2016-44%}	0.42	0.59	0.69	0.77
OSSE _{4×5-100%-IAV}	0.75	0.69	0.70	0.48
CT2016	0.31	0.19	0.50	0.15
Northern Extratropics				
Inversion	Truth		OSSE _{JULES-100%}	
	slope	R^2	slope	R^2
OSSE _{JULES-100%}	0.35	0.39		
OSSE _{JULES-44%}	0.27	0.48	0.76	0.80
OSSE _{CT2016-100%}	0.30	0.30	1.04	0.88
OSSE _{CT2016-44%}	0.31	0.43	1.06	0.62
OSSE _{4×5-100%-IAV}	0.63	0.15	0.55	0.41
CT2016	0.48	0.46	0.18	0.05

Northern extratropics

Table 5.4 shows the slope and R^2 for linear regressions of flux anomalies from the OSSEs against the truth on sub-continental regions in the northern extratropics during JJA. In all cases the slope is less than one, indicating that the OSSEs are not recovering the full magnitude of IAV. The R^2 values are less than between the GOSAT inversions and proxies. This may be due to the fact that temporal anomalies in JULES NEE are highly variable month-to-month and may have a shorter temporal correlation length scales than the true anomalies. Comparing the different OSSE set ups, the OSSE_{4×5-100%-IAV} performs substantially worse than the OSSEs with no prior IAV ($R^2 = 0.15$ versus $R^2 = 0.30$ – 0.48). This is consistent with comparisons between GOSAT inversions and proxies, and suggests that employing prior IAV in the northern extratropics degrades posterior IAV on sub-continental scales during JJA. OSSEs with tighter prior constraints give larger R^2 , consistent with the GOSAT inversions. OSSEs with JULES mean seasonal cycle also agree better with the truth than those which employ the mean seasonal cycle from CT2016.

5.4 Discussion

5.4.1 Implications of correlations between flux inversions and proxies

The results of this study show varying degrees of agreement between anomalies in GOSAT flux inversions and anomalies in proxies and FLUXCOM. We consistently find that T_{soil} and FLUXCOM NEE show the strongest agreement with the flux inversions, whereas scPDSI and SIF show weaker agreement. In this section we discuss agreement between the proxies and flux inversions in detail.

Agreement with T_{soil} and scPDSI

The results show high consistency in the timing of anomalies between T_{soil} and GOSAT flux inversions on continental and larger scales in the tropics, and on sub-continental scales in the northern extratropics during JJA. These results indicate that T_{soil} is a useful proxy for corroborating IAV in flux inversions in both the tropics and northern extratropics. Linear regressions between GOSAT flux inversion and scPDSI IAV indicate moderate agreement on the scale of the entire tropics ($R^2 \leq 0.27$) and on sub-continental scales in the northern extratropics ($R^2 \leq 0.29$). The GOSAT flux inversion IAV consistently shows closer agreement with T_{soil} anomalies than with scPDSI in both the tropics and northern extratropics. This is consistent with previous research that has mostly shown that IAV is most closely related to temperature anomalies on large scales (Wang et al., 2013; Jung et al., 2017).

Although the results of this study indicate that T_{soil} is a useful metric for corroborating IAV in flux inversions, inferring the sensitivity of NEE anomalies to temperature anomalies directly is not advised for the fits given in Tables 5.3 and 5.2. This is because a number of factors have not been considered in this analysis. One factor is that temperature anomalies are also correlated with moisture and biomass burning anomalies. Keppel-Aleks et al. (2014a) show that accounting for these covariances results in reduced sensitivity of NEE anomalies to temperature anomalies. A second factor is that the relationship between NEE anomalies and temperature and moisture anomalies is variable, depending on large-scale climate modes. For example, Fang et al. (2017) show that either temperature or precipitation anomalies can be the primary driver NEE anomalies based on ENSO phase. A third factor is that the impact of temperature and moisture on NEE anomalies may be lagged. Ecosystems can take a months to years to recover from droughts (Frank et al., 2015; Schwalm et al., 2017; Sippel et al., 2018). Baldocchi

et al. (2018) found that a flux anomalies at number of FLUXNET sites are negatively correlated with themselves after a one-year lag implying a highly oscillatory behavior in the net carbon fluxes from year to year.

This leaves many opportunities for future work to further investigate the relationship between NEE anomalies and climate variability in more detail. A further limit to the comparisons of flux inversions with T_{soil} and scPDSI anomalies in the tropics is that we do not distinguish between seasons. The relationship between NEE, T_{soil} and scPDSI anomalies likely have substantial seasonal differences (Rödenbeck et al., 2018). We encourage future studies to examine the seasonally dependent relationships using longer flux inversions, as well as studies which investigate lagged correlations and climate mode relationships between inversion NEE anomalies and temperature and water availability anomalies.

Agreement with SIF

It is notable that correlations with SIF are weaker than those with the other proxies. Linear regressions indicate that SIF anomalies show some correspondence to GOSAT flux inversion anomalies on sub-continental scales in the northern extratropics during JJA ($0.14 \leq R^2 \leq 0.27$), however, little agreement is found in the tropics ($R^2 \leq 0.05$). This suggests that SIF may not be a good indicator of NEE IAV in the tropics, which is likely due to the strong covariances between GPP and R_e anomalies. However, it is also possible that the IAV in SIF observed by GOME-2 is not reliable as spurious trends have been found in the observations (Zhang et al., 2018). We examined the correlation between FLUXCOM MARS GPP and SIF anomalies to test if the SIF anomalies used here are consistent with independent estimates of GPP anomalies (Fig. 5.8). Spatially heterogeneous agreement is found between the two datasets, with the closest agreement occurring over semi-arid regions. However, correlations are generally positive over the majority of the globe, suggesting that IAV from GOME-2 SIF is reliable.

Agreement with FLUXCOM NEE

The GEOS-Chem GOSAT flux inversions with no prior IAV showed close agreement with FLUXCOM NEE anomalies in the tropics on regional and larger scales, and in the northern extratropics on regional scales during JJA. This is a remarkable finding as these data driven estimates of NEE IAV are independent, and agreement between the two estimates provides a strong indication that the results are robust. Therefore, comparisons with FLUXCOM NEE may provide a method for corroborating results from

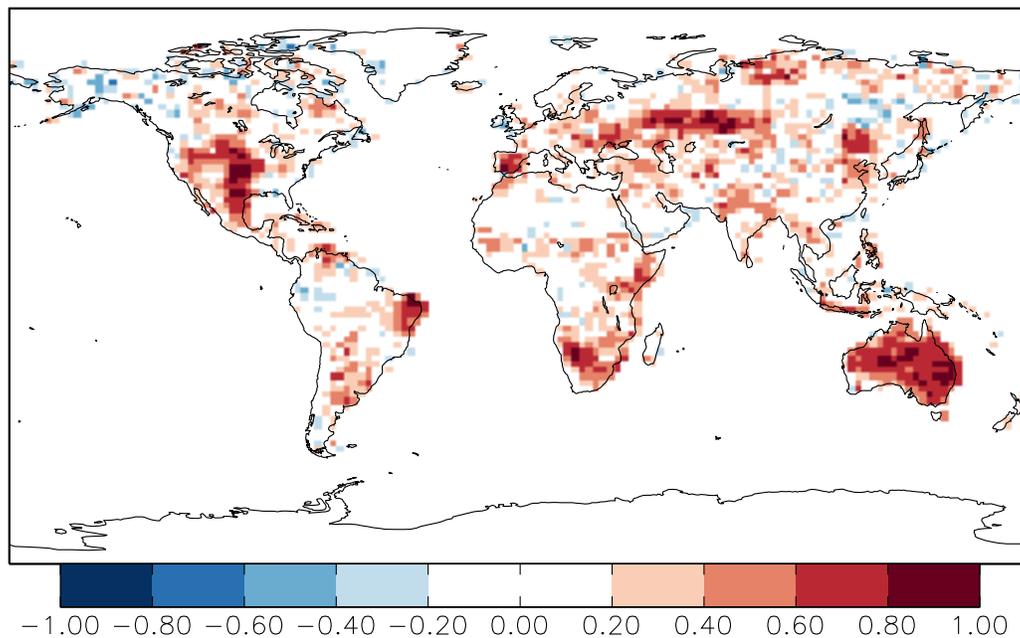


Figure 5.8: Correlation between FLUXCOM MARS GPP anomalies and SIF anomalies at $2^\circ \times 2.5^\circ$ spatial resolution.

flux inversion studies. However, it should be noted that the net annual NEE fluxes produced by FLUXCOM are quite unrealistic (Tramontana et al., 2016; Jung et al., 2017), with annual net draw-down by the biosphere of 18–28 Pg.

It may also be possible to evaluate the magnitude of IAV in FLUXCOM NEE through comparisons with flux inversions. Here we compare the magnitude of IAV between the GOSAT flux inversions and FLUXCOM NEE. The slope of the linear regression between the inversions indicates the relative magnitude of the inversion and FLUXCOM NEE anomalies. Over the entire tropics, the GOSAT inversions give slopes of 1.03–2.10 (mean of 1.56), suggesting that the magnitude of NEE anomalies are underestimated by FLUXCOM NEE. For JJA in the northern extratropics, the GOSAT inversions give slopes of 0.79–1.59 (mean of 1.31), again suggesting that the magnitude of NEE anomalies are underestimated by FLUXCOM. Furthermore, the OSSEs suggested that the inversions do not recover the full magnitude of IAV, providing further evidence that FLUXCOM underestimates the magnitude of IAV. This result is consistent with previous studies which indicate that FLUXCOM underestimate the magnitude of IAV (Jung et al., 2011, 2017).

5.4.2 Scales constrained

We investigated the agreement between monthly anomalies in flux inversions and proxies/FLUXCOM over a range of spatial scales in the tropics. The results showed that the agreement between the inversions and the proxies/FLUXCOM were scale dependent, which was corroborated by OSSEs. Here we synthesize these results and discuss the ability of GOSAT flux inversions to recover IAV in NEE over the range of scales examined in this study.

The results provide strong evidence that GOSAT flux inversion can constrain monthly flux anomalies on the scale of the entire tropics. All of the GEOS-Chem GOSAT flux inversions obtained $R^2 \geq 0.55$ for linear regressions with T_{soil} , and $R^2 \geq 0.51$ with FLUXCOM NEE. The OSSEs provide further evidence that the true NEE anomalies could be recovered, as linear regressions between the posterior and true anomalies give $R^2 \geq 0.53$. These results provide strong evidence that the GOSAT inversions are recovering the timing of tropical NEE anomalies, however, there is less agreement on the magnitude of flux anomalies over the tropics. The OSSEs indicate that GOSAT flux inversions can recover 42–68% of the magnitude of NEE anomalies, depending on the inversion set up.

On continental scales in the tropics, the results suggest that GOSAT flux inversion can constrain monthly flux anomalies. The GEOS-Chem inversions show good agreement with FLUXCOM NEE and T_{soil} anomalies. However, the agreement between the inversions and proxies/FLUXCOM on this scale is strongly influenced by the large-scale ENSO anomalies. We isolated the continental-scale anomalies by subtracting a mean tropical anomaly for GC_{2×2.5–200%} and FLUXCOM (Fig. 5.4), and for the OSSEs (Fig. 5.7). We found that the anomalies were still correlated after removing the mean tropical signal, suggesting that the continental-scale anomalies are largely recovered in the inversions. However, we also found that the inversion set up can have a significant influence on posterior anomalies on continental scales. The strongest correlations between the proxies and inversions were obtained with higher resolution, looser prior constraints, and no prior IAV. Similarly, OSSEs showed the best agreement with the truth when looser prior constraints were employed, but results were less clear for the impact of prior IAV. The OSSEs also showed that correlations with the truth were improved on continental scales when the prior mean seasonal cycle was closer to the truth. Overall, these results suggest that GOSAT observations contain information on continental-scale NEE anomalies in the tropics, however, recovering the correct IAV from these observations may be sensitive to the flux inversion set-up. Furthermore, the magnitude of NEE anomalies are likely underestimated.

On regional scales in the tropics, the results were much more ambiguous. The GOSAT

inversions generally showed good agreement with FLUXCOM NEE IAV on regional scales, but not with T_{soil} . The OSSEs also indicate marginal ability to recover regional scale fluxes. From these results, we caution against making conclusions about IAV on regional scales in the tropics using GOSAT flux inversions without corroborating evidence. On smaller scales, there is little evidence that the flux inversions recover IAV.

In the northern extratropics during JJA, the results of this study suggest that regional-scale constraints are possible. We found that large flux anomalies due to major climate events are recovered in the inversion for $GC_{2 \times 2.5-200\%}$ (Fig 5.5), while linear regressions showed close agreement for the GOSAT flux inversions with FLUXCOM NEE and T_{soil} . However, we also found evidence that the posterior NEE IAV was sensitive to the inversion set up. The inversion analyses with prior IAV ($GC_{4 \times 5-44\%-IAV}$, $GC_{4 \times 5-100\%-IAV}$, and GOSAT L4) showed weaker agreement with the proxies relative to the inversions without prior IAV. Similarly, the OSSEs showed prior IAV reduced agreement between the posterior and the “true” IAV.

5.4.3 Influence of the inversion configuration

Model horizontal resolution

The results of this study indicate that the spatial resolution of the model used in the inversion analysis ($2^\circ \times 2.5^\circ$ or $4^\circ \times 5^\circ$) has a relatively minor impact on posterior NEE anomalies. This is somewhat surprising since recent studies (Yu et al., 2018; Stanevich, 2018a,b) have shown significant transport differences for different resolution versions of GEOS-Chem. Also, Deng et al. (2015) showed that there are large biases in CO_2 in the upper troposphere and lower stratosphere in GEOS-Chem that impact inferred flux estimates. It is possible that although the model transport errors influence the flux estimates, the resolution-dependent transport processes are not sensitive to IAV for the time period considered here. It could also be related to the information content of GOSAT observations. As we have shown in this study, GOSAT observations only constrain IAV on regional and larger scales. If transport errors have the largest impact on smaller scales, it may explain why model resolution did not have a major impact on our results.

Prior error covariances

All of the GEOS-Chem inversions were performed with tight (44% for $4^\circ \times 5^\circ$ and 66% for $2^\circ \times 2.5^\circ$) and loose (100% for $4^\circ \times 5^\circ$ and 200% for $2^\circ \times 2.5^\circ$) prior error covariances. The prior error covariances generally had a significant impact on the posterior IAV. In

the tropics, inversions with loose prior constraints gave larger correlations with T_{soil} and FLUXCOM NEE on regional and continental scales. Similarly, for the OSSEs, looser prior constraints gave larger correlations with the truth on regional and continental scales. This suggests that the information content of the GOSAT observations is sufficiently large in the tropics that prior error covariances of 100% for $4^\circ \times 5^\circ$ or 200% for $2^\circ \times 2.5^\circ$ can be applied without degrading the posterior results by over fitting the observations.

In the northern extratropics, the inversions with tighter prior constraints gave larger correlations with T_{soil} and FLUXCOM NEE on regional and continental scales. Similarly, tight prior constraints gave larger correlations with the truth for the OSSEs. These results are the opposite of what was found for the tropics, and suggests that tighter error constraints (as a percentage of NEE) should be applied in the northern extratropics than in the tropics. These results suggest that, when the prior error covariances are loose in the northern extratropics, the inversion over fits the GOSAT observations which degrades the agreement with proxies (or the truth for OSSEs).

The largest impact of varying the prior error covariances is in the magnitude of posterior IAV. When loose prior constraints are applied the magnitude of NEE anomalies increases by 30–80% (15–30% for OSSEs) in the tropics and 5–60% (0–30% for OSSEs) in the northern extratropics. These results imply that care should be taken when making conclusions about the magnitude of NEE anomalies from this analysis. Based on the OSSEs, it seems likely that the inversions underestimate the magnitude of IAV on all scales.

Prior fluxes

We investigated the influence of prior IAV on posterior NEE anomalies in flux inversions by performing inversions with prior IAV ($GC_{4 \times 5-100\%-IAV}$ and $GC_{4 \times 5-44\%-IAV}$) and without prior IAV ($GC_{4 \times 5-100\%}$ and $GC_{4 \times 5-44\%}$), as well as OSSEs with and without prior IAV. In the tropics, the impact of prior IAV is generally small. For the GOSAT inversions, the presence of prior IAV degrades agreement with the proxies on all scales. In the OSSEs, the presence of prior IAV degrades the agreement with the truth on continental scales, but improves agreement on regional scales and over the entire tropics. In the northern extratropics, the presence of prior IAV has a large negative impact on agreement with proxies for GOSAT inversions and on agreement with the truth in the OSSEs.

The impact of prior IAV on the inversion is likely strongly dependent on how well the prior IAV reflects the true IAV. Presumably, if the prior IAV was close to the true IAV it would improve the posterior IAV. However, even when the prior and true IAV are

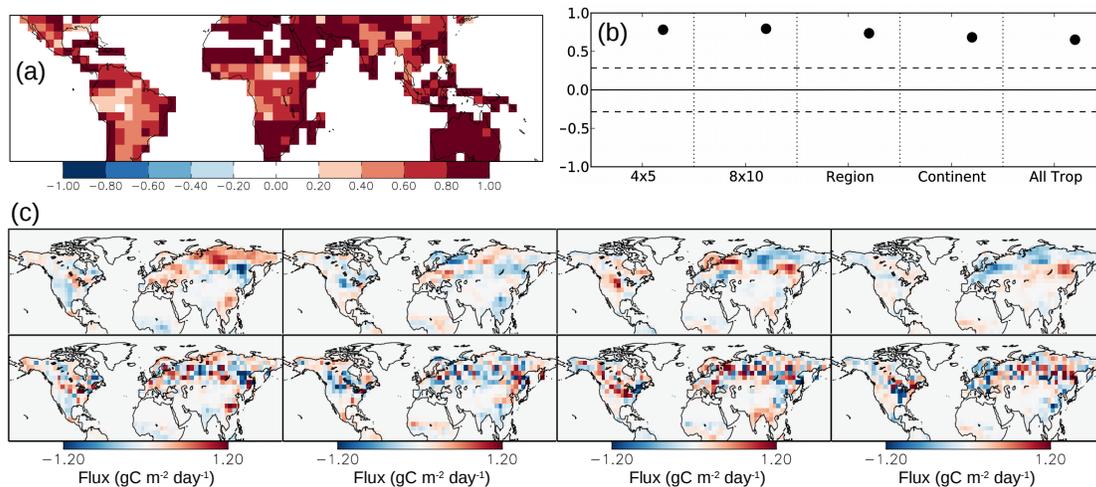


Figure 5.9: Comparison of GC_{IAV} posterior and prior IAV. (a) Correlation coefficient (R) between the posterior and prior IAV in the tropics at the spatial scale of $4^\circ \times 5^\circ$. (b) Mean correlation coefficient (R) between posterior and prior IAV in the tropics for different degrees of spatial aggregation. (c) Northern extratropical anomalies during JJA for (top) prior and (bottom) posterior NEE for (left–right columns) 2010–2013.

correlated, the posterior IAV can still be degraded. Presumably, the reason that prior IAV degrades posterior IAV is related to the fact that the observations under-constrain IAV, such that the prior IAV strongly influences the spatiotemporal distribution of IAV in the posterior NEE. To investigate this, we examined how closely the posterior IAV resembles the prior IAV. Figure 5.9 shows the agreement between the posterior and prior IAV for $\text{GC}_{4 \times 5-100\%-\text{IAV}}$ in the tropics and northern extratropics. Posterior IAV is strongly correlated with IAV in the prior, particularly on smaller scales. The fact that correlations between the prior and posterior IAV are strong at $4^\circ \times 5^\circ$ and $8^\circ \times 10^\circ$ is not surprising, as the NEE fluxes are strongly under-constrained at these spatial scales. However, the correlation with the prior IAV is substantially larger than with FLUXCOM on regional ($R^2 = 0.55$ versus $R^2 = 0.15$) and continental ($R^2 = 0.46$ versus $R^2 = 0.26$) scales as well. This suggests that IAV is still under-constrained even on continental scales. Only on the scale of the entire tropics is the correlation with the prior ($R^2 = 0.42$) less than with the proxies ($R^2 = 0.61$ for FLUXCOM NEE and $R^2 = 0.56$ for T_{soil}), indicating that the observations are influencing the IAV more than the prior.

The reason prior IAV degrades posterior IAV could also be linked to the prior error covariances. In our GEOS-Chem inversions, we prescribe error estimates as a fraction of the prior NEE (e.g., 44% for the $4^\circ \times 5^\circ$ inversions). Therefore, differences in the prior fluxes between years imply differences in the errors. Differences in the spatiotemporal

distribution of errors could lead to differences in the spatiotemporal distribution of NEE between years, which could make isolating the contribution for IAV in NEE challenging.

We also investigated the impact of the prior mean seasonal cycle on posterior IAV. We performed a series of OSSEs to examine the impact of the mean seasonal cycle of the prior fluxes on the inversion and found that correlations with the truth were significantly improved on continental scales when the mean seasonal cycle was closer to the truth. In particular, $\text{OSSE}_{CT2016-100\%}$ gives much weaker correlations with the truth than $\text{OSSE}_{JULES-100\%}$ after removing the mean tropical signal (Fig. 5.7). These results suggest that it is important to use prior fluxes with a realistic seasonal cycle to recover IAV in NEE from GOSAT observations.

5.5 Conclusions

In this study, we examined the constraints on interannual anomalies in NEE provided by GOSAT observations by performing a series of flux inversions. We addressed three main objectives in this analysis. The first objective was to quantify the agreement between GOSAT flux inversions and flux proxies, which are associated with IAV in the terrestrial carbon cycle, and FLUXCOM NEE. We found strong correlations ($P < 0.05$, $R > R_{\text{NINO3.4}}$) with FLUXCOM NEE and T_{soil} in the tropics on continental and larger scales, and in the northern extratropics on sub-continental scales during the summer ($R^2 > 0.49$), when there is no prior IAV. These results demonstrate that both FLUXCOM NEE and T_{soil} can be useful tools for corroborating flux inversion results. We found flux anomalies from GOSAT inversions were less correlated with scPDSI and SIF. For scPDSI we found some agreement on the scale of the entire tropics ($R^2 \leq 0.27$) and on sub-continental scales in the northern extratropics ($R^2 \leq 0.29$). For SIF, there was some agreement on sub-continental scales in the northern extratropics during JJA ($0.14 \leq R^2 \leq 0.27$), however, little agreement was found in the tropics ($R^2 \leq 0.05$).

The second objective was to determine the spatial scales over which the GOSAT inversion constrain flux anomalies. In the tropics, we found that continental and larger scale flux anomalies can be captured in GOSAT inversions. This conclusion is supported by strong agreement ($P < 0.05$, $R > R_{\text{NINO3.4}}$) with T_{soil} and FLUXCOM NEE, and a series of OSSEs which showed that the true IAV can be largely recovered on these scales. On regional scales in the tropics, the GOSAT flux inversions showed some agreement with the proxies and FLUXCOM, but the OSSEs indicated that the GOSAT observations likely strongly underconstrain IAV on these and smaller scales. In the northern extratropics, we found that flux anomalies are recovered by GOSAT flux inversions on sub-continental

regions during JJA. Strong agreement was found with anomalies in T_{soil} ($0.57 \leq R^2 \leq 0.65$) and FLUXCOM NEE ($0.49 \leq R^2 \leq 0.65$), when no prior IAV is used. OSSEs supported these findings, indicating that GOSAT observations can recover regional scale flux anomalies in the northern extratropics during JJA.

The third objective was to quantify the sensitivity of the results from the first two objectives to the inversion set-up. We found that the agreement between the flux inversions and proxies can be sensitive to the inversion set-up. Posterior flux anomalies were most sensitive to the prior fluxes and error covariances. In general, the inclusion of prior IAV in the inversion degraded the agreement with FLUXCOM NEE and T_{soil} , particularly in the extratropics, and this result was supported by the OSSEs. We compared the impact of the mean seasonal cycle on the posterior IAV by performing OSSEs and found that having a prior climatological seasonal cycle that was close to the truth improved posterior NEE anomalies on continental scales in the tropics. The prior error constraints also had a significant impact on the results. We found that looser constraints in the tropics gave better agreement with the proxies, while tighter constraints in the northern extratropics gave better agreement with the proxies (as a percentage of the prior flux). The magnitude of the prior constraints had a large impact on the magnitude of NEE anomalies. Also, the OSSEs showed that the magnitude of NEE anomalies are underestimated even with loose prior constraints. These results indicate that the prior fluxes and error covariances need to be carefully considered. We recommend that annually repeating prior fluxes be used to investigate IAV, and that the mean seasonal cycle of the prior fluxes be evaluated before performing the inversion. The results also indicate that defining the prior error covariance to be a fraction of the prior flux may produce either overfitting of GOSAT data in the northern extratropics or underfitting of the data in the tropics.

Overall, our results show that T_{soil} and FLUXCOM NEE are useful for evaluating IAV in flux inversions. Furthermore, comparisons with the anomalies in T_{soil} and FLUXCOM NEE suggest that GOSAT flux inversions are useful tools for constraining IAV in NEE on continental and larger scales in the tropics, and on regional scales in the northern extratropics during JJA.

Chapter 6

OP-FTIR

6.1 Introduction

Urban greenhouse gas (GHG) emissions account for 37–49% of direct global GHG emissions, and a larger fraction for indirect emissions (Seto et al., 2014). This is despite the fact that urban areas only occupy 0.2–2.7% of ice-free land (Seto et al., 2014). Thus, mitigating GHG emissions from urban areas will play an important role in mitigating climate change. Bottom-up accounting of emissions will be required to monitor emission reductions, however, independent verification of these estimates will also be needed. Measurements of GHG enhancements in urban areas provide an independent observational constraint to estimate emissions. Over the past decade, a number of projects have been initiated to monitor GHGs in urban areas with the goal of providing top-down emission constraints. Monitoring urban GHG emissions using atmospheric observations is still at the experimental stage, with different approaches being applied for different projects. The differences in approaches are partially due to available resources and the specific problem addressed, but also because no approach is clearly superior to any other. A number of different measuring platforms have been used, including surface-based in-situ and flask measurements (Shusterman et al., 2016; Miles et al., 2017; Verhulst et al., 2017; Bares et al., 2018; Xueref-Remy et al., 2018; Sargent et al., 2018), remote sensing of solar absorption spectra (Viatte et al., 2017; Hedelius et al., 2018), and aircraft-based in-situ observations (Cambaliza et al., 2015).

Here, we present an Open-Path Fourier Transform Infrared (OP-FTIR) spectroscopy system for GHG monitoring in Toronto, Canada. OP-FTIR spectroscopy allows for the monitoring of the abundances of atmospheric gases over an integrated atmospheric path of up to several kilometers in length. An advantage of open-path systems over more traditional in situ sampling observations is that open-path measurements are less sen-

sitive to highly localized emissions in heterogeneous landscapes such as cities (Waxman et al., 2017; Griffith et al., 2018), which have many different emission sources of greenhouse gases. The integrated atmospheric path of these systems is more representative of mesoscale atmospheric tracer transport models with grid cells on the order of 1 km^2 (i.e., $1 \text{ km} \times 1 \text{ km}$, Lauvaux et al., 2016) than in situ observations. Therefore, comparisons between forward models, simulating atmospheric trace gas abundances, and OP-FTIR measurements are less affected by representativeness errors. The advantage of an integrated atmospheric path has led to the deployment of OP-FTIR systems for a number of applications, including monitoring of trace gases released by biomass burning (Paton-Walsh et al., 2014), agricultural emissions (Flesch et al., 2016), shipping emissions (Wiacek et al., 2018), city emissions (Griffith et al., 2018), and vehicle emissions (You et al., 2017). You et al. (2017) demonstrate the advantage of OP-FTIR measurements over in situ measurements in an urban environment, showing that OP-FTIR measurements of trace species over a highway have much less wind direction dependence than in situ measurements.

This study has two main objectives. The first is to describe the deployment of an OP-FTIR observing system in downtown Toronto, and retrieval of CO_2 , CH_4 , CO , N_2O , and H_2O from near continuous observations starting in November 2017. The second is to examine the utility of estimating urban emissions of CO_2 by comparing OP-FTIR measurements with in situ measurements 5.4 km south of the OP-FTIR, at the southern edge of Toronto. The chapter is structured as follows. Sec. 6.2 and Sec. 6.3 provide an overview of the instrumentation and experimental set-up. Sec. 6.4 and Sec. 6.5 provide descriptions of the data collection and trace gas retrievals. Sec. 6.6 examines nearby meteorological data. Sec. 6.7 presents the results of this study. Retrieved gas concentrations are compared with nearby meteorological and in situ GHG measurements, and preliminary flux estimation results are presented. Sec. 6.8 gives the conclusions of this study.

6.2 Instrumentation

6.2.1 Bruker 125M

The Bruker IFS 125M is a mobile high-resolution spectrometer designed to have a spectral range covering the mid-infrared through ultra-violet region with a spectral resolution of up to 0.008 cm^{-1} in the mid-infrared and a maximum optical path difference (OPD) of 112.5 cm. However, for this set-up, measurements are recorded over the range 1,900–

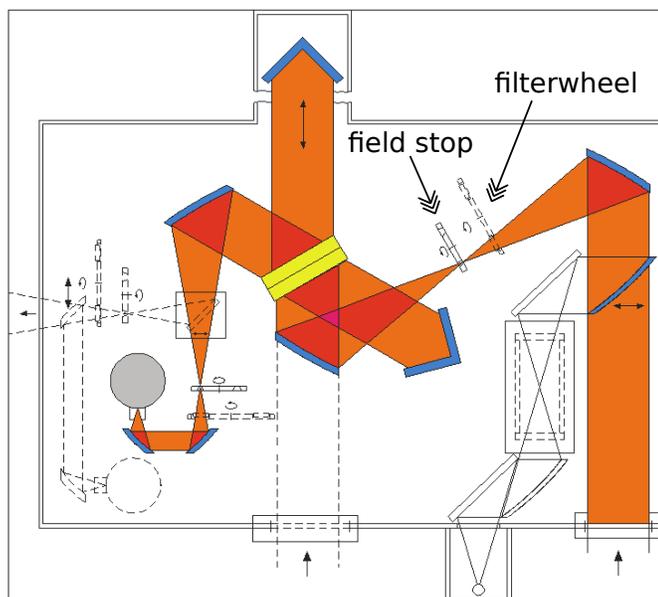


Figure 6.1: Optical path of the Bruker IFS 125M spectrometer. Reprinted from the IFS 125M manual.

$6,000\text{ cm}^{-1}$ with 0.4 cm^{-1} resolution. Figure 6.1 shows a schematic of the optical set-up of the instrument. A collimated input beam is focused by an off-axis-paraboloid (OAP) mirror onto computer-controlled filter and aperture wheels (no filter is used in the OP-FTIR set-up). After passing the field stop, the input beam is re-collimated onto the beamsplitter. For the OP-FTIR set-up used here, a CaF_2 beam-splitter ($14,000\text{--}1,850\text{ cm}^{-1}$) is used. After recombining the beam from the two arms of the interferometer, a computer-controlled mirror stage selects between two detectors. We use the photovoltaic indium antimonide (InSb) detector ($1,850\text{--}9,600\text{ cm}^{-1}$).

Measurement technique

The spectrometer records an absorption spectrum using the Fourier transform technique. The scanning mirror produces an OPD between two beams separated by the beam-splitter. Therefore, there is interference when the two beams are recombined by the beam-splitter. The interference of these beams determines the intensity of radiation recorded by the detector. By measuring the intensity of radiation as a function of OPD an interferogram can be generated. The interferogram is converted into a spectrum using a Fourier transform. In the practical implementation of this technique, additional considerations are required. In particular, the Norton-Beer Medium apodization function is applied to the interferogram to reduce the impact of the finite maximum OPD on absorption line sidelobes.

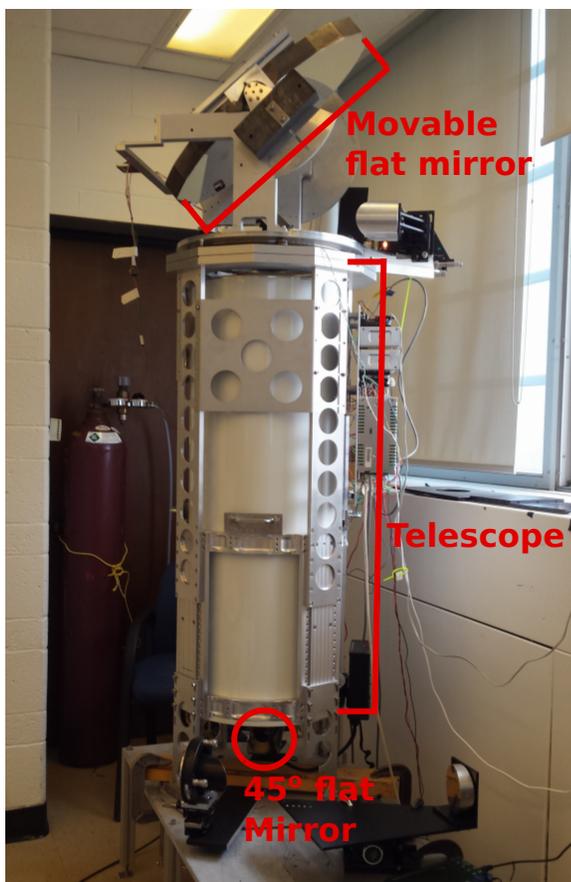


Figure 6.2: The OP-FTIR telescope and movable mirror mounted within a custom-built frame.

6.2.2 Telescope

We use an RC Optical 12-inch Ritchey-Chrétien telescope, with gold-coated mirrors. The telescope is oriented vertically and is housed within a custom built frame (Fig. 6.2). The incoming infrared (IR) beam is directed into the top of the telescope by a 14-inch movable mirror placed at the input of the telescope, which can be rotated and tilted, thereby acquiring and directing the IR beam. The mirror motors are controlled by two Applied Motion 3540i stepper motor drivers and Wantai 57BYG621 1.8 deg 2.2 Amp stepper motors. A flat 45° mirror is located at the bottom of the telescope which re-directs the vertical IR beam exiting the telescope horizontally towards the 125M.

6.2.3 Retro-reflector

We use a PLX AR-30-5 corner cube array retro-reflector housed in a National Electrical Manufacturers Association (NEMA) enclosure that can be remotely opened via a radio

control link to a garage door opener attached to the frame.

6.3 Instrument set-up

The IR globar source is mounted at the top of the telescope (Fig 6.3a). To approximate an infinitesimal source, an aperture is placed directly in front of the globar. There is a shutter located directly in front of the aperture, which can be raised to let the IR beam pass or lowered to block the IR beam (the shutter is used to block the IR beam to collect ambient spectra). After passing through the aperture, the beam diverges until it reaches an OAP mirror. The OAP mirror's focal point is located at the position of the aperture so that the IR beam is collimated and directed towards a 45° flat mirror placed at the center of the telescope. The beam is reflected off this mirror and onto the movable flat mirror which directs the beam to the retro-reflector. Upon returning from the retro-reflector, the beam fills the movable mirror and is directed into the telescope (Fig 6.3b). Note that a fraction of the returning IR beam is directed back to the globar source, while the rest of the beam is directed into the telescope.

When the IR beam enters the telescope, it is focused by the gold-coated hyperbolic mirrors (Fig. 6.4). Upon exiting the telescope, the beam is re-directed to an OAP mirror by a 45° flat mirror. The OAP mirror is located such that its focal point matches the focal point of the telescope. Theoretically, the IR beam should be collimated after reflecting off the OAP mirror. However, we were unable to position the OAP such that the beam was collimated. Instead, we found that the beam would focus after reflecting of the single OAP. To solve this problem, we placed a second OAP so that the focal point of the second OAP matched the position where the beam comes to a focus. The beam is better collimated after reflecting off this second OAP. The collimated beam is then directed into the Bruker IFS 125M.

6.3.1 Location

The system is set up in a room on the 12th floor of the Burton tower of the McLennan Physical Laboratories about 45 m AGL (43.6604° N, 79.3983° W) and the retro-reflector is set up on the roof of the Galbraith building about 20 m AGL (43.6600° N, 79.3964° W) (Fig. 6.5). The one-way path between the source and retro-reflector is about 160 m. The path crosses two streets (Russell St. and St. George St.) which are two-lane streets, and is surrounded by multi-story buildings.

The OP-FTIR is located within the St. George Campus of the University of Toronto.

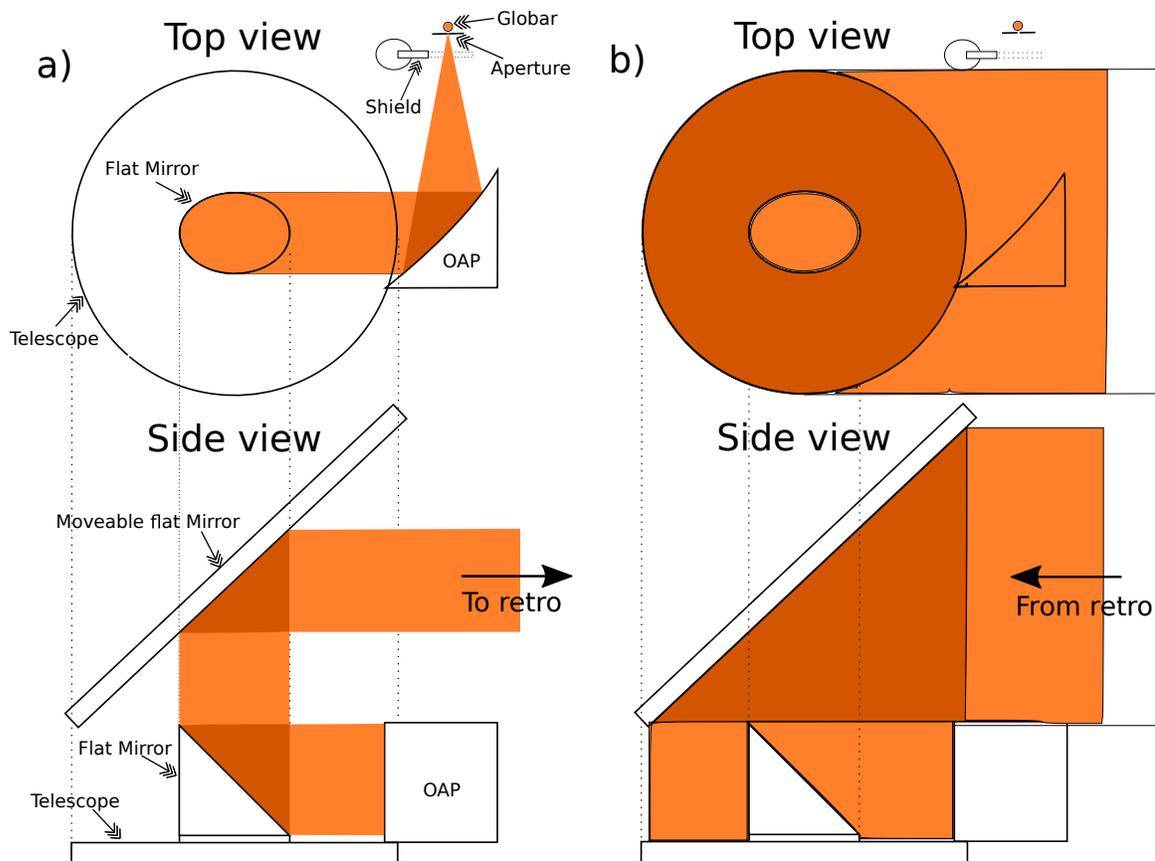


Figure 6.3: View of the top of the telescope. Ray tracing (a) from the globar to the retro-reflector and (b) from the retro-reflector into the telescope are shown. The shading indicates the path of the IR beam.

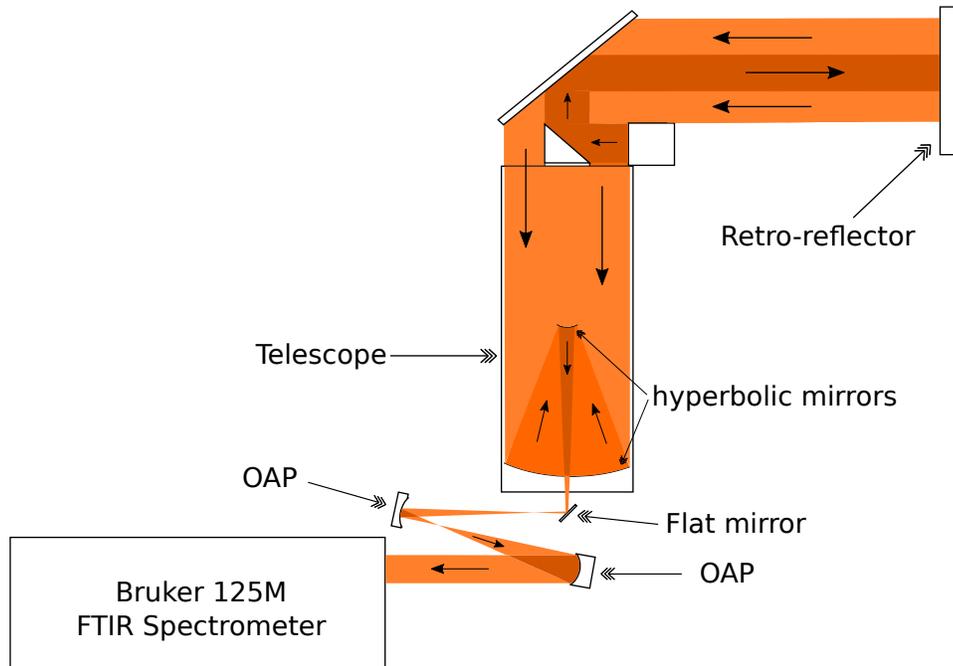


Figure 6.4: Light path for the entire optical system. The light path is shown as a collimated beam emerging from the OAP at the top of the telescope, traveling to and returning from the retro-reflector, being focused by the telescope and collimated again before being directed into the Bruker 125M.

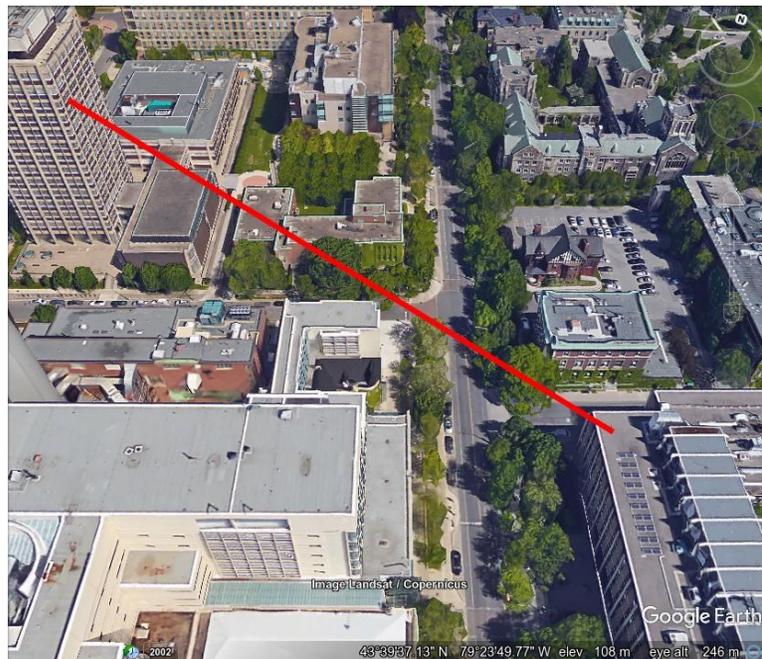


Figure 6.5: Path between the instrument and retro-reflector. (Google Earth V 7.3.2.5483. (June 22, 2018). Toronto, Canada, 43°39'37.13" N 79°23'49.77" W, eye alt 246 m.)

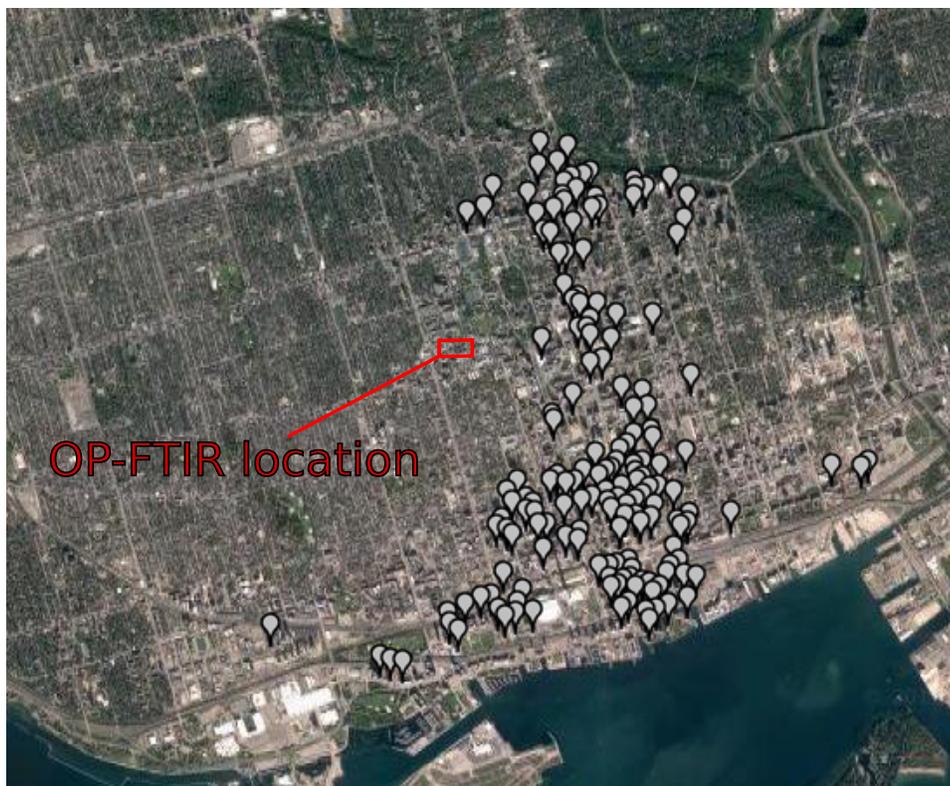


Figure 6.6: Location of the OP-FTIR system in downtown Toronto, with buildings which are 100 m or greater in height indicated by markers (<http://skyscraperpage.com/cities/maps/?cityID=12>).

Figure 6.6 shows the location of the OP-FTIR setup within downtown Toronto. The setup is approximately 3.5 km north of Lake Ontario. The downtown core of Toronto is located to the east and southeast with many tall buildings (Fig. 6.6). The city of Toronto is home to nearly 3 million people and is located within the “Golden Horseshoe”, a densely populated region of about 9 million which straddles the northeast and southeast coasts of Lake Ontario.

6.4 Data collection

Spectra are collected using an OPUS macro (Bruker software that comes with the 125M). The macro performs the following series of tasks in a repeating loop. First, the macro raises the global shutter and records 40 co-added scans at 0.4 cm^{-1} resolution. Then, the macro lowers the shutter and records 40 co-added scans at 0.4 cm^{-1} resolution. These steps are repeated until the system is turned off. An example of the data collected from one iteration of the macro is shown in Fig. 6.7a. The spectra collected with the shutter

raised and lowered are shown in blue are red respectively.

6.4.1 Spectra processing

To isolate the detection of radiation from the global source, we must subtract ambient blackbody emission and stray light that enters the instrument. To do this, we take the difference between measurements with the shutter raised and shutter lowered:

$$\mathbf{y}_{i1}^{r-l} = \mathbf{y}_i^r - \mathbf{y}_{i-1}^l, \quad (6.1)$$

$$\mathbf{y}_{i2}^{r-l} = \mathbf{y}_i^r - \mathbf{y}_{i+1}^l \quad (6.2)$$

where \mathbf{y}^l is the measured spectrum with the IR source blocked (recorded when shutter is lowered, such that the signal is only due to ambient emission between the telescope and retro-reflector), \mathbf{y}^r is the measured spectrum with the IR source (recorded when the shutter is raised, so that the signal is due to the IR source and ambient emissions between the telescope and retro-reflector), and \mathbf{y}^{r-l} is the calculated spectrum due to the IR source only (includes absorption by gases along the path between telescope and retro-reflector). The subscript “ i ” represents the iteration for data collection. Figure 6.7b shows the difference between the spectra recorded with the shutter raised and lowered.

Because the ambient spectra can change quite rapidly, we calculate the absorption spectrum using the ambient radiation collected before and after the spectrum \mathbf{y}_i^r . Equation 6.1 is used when the shutter-lowered ambient-light spectrum is recorded before the shutter-raised spectrum, and Eq. 6.2 is used when the shutter-lowered ambient-light spectrum is recorded after the shutter-raised spectrum. The transmission spectrum (\mathbf{y}^t) is then obtained by taking the ratio of \mathbf{y}^{r-l} and a background spectrum (\mathbf{y}^0):

$$\mathbf{y}_{i1}^t = \frac{\mathbf{y}_{i1}^{r-l}}{\mathbf{y}^0}, \quad (6.3)$$

$$\mathbf{y}_{i2}^t = \frac{\mathbf{y}_{i2}^{r-l}}{\mathbf{y}^0}. \quad (6.4)$$

The background spectrum (\mathbf{y}^0) is the spectrum due to the global source before transmission through the atmospheric path. To obtain this spectrum, the retro-reflector was placed at a distance of 2 m from the movable flat mirror. Spectra were measured with the shutter raised and lowered, and \mathbf{y}^0 was calculated using Eqs. 6.1 and 6.3. The trace gas retrieval is performed on \mathbf{y}_{i1}^t and \mathbf{y}_{i2}^t , and the retrieved mole fractions are averaged. Currently, the same \mathbf{y}^0 spectrum is used for all retrievals. However, variations in the source (due to factors such as temperature changes) could impact the source signal. Therefore,

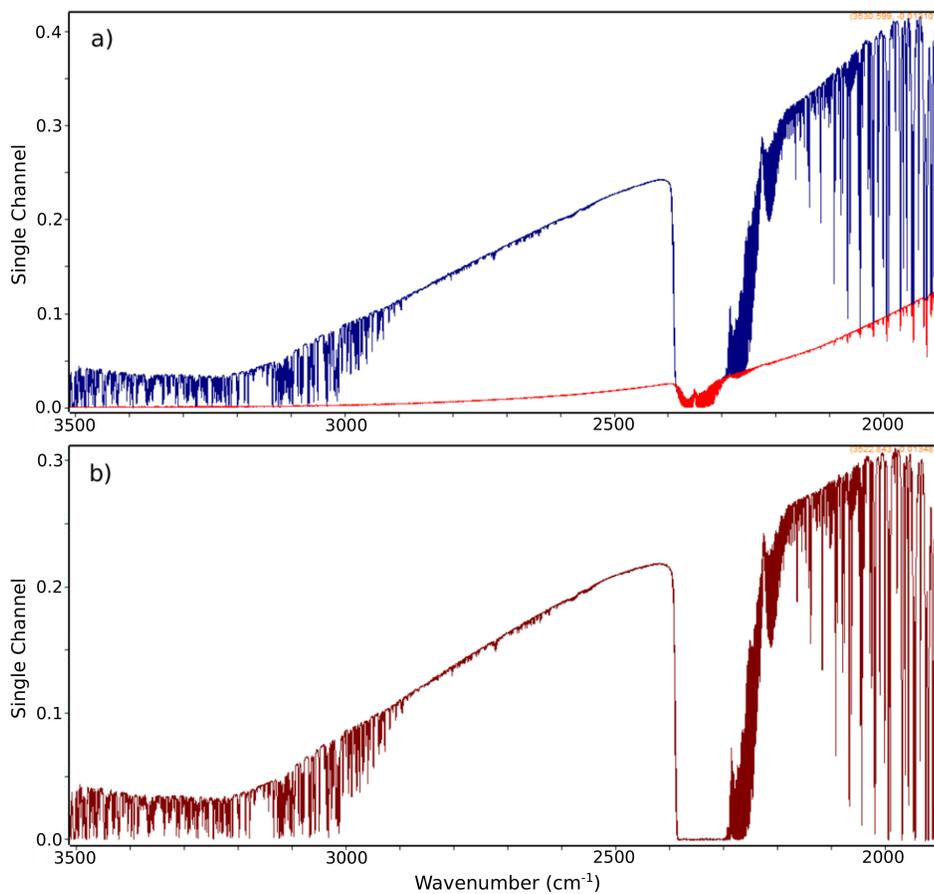


Figure 6.7: a) OP-FTIR absorption spectra recorded with the global shutter raised (blue) and lowered (red). b) The difference between spectra recorded with the shutter raised and lowered.

in the future it would be preferable to measure background spectra on a regular basis.

6.5 Trace gas retrievals

Trace gas dry-air mole fractions are retrieved using version 5 of the Multiple Atmospheric Layer Transmission (MALT) code (Griffith, 1996, 2003). CO₂, CO, N₂O, and H₂O dry-air mole fractions are retrieved in the range 2132.52–2234.97 cm⁻¹. CH₄ and H₂O are retrieved from 2900–3027 cm⁻¹. These wavenumber ranges were selected based on personal communication with Aldona Wiacek and Li Li at St. Mary’s University, and are the same ranges used for their retrievals (Wiacek et al., 2018). Note that Appendix C provides a description of how to run the codes to perform the retrieval.

6.5.1 Radiative transfer theory

The description here follows from Goody and Yung (1995), Petty (2006) and Pierrehumbert (2010). The absorption and emission of IR radiation by atmospheric gases is caused by transitions in the vibrational–rotational states of the molecules. Each possible energy transition is associated with a unique frequency of radiation and results in an absorption or emission line. The ability of a given absorption line to attenuate radiation is characterized by an absorption coefficient (σ) and has dimensions of area per mass (or area per molecule). The absorption coefficient for an absorption line centered at frequency ν_0 is defined as,

$$\sigma(\nu - \nu_0) \equiv S \cdot f(\nu - \nu_0), \quad (6.5)$$

where S denotes the strength of the absorption line and $f(\nu - \nu_0)$ denotes the line shape. The strength of a given absorption/emission line is dependent on the transition probability between two energy states and the fraction of molecules that occupy the initial energy state (which is dependent on temperature). The line shape deviates from the Dirac delta function due to two main broadening mechanisms: Doppler broadening and pressure broadening. Doppler broadening is caused by the Doppler shifts resulting from the velocity distribution of the molecules. Pressure broadening is a result of elastic collisions between molecules.

The optical depth, $\tau(\nu)$, characterizes the absorption over an atmospheric path and is given by

$$\tau(\nu) = \sum_i \sum_k \sigma_i^k(\nu) \cdot a_i, \quad (6.6)$$

where i denotes atmospheric species, k denotes absorption lines, and a_i is the column

density and is equal to the path length times the number density. The transmission spectrum, T , is given by

$$T(\nu) = \frac{I(\nu)}{I_0(\nu)} = e^{-\tau(\nu)}, \quad (6.7)$$

where $I(\nu)$ and $I_0(\nu)$ are the intensities before and after transmission through the path.

In real-world applications of remote sensing, the influence of the observing system on the recorded spectra also needs to be considered. The intensity of radiation recorded by an FTIR spectrometer is convolved with an instrumental line shape function ($f_I(\nu)$). The instrument line shape is a function of the maximum optical path difference, field-of-view, instrument misalignment and apodization applied to the interferogram. The measured transmission, $T'(\nu)$, is given by

$$T'(\nu) = \frac{I'(\nu)}{I'_0(\nu)} = \frac{I(\nu) * f_I(\nu)}{I_0(\nu) * f_I(\nu)}, \quad (6.8)$$

where $I'(\nu)$ and $I'_0(\nu)$ are the measured intensities before and after transmission through the path.

6.5.2 MALT

We use MALT5 (Griffith, 1996, 2003) to retrieve gas dry-air mole fractions. MALT consists of a forward model that calculates transmission spectra and an inverse algorithm that optimizes gas dry-air mole fractions from measured spectra.

Forward model

The MALT forward model calculates the atmospheric transmission spectra from inputs of species dry-air mole fractions, temperature, pressure, and path length, in addition to instrument parameters to account for the instrument line shape. MALT calculates a set of reference spectra based on the HITRAN line parameter database; we use version 2012 of the HITRAN database (Rothman et al., 2013) for our analysis. The forward model is described in detail in Griffith (1996).

Inversion

A measured spectrum is fitted to a MALT forward-model-calculated spectrum by iterative non-linear least squares to obtain a best-fit (minimum residual) match to the measured spectrum, as detailed in Griffith et al. (2012). Component gas abundances and instrument line shape parameters can be fixed or fitted as part of the fitting process. The inversion is

performed using the Levenberg-Marquardt non-linear least squares algorithm. Optimized gas path-densities (and instrument line shape parameters) are found by iteratively solving the equation

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{H}^T \mathbf{H} + \gamma_i \mathbf{I})^{-1} \mathbf{H}^T [\mathbf{y}^t - H(\mathbf{x}_i)]. \quad (6.9)$$

where \mathbf{y} is the observed spectrum, \mathbf{x} is a vector of the gas path-densities for multiple species, $H()$ is the MALT forward model, \mathbf{H} is the weighting function matrix, and γ is a weighting parameter which is adjusted between optimization steps. When $\gamma_i \rightarrow 0$, the step tends to the Gauss-Newton method, whereas when $\gamma_i \rightarrow \infty$, the step tends to the steepest descent and the step size tends to zero (Rodgers, 2000). The implementation of this algorithm within MALT is based on the description of the algorithm in Press et al. (1992). The same a priori dry-air mole fractions for retrieved gases are used for all retrievals, they are taken to be 1 % of the atmosphere for H_2O , 400 ppm for CO_2 , 150 ppb for CO , 300 ppb for N_2O and 1700 ppb for CH_4 .

Processing

The spectra are processed through the MALT5 executable in batch mode. Two text files are input to the executable: first, a parameter file with the instrument information and path length and, second, a text file containing the path (file and directory on the computer) of the source spectra (\mathbf{y}^{r-l}), source background data files (\mathbf{y}^0), temperature and pressure measurements (see Sec. 6.6). An example of an optimized spectral fit is shown in Fig 6.8. MALT5 outputs optimized wet-air mole-fractions for each gas as well the residuals for the fits. Wet-air mole fractions are converted to dry-air mole fractions using:

$$\text{GAS}_{dry} = \text{GAS}_{wet} \cdot \frac{1}{1 - \text{H}_2\text{O}_{wet}} \quad (6.10)$$

where

$$\text{GAS}_{wet} = \frac{mol_{gas}}{mol_{wet\ air}}, \quad \text{GAS}_{dry} = \frac{mol_{gas}}{mol_{dry\ air}}. \quad (6.11)$$

6.6 Meteorological data

A Davis Vantage Pro Plus 2 meteorological station (met-station) is installed on the roof of the Burton tower of the McLennan Physical Laboratories, approximately 61 m AGL (174 m above sea level). The met-station is installed on the side-railing on the southeast corner of the building, approximately 2 m above the roof. The location is not ideal, as there are obstacles on the building within ~ 2 m of the met-station which likely impact

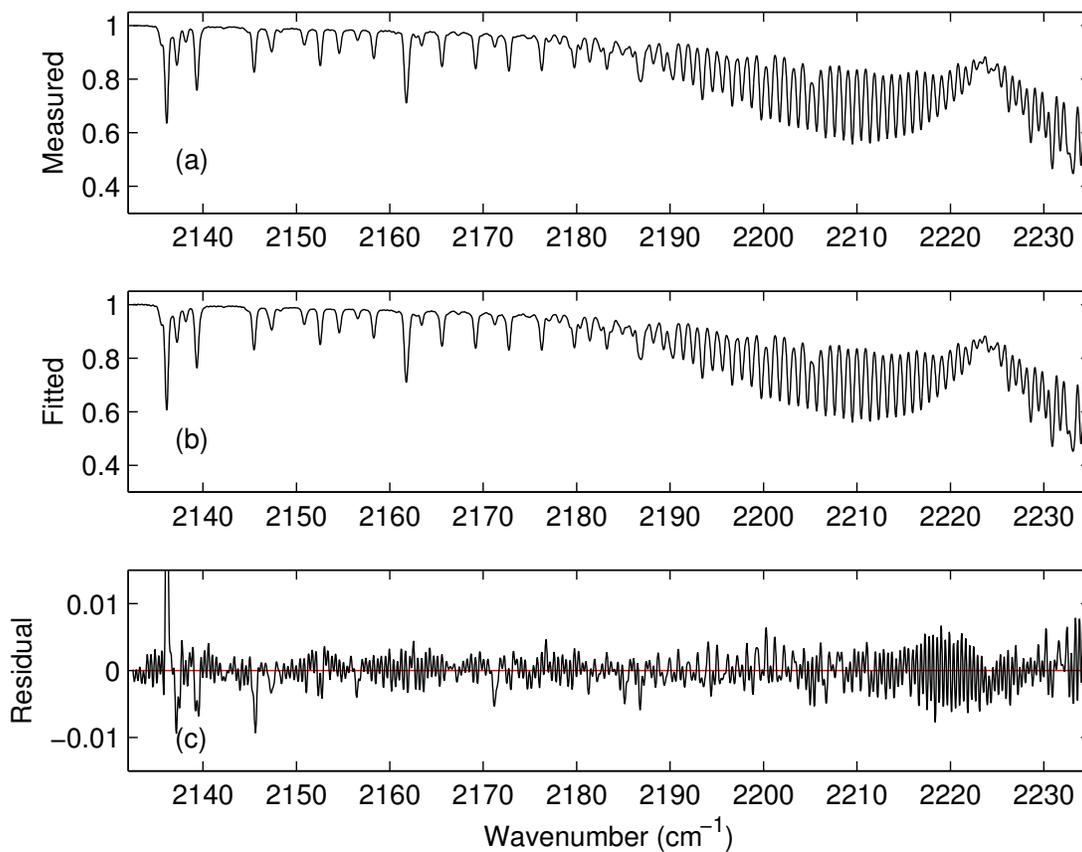


Figure 6.8: (a) Measured, (b) fitted, and (c) residual spectra in the wavenumber range 2132.52–2234.97 cm⁻¹ for 1 April 2018 at 22:25 local time. The fitted spectrum is obtained by optimizing CO₂, CO, N₂O, and H₂O dry-air mole fractions.

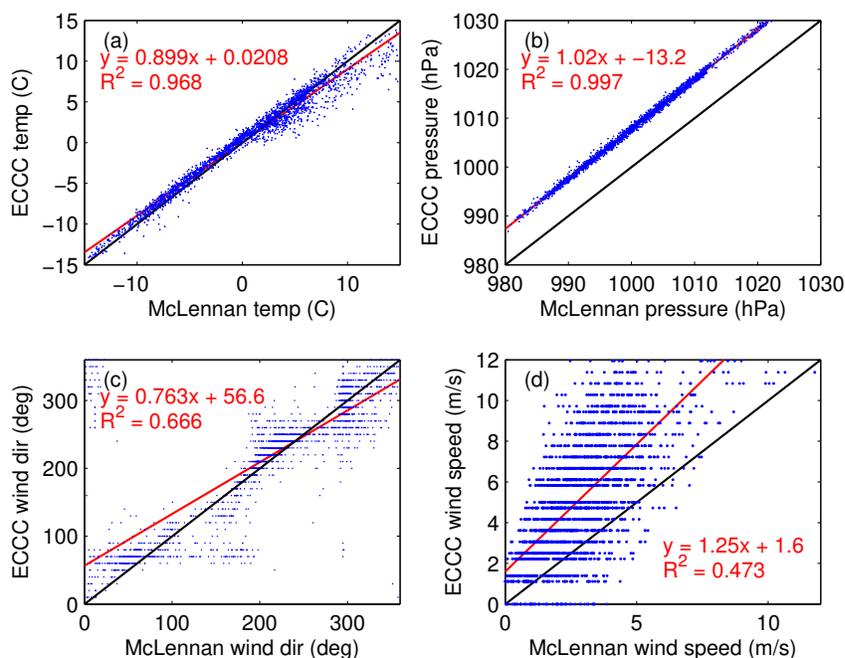


Figure 6.9: Scatter plots of ECCC vs. McLennan met-station data. The panels show (a) temperature, (b) pressure, (c) wind direction, and (d) wind speed after averaging to hourly temporal resolution. The black line indicates the 1:1 line and the red line shows the linear regression.

the observed winds. The met-station records meteorological variables (including temperature, pressure, wind direction, and wind speed) at one-minute temporal resolution. Measurements from the McLennan met-station have been compared against data from two nearby met-stations. One station, referred to as “Wallberg”, is located on the roof of the four-story Wallberg building, approximately 300 m from the McLennan building. This site is run as part of the Southern Ontario Centre for Atmospheric Aerosol Research (SOCAAR). The other site, referred to as the “ECCC” station, is run by ECCC and is located about 4 km south of the McLennan building on Toronto Island (note that the met-station is not at Hanlan’s Point). Figures 6.9 and 6.10 show comparisons of temperature, pressure, wind direction, and wind speed from the McLennan building vs. ECCC and Wallberg data, respectively. Note that the ECCC met-station records wind observations with lower wind speed and wind direction resolution than the other sites, leading to apparent gaps in Fig. 6.9. There is high consistency in measurements of temperature and pressure between the stations. Based on the Hypsometric equation and pressure offsets between stations, the Wallberg and ECCC stations are ~ 13.4 m and ~ 104.9 m and lower than the McLennan station, respectively. In reality, Wallberg station is 11 stories below

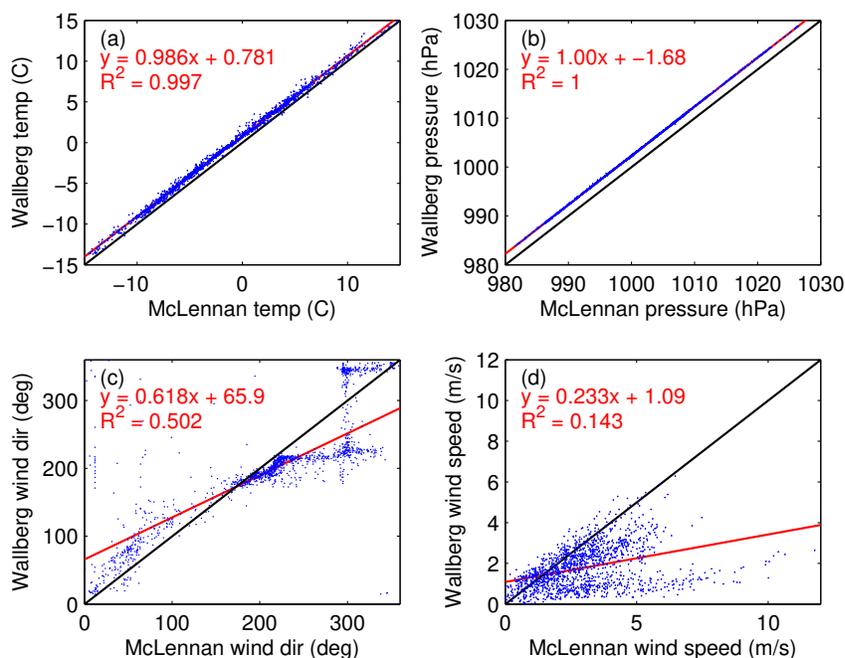


Figure 6.10: Scatter plots of Wallberg vs. McLennan met-station data. The panels show (a) temperature, (b) pressure, (c) wind direction, and (d) wind speed after averaging to hourly temporal resolution. The black line indicates the 1:1 line and the red line shows the linear regression.

the met-station (~ 36 m) and the ECCC station just above lake Ontario (~ 100 m below McLennan station). Thus, the pressure difference between McLennan and ECCC stations seems reasonable, but it appears that pressure measurements at Wallberg station are biased low.

The sites show less consistency in wind data. Part of the reason for this is that the variance is larger for wind observations. For example, in calculating hourly means from one minute measurements for the McLennan data, standard deviations of 0.22 °C are obtained for temperature, 0.17 hPa for pressure, 0.43 m s^{-1} for wind speed, and 17° for wind direction. However, this variance cannot explain all of the differences in wind fields between met-stations. It is likely that the placement of the met-stations also impacts the differences between sites. First, both the McLennan and Wallberg met-stations are close to surfaces. This could result in local turbulent flow that is not representative of the flow further above the surfaces. Second, differences in the height above ground level could impact observed winds. The heights of the met-stations are comparable to the mean building height in the surrounding area. This may lead to complicated flows which are variable over short distances, resulting in differences in the winds between sites.

6.7 Results

Note that none of the trace gas retrievals presented here have been calibrated against independent measurements, however, in Appendix B I describe some rudimentary attempts to calibrate the measurements and challenges faced. Figure 6.11 shows the time series of the OP-FTIR dry-air mole fractions from November 6, 2017, through June 30, 2018. A number of large enhancements are observed over this time period for all gases. CO and CO₂ are highly correlated ($R^2=0.83$ for hourly means) throughout the time series. This is expected because CO₂ and CO are both produced by the combustion of fossil fuels. However, the correlation is greater over the winter ($R^2=0.87$ before April 1, 2018) than during the growing season ($R^2=0.79$ after April 1, 2018), due to the fact that the biosphere strongly influences CO₂ but not CO. N₂O shows a much smaller dynamic range than CO and CO₂, and only varies by $\sim 3\%$ of its mean dry-air mole fractions. That said, there is still considerable variability in N₂O over the winter, which is indicative of a local source. There also appears to be an enhancement between the middle of April and early June. This could be due to fertilizer being deployed on farms or on household lawns. CH₄ is found to be highly variable throughout the time series, with dry-air mole fractions frequently exceeding 2600 ppb. This range in abundance is larger than would be expected from background city emissions, and suggests a spatially localized source. Interestingly, the variability in CH₄ seems to decrease through the spring, with less large enhancements. This could be due to reduced emissions from a localized source.

6.7.1 Sensitivity of GHGs to meteorology

Figure 6.12 shows hourly-mean gas dry-air mole fractions as a function of wind speed (from the McLennan met-station). All gases show enhancements at low wind speeds, indicating that local urban emissions are significantly impacting observed dry-air mole fractions. Figure 6.13 shows pollution roses of the mean gas dry-air mole fractions as a function of wind speed and direction. CO and CO₂ show the greatest wind speed dependence of the gases, with large enhancements at low wind speeds (mean dry-air mole fraction increase by 150% for CO and 13% for CO₂ at low wind speeds relative to high wind speeds). This is expected, as fossil fuel combustion is expected to be a large source of these species nearby, within downtown Toronto. Furthermore, CO and CO₂ do not show large wind direction dependence. N₂O shows wind speed dependence (mean dry-air mole fraction increase by 1.5% at low wind speeds relative to high wind speeds), suggesting that there is a substantial urban source of N₂O, presumably from combustion (Becker et al., 1999; Colorado et al., 2017). N₂O also shows some wind

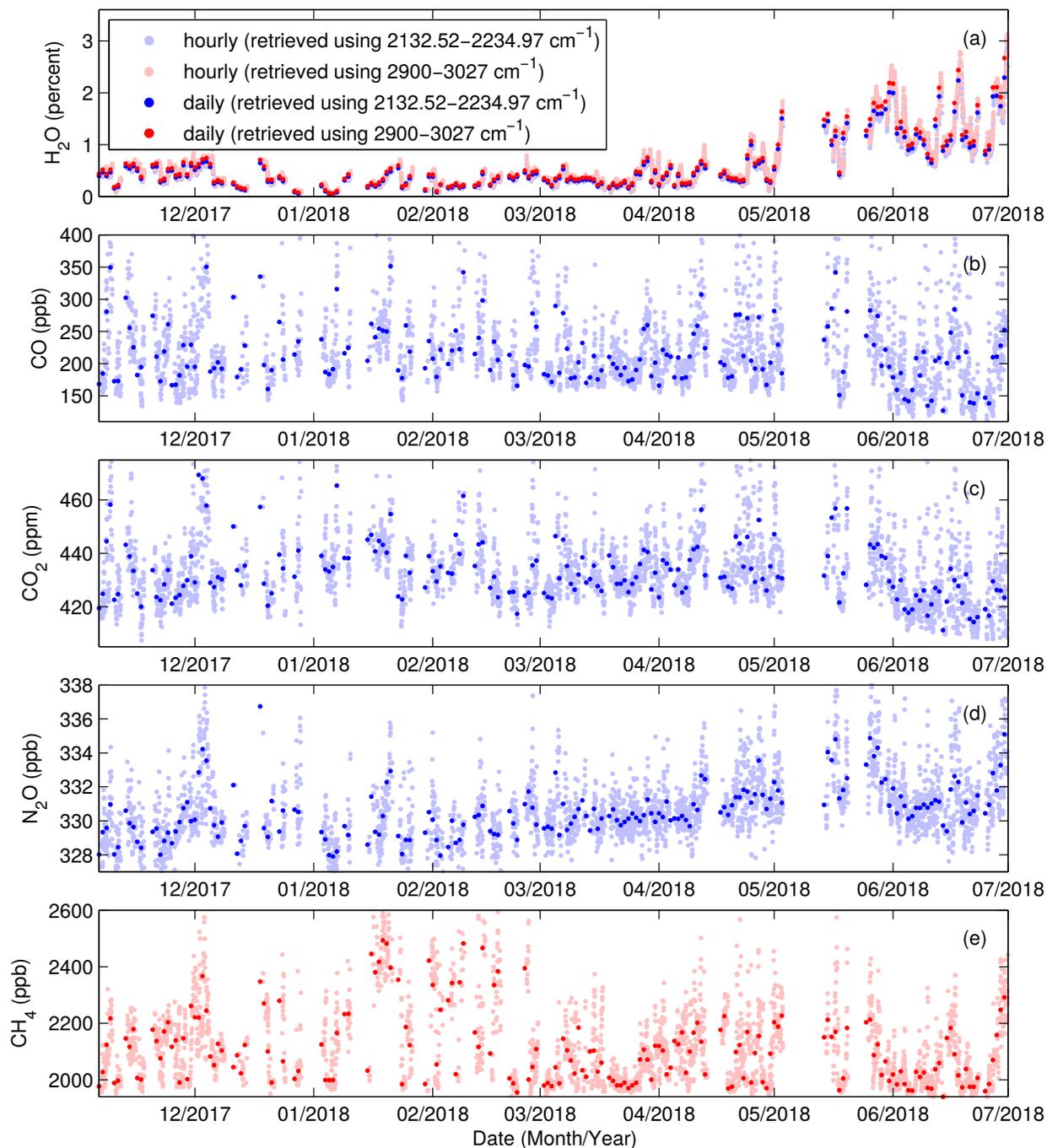


Figure 6.11: Time series of (a) H₂O (b) CO, (c) CO₂, (d) N₂O, and (e) CH₄ dry-air mole fractions retrieved from OP-FTIR measurements. Gases are shown in blue for retrievals using 2132.52–2234.97 cm⁻¹ and red for 2900–3027 cm⁻¹. Pastel colors indicate hourly means and darker colors indicate daily means.

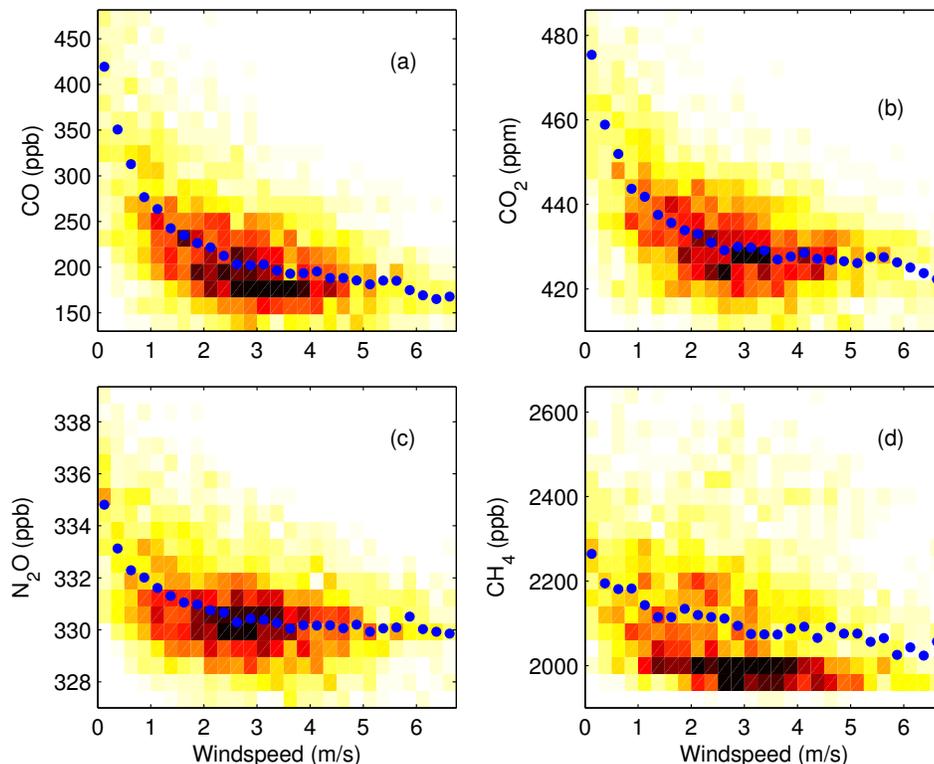


Figure 6.12: Number density for OP-FTIR measurements of (a) CO, (b), CO₂, (c) N₂O, and (d) CH₄ dry-air mole fractions as a function of wind speed for November 2017 through June 2018. Blue dots show the mean dry-air mole fractions for each wind speed. Meteorological data is taken from the McLennan met-station.

direction dependence, with enhanced dry-air mole fractions for wind from the northeast. The cause of this wind direction dependence is unclear. CH₄ only shows weak wind speed dependence, but strong wind direction dependence with enhanced dry-air mole fractions for wind from the southwest. This suggests that there is a large spatially localized source to the southwest of the OP-FTIR.

6.7.2 Comparison to Hanlan’s Point station

ECCC runs an in-situ GHG observing site, called Hanlan’s Point station, approximately 5.4 km south of the OP-FTIR system (Figure 6.14). In this section, the gradients in CO, CO₂, and CH₄ between the OP-FTIR and ECCC’s Hanlan Point station are investigated for Nov 6, 2017–June 30, 2018.

Hanlan’s Point station is on the southern edge of the city of Toronto. Therefore, measurements at this site are not influenced by urban emissions when the air mass is advected over the site from the south. Figure 6.15 shows the gradient in CO, CO₂,

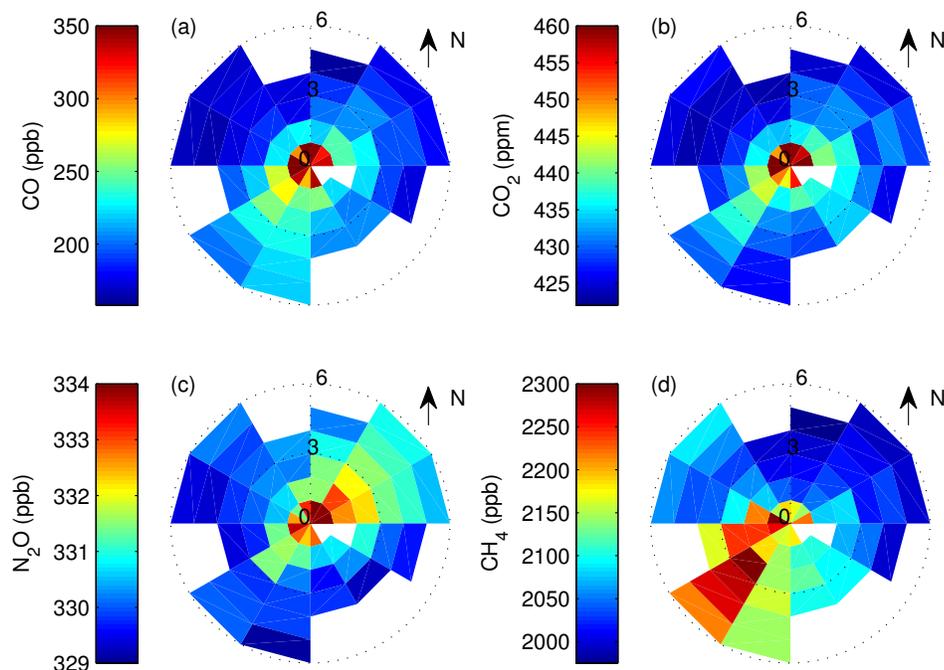


Figure 6.13: Pollution roses for OP-FTIR measurements of (a) CO, (b), CO₂, (c) N₂O, and (d) CH₄ spanning November 2017 through June 2018. Meteorological data is taken from the McLennan met-station.

and CH₄ between the OP-FTIR and Hanlan's Point station (OP-FTIR minus Hanlan's Point) as a function of wind speed and direction. Meteorological data was taken from the McLennan met-station. To ensure that the wind direction measurements were reliable, data points were only included in the plot if the wind direction at the Toronto Island ECCC met-station and McLennan met-station were in within the same 45° wind direction bin. For all gases and wind conditions, the gradient between sites is generally greater than zero. This is expected because the OP-FTIR is closer to anthropogenic emission sources than Hanlan's Point station. However, there may also be a positive bias in the OP-FTIR measurements which increases the gradient. For CO and CO₂, the gradient between the sites is generally larger if the wind is from the south. This is expected because air from the south should not be impacted by Toronto emissions when passing over Hanlan's Point. For CH₄, the gradient is largest when the wind is from the southwest. This is likely due to the fact that the OP-FTIR measurements seem to be influenced by a highly localized CH₄ source as discussed above.

To isolate the atmospheric imprint of Toronto GHG emissions on these two sites, the gradient between the sites can be isolated for wind from the north and from the south. This is obtained by calculating the gradient between the sites for all hours in

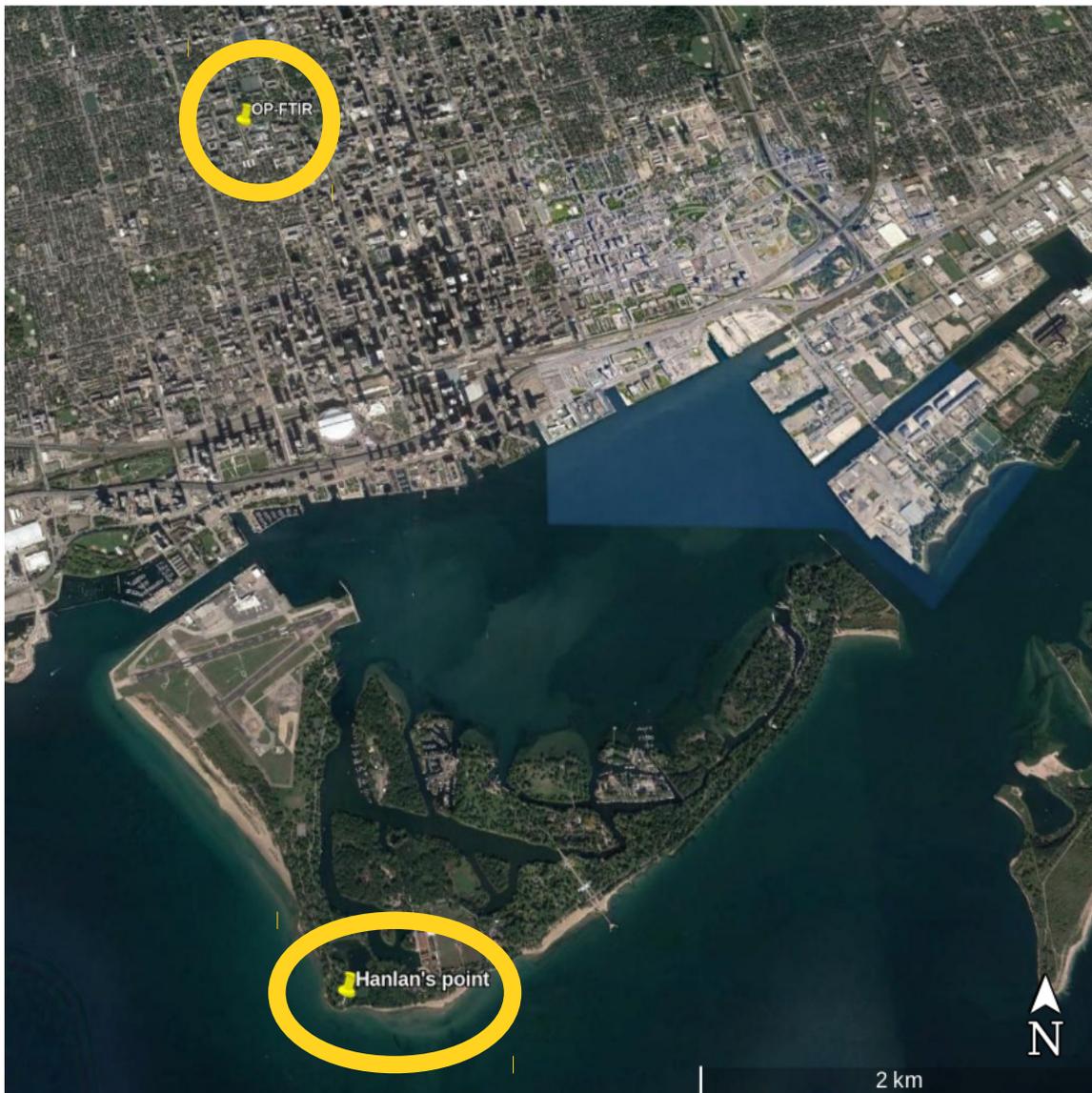


Figure 6.14: Locations of the OP-FTIR and Hanlan's Point station within Toronto, Ontario. Google Maps (2018).

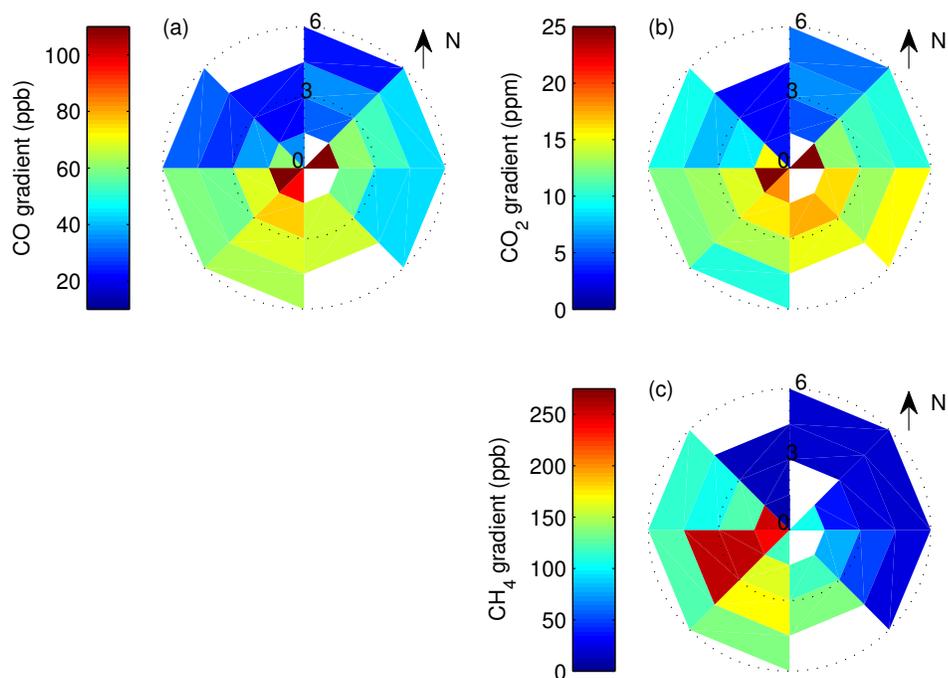


Figure 6.15: Pollution roses of the gradient between the OP-FTIR and Hanlan's Point for (a) CO, (b), CO₂, and (c) CH₄ spanning November 2017 through June 2018. Meteorological data is taken from the McLennan met-station, but the wind direction must also be in the same within the same 45° wind direction bin for the Toronto Island ECCC met-station.

which the wind is from the north ($\pm 45^\circ$) or south ($\pm 45^\circ$) and the wind speed is greater than 1.5 m s^{-1} . Note that the results were not sensitive to small changes in wind direction/speed cutoff. Figure 6.16 shows the resulting histogram of the gradients between the sites when these criteria are met. In general, the gradient between sites is significantly larger for wind from the south relative to wind from the north. This demonstrates that Toronto emissions significantly impact the gradient in GHG dry-air mole fractions as air-masses from the south move over these sites. However, there is also significant variance in the gradient between the sites after accounting for wind direction. This could be due to a number of factors, including the stability of the boundary layer and wind speed. If surface fluxes are to be estimated using the dry-air mole fractions gradient between these sites, then these additional factors will need to be accounted for.

6.7.3 Potential for flux estimation

In this section, the plausibility of estimating surface emissions from the gradient in GHG dry-air mole fractions between the OP-FTIR and Hanlan's Point is discussed. The simplest method for estimating emissions based on the GHG gradient would be to use mass balance. Take the continuity equation for conservation of mass of tracer C_{gas} (Eq. 2.1) and assume there is no divergence:

$$\frac{\partial C_{\text{gas}}}{\partial t} + \nabla \cdot C_{\text{gas}} \mathbf{v} \stackrel{0}{=} F, \quad (6.12)$$

where $C_{\text{gas}} = \rho q_{\text{gas}}$ and F is the surface flux (in units of g s^{-1}) occurring at the surface over area A . This equation can then be re-written as:

$$F = \frac{\partial C_{\text{gas}}}{\partial t} = \frac{\partial C_{\text{gas}}}{\partial x} \frac{dx}{dt} = \frac{\partial C_{\text{gas}}}{\partial x} U. \quad (6.13)$$

where $U = dx/dt$. If C_{gas} is in units of grams of CO_2 in a boundary layer column, then it can be expressed as a function of observable quantities:

$$C_{\text{gas}} = \langle S_{\text{gas}} \rangle \frac{M_{\text{gas}}}{M_{\text{air}}} \rho_{\text{dry}} A h, \quad (6.14)$$

where $\langle S_{\text{gas}} \rangle$ is the bulk boundary layer dry-air mole fraction of a given gas, M_{gas} is the molar mass of the gas of interest, M_{air} is the molar mass of dry air, ρ_{dry} is the air density, A is the surface area, and h is the boundary layer height. Substituting Eq. 6.14 into

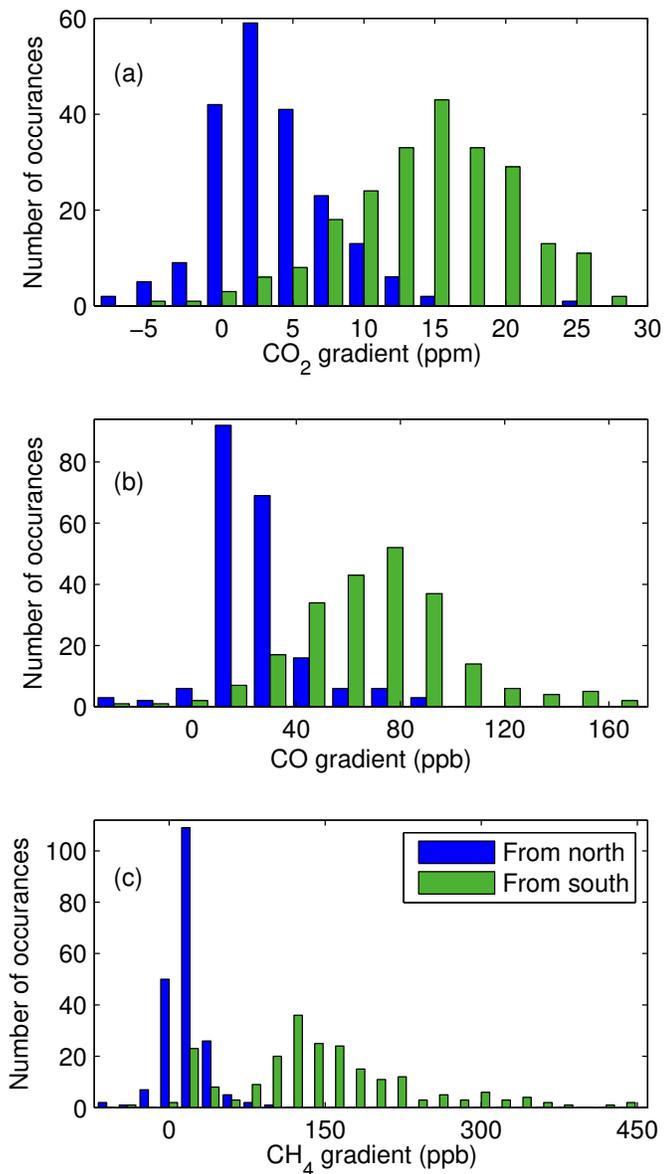


Figure 6.16: Number of occurrences of hourly gradients in (a) CO₂, (b) CO, and (c) CH₄ for wind from the north (blue) and from the south (green) and for wind speeds between 1.5 m s⁻¹ and 4.5 m s⁻¹.

Eq. 6.13 gives:

$$F = \frac{\partial C}{\partial x} U = \frac{\rho_{dry} A h U M_{gas}}{M_{air}} \frac{\partial \langle S_{gas} \rangle}{\partial x} \approx \frac{\rho_{dry} A h U M_{gas}}{M_{air}} \frac{\Delta \langle S_{gas} \rangle}{\Delta x}. \quad (6.15)$$

The flux per unit area can be written as:

$$\frac{F}{A} = \frac{\rho_{dry} h U M_{gas}}{M_{air}} \frac{\Delta \langle S_{gas} \rangle}{\Delta x} = \frac{\rho_{dry} M_{gas}}{M_{air}} \left[h U \frac{\Delta \langle S_{gas} \rangle}{\Delta x} \right]. \quad (6.16)$$

Note that the quantities in Eq. 6.16 represent bulk boundary layer values. h can be estimated by a numerical model or observed by a lidar. U can be approximated based on the wind measurements from the McLennan met-station. $\Delta \langle S_{gas} \rangle$ can be estimated based on the GHG gradient between the OP-FTIR and Hanlan's Point. If we take typical values for these quantities and plug them into Eq. 6.16, we get

$$\frac{F}{A} = \frac{(1225 \text{ g m}^{-3})(44 \text{ g mol}^{-1})}{(28 \text{ g mol}^{-1})} \left[(1000 \text{ m})(4 \text{ m s}^{-1}) \frac{15 \times 10^{-6}}{5400 \text{ m}} \right] \quad (6.17)$$

$$= 2.1 \times 10^{-2} \text{ g m}^{-2} \text{ s}^{-1} \quad (6.18)$$

For comparison, Pugliese et al. (2018) estimate CO₂ emissions of $5\text{--}10 \times 10^{-4} \text{ g m}^{-2} \text{ s}^{-1}$ for downtown Toronto based on bottom-up estimates. Thus, this calculation seems to overestimate Toronto emissions by a factor of 20–40. It is challenging to isolate the exact cause of the large error in the estimated flux as there are many possible sources of error. One source of error is that the gradient between the OP-FTIR and Hanlan's Point is not representative of the boundary layer mean gradient. Based on Eq. 1.7, the observed gradient should be expected to be ~ 5 times larger than the bulk boundary layer gradient. Furthermore, the OP-FTIR system is within 20 m of a road which could enhance CO₂ dry-air mole fractions above background city emissions. High-resolution atmospheric transport models which can resolve atmospheric transport within street canyons, combined with inversion methods, will likely be required to produce robust flux estimates based on these data.

6.8 Conclusions

An OP-FTIR system has been installed on the St. George Campus of the University of Toronto. The system has been used to retrieve CO₂, CO, N₂O, and CH₄ dry-air mole fractions since 6 November 2017, producing a time series with over 1500 measurements. Comparisons with meteorological observations have shown that retrieved gas

dry-air mole fractions are sensitive to urban emissions from Toronto. CO_2 , CO , and N_2O appear to be influenced by large-scale urban emissions, while CH_4 observations are contaminated by a localized source southeast of the observing system, presumably a natural gas leak. By performing comparisons with nearby measurements from Hanlan's Point, it was demonstrated that the gradient in CO_2 and CO measurements between these sites contains information on urban emissions which may allow for flux estimation. However, preliminary flux calculations result in an unrealistically large source of CO_2 relative to bottom-up estimates.

Chapter 7

Conclusions

The magnitude of future climate change is dependent on the magnitude of changes in GHGs. Predicting the future rise in atmospheric CO₂ will require both CO₂ emission monitoring and TBMs that can reliably simulate uptake by the terrestrial biosphere. This thesis addressed the challenge of diagnosing changes in atmospheric GHGs. In particular, the following two questions were addressed:

1. How can deficiencies in TBMs be identified from top-down constraints?
2. Can an OP-FTIR system monitor urban GHG emissions in Toronto?

7.1 Identifying deficiencies in TBMs

To address the first question, three studies were undertaken:

- In Chapter 3, the sensitivity of NEE constraints to the observational coverage of several observing systems was examined. It was shown that the spatiotemporal differences in observational coverage between observing systems can have a large impact on estimated fluxes. This implies that careful consideration is required in interpreting constraints on NEE from any given observing system.
- In Chapter 4, observational constraints from atmospheric CO₂ and SIF observations were combined to evaluate the seasonality of NEE, GPP and R_e fluxes over the northern mid-latitudes for a set of TBMs. It was found that the subset of models which assimilated phenology observations produced realistic GPP seasonality, however, these model were systematically biased in the seasonality of R_e . This bias was attributed to parameterizations of R_e within the models.

- In Chapter 5, GOSAT flux inversions were shown to have strong agreement in the timing of anomalies with MERRA T_{soil} and FLUXCOM NEE. These results showed that GOSAT flux inversion largely recover NEE anomalies in the tropics on continental and larger scales, and in the northern extratropics on sub-continental scales during the summer. Furthermore, NEE monthly anomalies were generally better captured by the GOSAT inversions than TBMs, indicating that the inversions can be used to evaluate IAV in TBMs.

These studies demonstrate that atmospheric CO₂ observations provide a constraint on NEE fluxes that is sufficient to evaluate the seasonal cycle and IAV in TBMs. Furthermore, combining these constraints with SIF observations or temperature anomalies can elucidate the origin of errors within TBM parameterizations. This type of analysis provides an effective application for atmospheric CO₂ constraints, which has been underappreciated within the flux inversion community as only a few of previous studies have directly evaluated TBMs (Messerschmidt et al., 2013; Peng et al., 2015; Keppel-Aleks et al., 2014b). Thus, there are many opportunities to build on the results presented here.

In this thesis, it was shown that the CO₂ flux constraints provided by an observing system should be considered within the spatiotemporal coverage of the observations. Currently, only limited constraints can be obtained in the extratropical winter from space-based observing systems, and surface observations are too sparse to provide constraints in the tropics. There is a need to combine the atmospheric CO₂ constraints from different observing systems to provide constraints on NEE with more uniform spatiotemporal sensitivity. In particular, combining space-based measurements with TCCON and surface measurements could provide increased observational constraints on the seasonal cycle in the northern extra-tropics. However, combining measurements from different observing systems with a single inversion framework is challenging, as the relative weighting which should be applied to different observing systems can be difficult to diagnose.

Based on the results of this thesis, several different approaches could be undertaken to evaluate TBMs in future studies. Combining constraints from atmospheric CO₂ and SIF observations was found to be a powerful approach to evaluate model GPP and R_e fluxes. This type of analysis could be extended to smaller scales if NEE flux estimates are robust on these scales. However, the minimum scale at which flux inversions can provide constraints on the seasonal cycle is not well documented, but is required to extend this analysis. A long-term objective of this research should be the development of data assimilation systems which can assimilate both atmospheric CO₂ and SIF observations. In the future, assimilation systems that utilize physical forward model relating GPP and observed SIF (e.g., Van der Tol et al., 2014) could improve constraints on GPP.

Another avenue of future research would be to examine flux anomalies in more detail. Recently, there has been substantial interest in looking at NEE anomalies from flux inversions (Liu et al., 2017; Bowman et al., 2017; Rödenbeck et al., 2018). However, these constraints have not been applied to evaluate TBMs. Future research could evaluate the response of TBM GPP, R_e , and NEE anomalies to climate anomalies, following a similar set-up to Liu et al. (2017). However, caution would be required to ensure that these events are well captured by the flux inversions. Based on the results found here, this type of analysis should be on continental or larger scales in the tropics or on sub-continental or larger scales in the northern extratropics during the summer.

7.2 Monitoring emission targets

To address the second question of this thesis, one study was undertaken:

- In Chapter 6, the installation of a OP-FTIR system in downtown Toronto was described and timeseries of CO₂, CO, CH₄, and N₂O were retrieved. Comparisons of the observed gases with meteorological observations were presented, as were comparisons with observed CO₂, CO, and CH₄ at ECCC's Hanlan's Point station.

The OP-FTIR system presented in this thesis will provide GHG measurements as an important component of a larger Toronto-wide network of observing sites. The OP-FTIR system is currently highly automated with minimal support needed to refill the LN₂ dewar and close the system down during bad weather. From the preliminary analysis presented in this thesis, three major issues have been identified which should be top priorities for future work:

1. Reliable meteorological data.
2. Mechanism to calibrate the OP-FTIR measurements.
3. Relating measured CO₂ to the boundary layer mean mole fraction.

Currently, the met-station is located on the southeast corner of the roof of the Burton tower. Based on comparisons with other nearby met-stations, there are likely artifacts in the wind data due to nearby obstacles. It is recommended that this station be raised on a pole (by several meters, if possible). This should help minimize the impact of nearby obstacles on the observed wind fields. It would also be useful to install a second met-station at the location of the retro-reflector (also raised several meters from the building roof). These steps would provide more reliable meteorological data.

Currently, there is not a reliable method to frequently calibrate the OP-FTIR system. CO_2 and CH_4 can be compared against the Los Gatos Research (LGR) multi-gas analyser, an in situ system which can in-turn be calibrated against a standard gas. If performed regularly, this could provide a good mechanism to calibrate the gases. An initial attempt at calibration is described in Appendix B. The LGR is unable to measure CO or N_2O . To calibrate N_2O observations, it may be possible to use an observing system that is currently under development and will be deployed for balloon-borne observations [personal communication, John Saunders, 2018]. In the future, an ideal calibration system could be obtained by installing in-situ systems at both the telescope and retro-reflector. This set-up would have the added benefit of allowing the investigation of the variability of GHG concentrations between the in-situ and OP-FTIR systems.

Besides these high priority issues, there are several optional changes to the OP-FTIR system which could be further investigated. The retro-reflector could be relocated to increase the measurement pathlength. Some tests performed by moving the retro-reflector over a range of distances suggests that the pathlength could be doubled. The main challenge in extending the pathlength is finding a suitable location for the retro-reflector. Ideally, the entire path should be distant from surfaces and any vents on buildings. The retro-reflector should also be raised by several stories from street level to minimize the impact of localized sources. Another possibility would be to change the optical set-up of the system. If the source were directed through the interferometer before being sent to the retro-reflector, the ambient (lowered shutter) spectrum would no longer be required and the frequency of observation could be doubled. Alternatively, the source could be set-up at the position of the retro-reflector and collimated such that only a one-way path is used, which would be expected to increase the signal. However, before changing the optical set-up, it should be noted that the current system was challenging to get running smoothly and these changes could take a long time to fully implement. A cost benefit analysis should be carefully thought through before changing the set-up.

There may be an opportunity to retrieve more trace gases from the recorded spectra. Some attempts were made to retrieve O_3 and C_2H_6 , but the signal was found to be insufficient. The feasibility of retrieving isotopologues has not been fully investigated, but could be a productive area of future research. There appears to be promise in retrieving HDO/ H_2O fractionation (which will be reported in a research paper using OP-FTIR observations), but isotopologues of long-lived atmospheric species have not been performed to a useful precision. Retrieving isotopologues of CO_2 and CH_4 would be useful for source attribution (e.g., Pataki et al., 2003; Newman et al., 2016). If there is insufficient signal to retrieve isotopologues with the current set-up, extending the

pathlength and/or installing a brighter source could help.

The discussion so far has focused on future work in terms of data collection, however, there are also opportunities for future work in estimating fluxes. Estimating surface fluxes from the OP-FTIR system alone is not feasible, instead, these observations should be combined with background observations from existing observing systems. ECCC's Hanlan's Point station likely offers the best site for comparisons. By examining gradients in the observed concentrations between sites, surface fluxes can be inferred. However, there are also a number of challenges associated with this. In particular, the vertical structure of the concentration field needs to be well understood to infer emissions. This will likely require a tracer-transport model with high vertical resolution within the boundary layer. Currently, the best option to estimate concentration gradients would be to use the Environment Canada Carbon Assimilation System (EC-CAS) on the nested Pan-Am domain, providing $2.5 \text{ km} \times 2.5 \text{ km}$ horizontal resolution. Another possibility would be to have a free-running version of Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) and impose boundary conditions from EC-CAS. The longer-term goal should be to use a high-resolution model coupled with an inversion method to estimate fluxes. This could be done in collaboration with other scientists working on Toronto emissions (e.g., Felix Vogel, ECCC; Debra Wunch, University of Toronto; Jennifer Murphy, University of Toronto).

Appendix A

Appendix

A.1 GOME-2 SIF errors

The precision of GOME-2 SIF observations is quite low, with single measurement errors comparable to the retrieved SIF value. Therefore, large numbers of observations need to be combined to average out random errors. Here we show the presence of random errors between years and thus demonstrate that averaging between years is required to calculate a precise mean seasonal cycle, using the NASA GOME-2 SIF data.

As discussed in the main text, we first aggregated the observations spatially to a $2^\circ \times 2.5^\circ$ grid and temporally to week of year by calculating the median value. However, even with this averaging, the random error remain substantial. This is demonstrated for evergreen needleleaf forests in Fig. A.1. There are substantial random fluctuations in the timeseries for any given year, which originate from random error in the observations. If all of these years are averaged together, then the random fluctuations are largely eliminated. The year-to-year variability in the signal (SIF emissions) will also be present between years and contributes to the variance estimate.

A.2 Seasonality of SIF and TBMs

Figure A.2 shows the start, end, and length of the growing season, where the start and end of the season are determined by the date that the GPP (or SIF) reaches 25% of the peak value. The two sets of fluxes generated with CTEM can have quite marked differences. The start of the growing season for CTEM-CRU GPP is within one week of SIF across the vegetation types, whereas the start for CTEM-GEM is one to three weeks early. The end of the growing season is coincident for CTEM-CRU and CTEM-GEM for

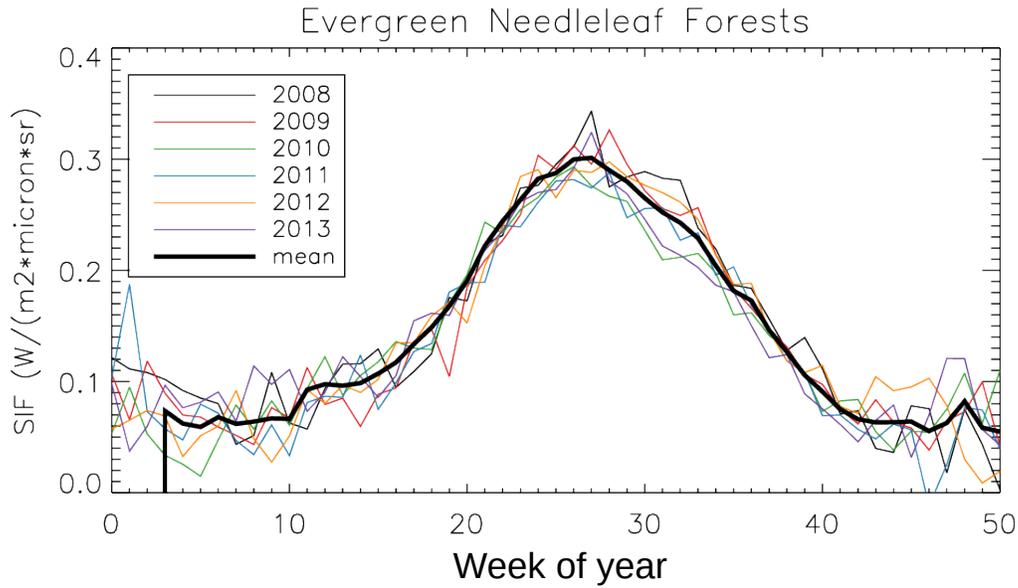


Figure A.1: Seasonal cycle of NASA GOME-2 SIF over the years 2008-2013 for the ENF vegetation region. Also plotted is the mean seasonal cycle averaged over this time period.

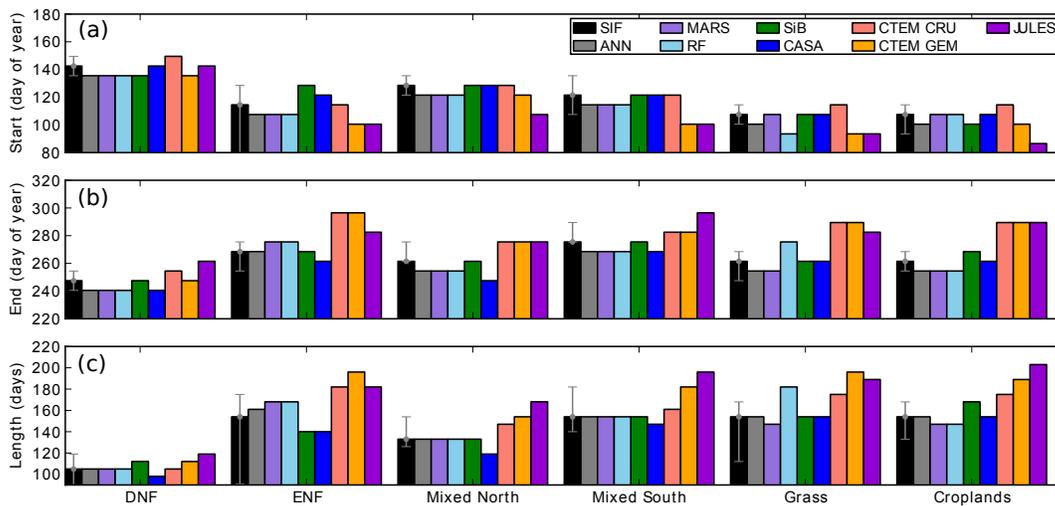


Figure A.2: (a) Start of season (day when SIF (GPP) reaches 25% of the maximum value), (b) end of season (day when SIF (GPP) reaches below 25% of the maximum value), and (c) length of the season (in days).

all but one vegetation type. However, the end of the growing season in the CTEM fluxes are up to five weeks later than in the SIF data. This results in CTEM fluxes predicting a longer growing season than SIF. On average, across the vegetation types, CTEM-CRU and CTEM-GEM predict growing seasons which are two weeks and four weeks longer than that suggested by the SIF data, respectively. JULES GPP has the same systematic differences relative to NASA GOME-2 SIF across all vegetation regions except for DNF. Over these regions, JULES predicts an earlier start (about two and a half weeks) and a later end (about three weeks) of the growing season. For DNF, the start of growing season is coincident for JULES GPP and the SIF data, however, the end of the growing season is delayed by a few weeks in JULES.

A.3 Model transport errors

The magnitude of the error associated with transport is difficult to quantify, although some insights can be garnered from previous studies. Houweling et al. (2010) performed an intercomparison of transport models using the same set of surface fluxes. They found that forward simulations of X_{CO_2} lead to mean model-to-model differences of 0.5 ppm over the continents and 0.27 ppm over the oceans. Looking at the seasonal cycle of X_{CO_2} at Park Falls, Basu et al. (2011) attempted to isolate the magnitude of transport error in their model-data mismatch estimated (based on TM5). They found that model transport errors combined with errors in calculating X_{CO_2} resulted in an error of less than or equal to 1 ppm in the amplitude of the seasonal cycle. We can also get some insights into the magnitude of transport errors from the model-data mismatch using CT2016 NEE.

We examine the difference in the mean seasonal cycle of X_{CO_2} simulated by GEOS-Chem and TM5 at Sodankylä, Białystok, Orléans and Park Falls. We use CT2016 NEE as the biosphere fluxes in both simulation, therefore, difference in the mean seasonal cycles of X_{CO_2} from GEOS-Chem and TM5 should largely be due to differences in model transport.

However, there are some differences in the method used to calculate mean seasonal cycle of X_{CO_2} from TM5 in comparison to GEOS-Chem. First, we use TM5 X_{CO_2} simulated from year specific CT2016 NEE fields rather than mean NEE. Second, additional data filtering is performed when calculating the mean seasonal cycles of X_{CO_2} for TM5. Only TCCON retrievals with a solar zenith angle of less than or equal to 60° are retained. If a day has fewer than 50 retrievals left, no data are processed from that day. There is also an outliers detection scheme involving a quadratic fit to time for each day's retrieval. Any retrieval whose residual X_{CO_2} from this fit is more than three standard deviations

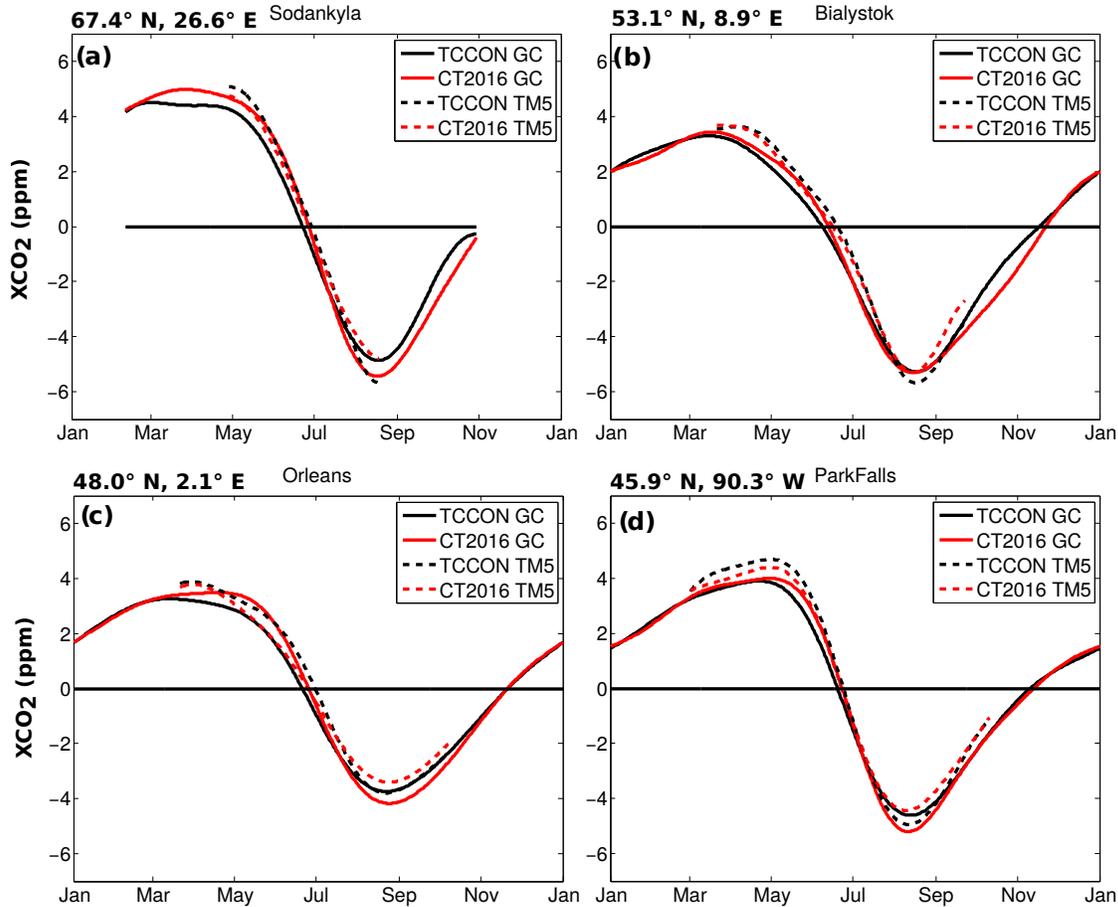


Figure A.3: Mean X_{CO_2} seasonal cycle at a) Sodankylä, b) Bialystok, c) Orléans and d) Park Falls. Solid lines show the mean X_{CO_2} seasonal cycle from TCCON observations (black) and simulated by GEOS-Chem (red) over the period 2008-2012 with minimal data screening. Dashed lines show the mean X_{CO_2} seasonal cycle from TCCON observations after additional data screening was performed (black) and simulated by TM5 (red) over the period 2008-2014 (also with additional data screening).

from the curve is rejected, and the procedure is iterated with a new fit until there are no such outliers remaining. Finally, we create 30-minute averages if and only if there are more than 10 retrievals within the 30-minute interval. Due to this additional filtering, the number of data points is significantly reduced. To achieve a good seasonal cycle fit of the simulated X_{CO_2} , we fit the years 2008-2014, which is two years longer than for the GEOS-Chem fits. We also fit a mean seasonal cycle to the TCCON observations which have undergone this additional data filtering.

Figure A.3 shows the mean seasonal cycle of X_{CO_2} for TCCON and simulated by GEOS-Chem and TM5 using CT2016 NEE at the four TCCON sites examined in this study. Due to differences in data filtering, there are differences between the mean seasonal

cycles calculated from the TCCON observations of up to 1 ppm at the same TCCON site. Differences between GEOS-Chem and TM5 mean seasonal cycles are of the same magnitude as these differences, thus it is difficult to determine whether these differences are primarily due to transport or data filtering. Nevertheless, differences in the mean X_{CO_2} seasonal cycle are always less than or equal to 1 ppm. This suggests that transport errors are likely smaller than 1 ppm, consistent with previous studies (e.g., Houweling et al., 2010; Basu et al., 2011).

We should also note that the X_{CO_2} seasonal cycle at these TCCON sites is not solely driven by NEE in the vegetation regions examined here, but is also dependent on NEE outside these regions (Keppel-Aleks et al., 2011, 2012). For example, Barnes et al. (2016) have shown that the X_{CO_2} seasonal cycle over northern mid-latitude regions is influenced by NEE throughout the midlatitudes as the CO_2 signal is transport upward and poleward by isentropic transport. To ensure that the X_{CO_2} seasonal cycle at the four TCCON sites examined here are driven by NEE in the regions of interest we performed tagged tracer runs with the GEOS-Chem forward model at $2^\circ \times 2.5^\circ$ resolution. In these runs, all of the NEE between 39° and 65° N is tracked so that its contribution to the seasonal cycle can be calculated. We find that NEE in these regions is responsible for 76% (Sodankylä) to 87% (Białystok) of the X_{CO_2} seasonal cycle at these TCCON sites.

A.4 SiB3 and CASA comparison for individual vegetation types

Figure A.4 shows the day of the year when the TBM GPP, R_e and NEE drawdown reach peak values for each vegetation region in SiB3 and CASA. Examining the timing of peak GPP, R_e and NEE drawdown for SiB3 reveals that (except for DNF) peak NEE drawdown occurs earliest followed by peak GPP and then R_e (between day 225–235). For CASA, peak R_e typically occurs around the same day of the year as peak GPP (except for DNF), whereas peak NEE typically occurs after peak GPP and R_e (except for DNF and southern mixed forests). Comparing SiB3 and CASA, we find the date of peak GPP is similar for both TBMs, as indicated by comparisons with SIF, but peak R_e and NEE are significantly different between the TBMs. Peak R_e for CASA occurs three to four weeks earlier than for SiB3 (except for ENF), and thus peak NEE occurs two to five weeks later. The differences between SiB and CASA found here are consistent with the results of Messerschmidt et al. (2013), who found that differences in seasonality of NEE between TBMs were primarily due to the differential phasing of R_e with respect to GPP.

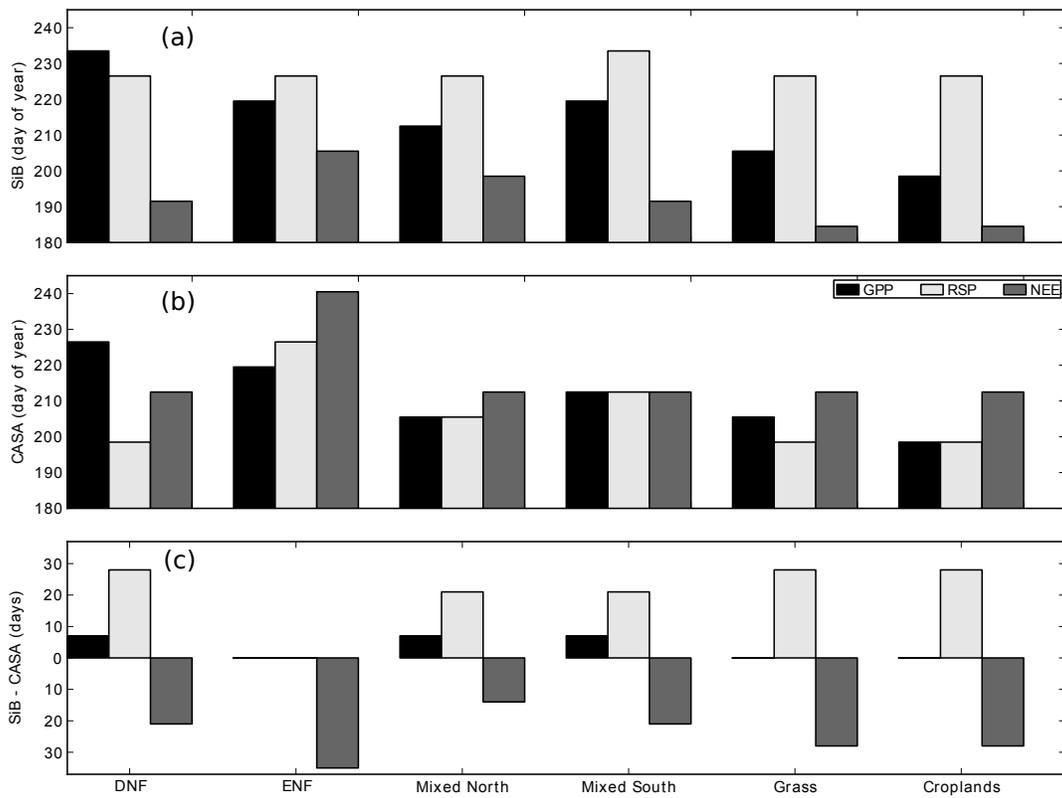


Figure A.4: Day of year at which the peak GPP (black), peak R_e (light gray) and peak NEE (dark gray) occur for (a) SiB3 and (b) CASA, and (c) the difference in timing between SiB3 and CASA.

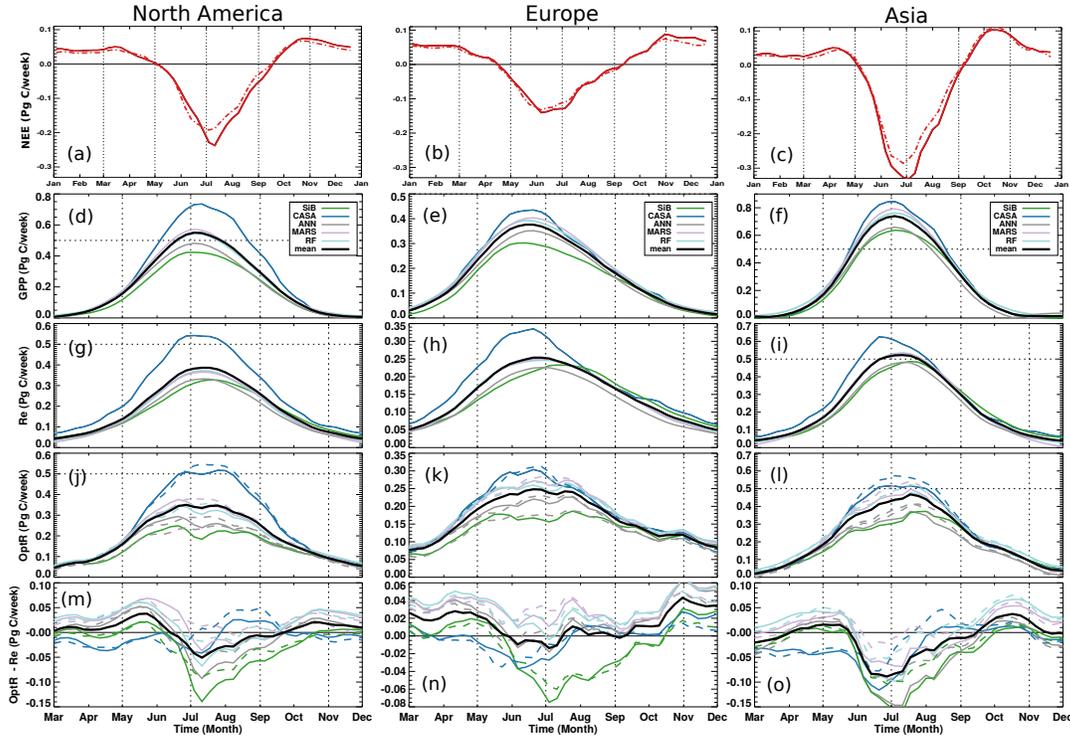


Figure A.5: Seasonal cycles of (a–c) inversion NEE (d–f) model GPP, (g–i) model R_e , (j–l) $optR_{inv-mod}$, and (m–o) the difference between $optR_{inv-mod}$ and model R_e for (left column) North America, (middle column) Europe, and (right column) Asia. For subplots a–c, CT2016 NEE is the solid line and GOSAT-Inv is the dash-dot line. For subplots d–f, $optR_{CT2016-mod}$ is represented by solid lines, whereas $optR_{GOSATinv-mod}$ is indicated by dashed lines. In all panels the solid, heavy black line represent the mean of all the curves shown.

A.5 Continental scales

Figure A.5 shows inversion NEE, model GPP, model R_e , and $optR_{inv-mod}$ for North America, Europe and Asia. As was found for the entire northern extratropics, there are large differences in the magnitude of GPP between models. The difference is particularly large for North America, where CASA GPP is much larger than SiB3 GPP and FLUXCOM GPP.

Appendix B

Initial side-by-side comparisons of the OP-FTIR and LGR

We performed initial side-by-side comparisons between the OP-FTIR and LGR, an in situ analyzer that has been calibrated against a standard gas. We first performed measurements with the LGR situated on the 15th floor balcony of the Burton Tower (Fig. B.1) and then with the LGR situated on the Galbraith roof near the retro-reflector (Fig. B.2). The results suggest that CO₂ and CH₄ fields have a significant amount of spatial heterogeneity over the OP-FTIR path. We find poorer agreement for CO₂ than CH₄. For CO₂, there is significant variability measured by the LGR at the position of the retro-reflector, suggesting that nearby emissions (presumably vehicular emissions) are resulting in large mole fraction enhancements. In contrast, the LGR CO₂ measurements on the Burton Tower balcony indicate that there is sensitivity to boundary layer dynamics, with lower CO₂ mole fractions observed in the afternoon (when vertical mixing is strongest). Presumably, the OP-FTIR is measuring a combination of these signals, which results in poor agreement with in situ measurements at both locations. For CH₄, agreement between the OP-FTIR and LGR is better. There is a low bias in OP-FTIR measurements relative to LGR measurements on the Burton Tower balcony, however, this could be explained differences in the sensitivity to the nearby CH₄ localized source.

APPENDIX B. INITIAL SIDE-BY-SIDE COMPARISONS OF THE OP-FTIR AND LGR173

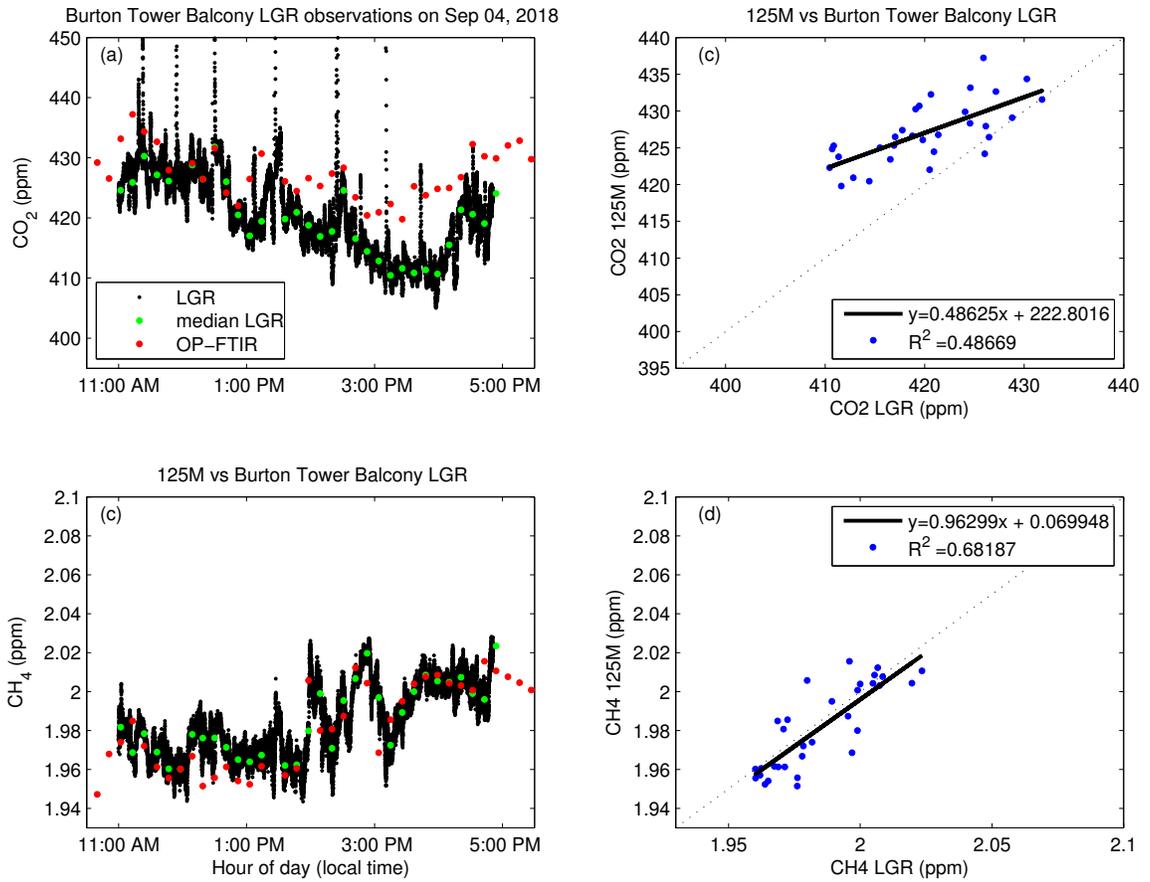


Figure B.1: Measurements of (a) CO₂ and (c) CH₄ from the (red) OP-FTIR and (black) 3 second averages and (green) 6 minute medians with the LGR positioned on the Burton Tower balcony on September 4th, 2018. Scatter plot and linear regression of OP-FTIR measurements against LGR measurements for (b) CO₂ and (d) CH₄.

APPENDIX B. INITIAL SIDE-BY-SIDE COMPARISONS OF THE OP-FTIR AND LGR174

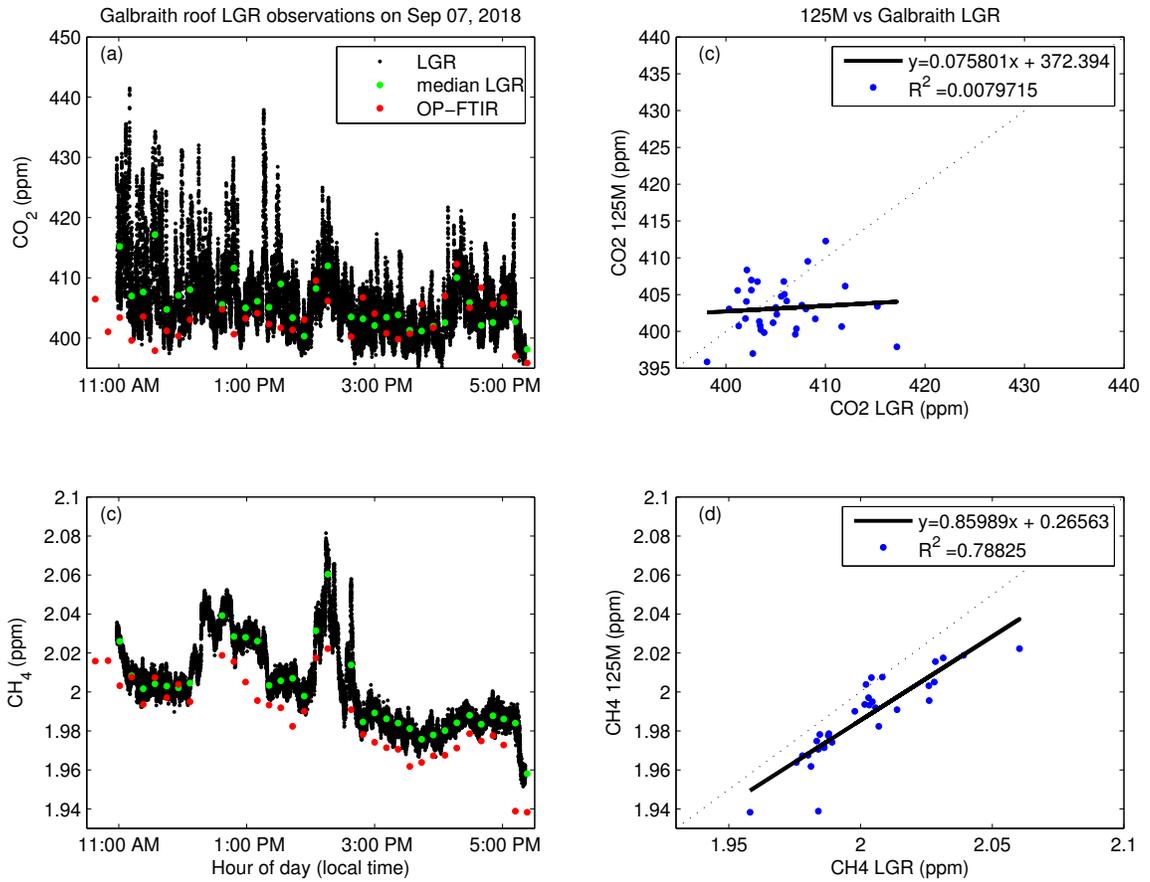


Figure B.2: Measurements of (a) CO_2 and (c) CH_4 from the (red) OP-FTIR and (black) 3 second averages and (green) 6 minute medians with the LGR positioned on the Galbraith roof on September 7th, 2018. Scatter plot and linear regression of OP-FTIR measurements against LGR measurements for (b) CO_2 and (d) CH_4 .

Appendix C

Processing of OP-FTIR data

This appendix provides some guidance on how to process the raw meteorological and OPUS files to perform the retrievals. This an outline and some changes may need to be made to the programs depending on the specific retrieval that is being performed.

C.1 Matlab processing

A matlab program (“/export/data/home/bbyrne/make_test_shutter.m”, or similar) is used to create input text files for the python code and MALT code. These text files are:

1. “test1.txt”, a python code that subtracts the shutter raised and lowered spectra.
2. “malt_input1.txt”, the input file to MALT. There are two versions of this program: one that subtracts the preceding lowered shutter spectrum (/home/bbyrne/make_test_shutter.m) and one that subtracts the lowered shutter spectrum measured after the open shutter measurement (/home/bbyrne/make_test_shutter_opp.m).

Before running the matlab code, the met fields and spectral data files need to be moved to the right locations. First, for the met fields:

- The met fields need to be copied over from “/net/deluge/pb-1/projects...
.../TAO/data/BomemDA8/Sun/YYYY_MM/Auxiliary/mmm_Log_Folder/mmmYYlog.txt”
to
“/data/01/bbyrne/Malt5_output”,
- the header needs to be removed and “:” and “/” need to be converted to commas.

- Note that the met-station records are in UTC and the OP-FTIR records in local time, so the difference is accounted for by this program (including daylight savings).

Second, for the data files:

- Raw data are stored in:

`/net/deluge/pb_1/projects/TAO/Equipment/spectrometers/bruker/ifs125m/data/.`

- Copy the desired files over to

`/data/01/bbyrne/dataSource/new/.`

Now, the matlab code “make_test_shutter.m” can be run. Before running make sure the adjustments are made to read in the desired met files.

C.2 Subtract shutter raised and lowered spectra

Now you need to run a python code which will subtract the shutter open and closed spectra.

- “cd” into “/data/01/bbyrne/test_OPUS_calc”.
- Run “python handler_mine_40scan_04res.py” (or for the “opp” files run “handler_min_40scan_04res_opp.py”).

These programs will output OPUS files into the same directory, these are the source on–source off files, and are the files that the retrieval is performed on.

C.3 Copy files to KS-XENA

- Copy the new OPUS files to “KS-XENA” (the OP-FTIR computer).
- Put the “out” files in C:/Malt5/Data_Nov2017.
- Put “malt_input1.txt” into C:/Malt5/N2O_path336m.

C.4 Run retrieval

- Open Command Prompt and type:
cd C:\Malt5\N2O_path336m.
- Run Malt:
Malt5 /L malt_para_N2O_1 malt_input1.txt.

Bibliography

- Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al. (2003). The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, 4(6):1147–1167.
- Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., et al. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Rev. Geophys.*, 53(3):785–818.
- Andres, R., Boden, T., and Marland, G. (2016). Monthly fossil-fuel CO₂ emissions: Mass of emissions gridded by one degree latitude by one degree longitude. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A*, doi:10.3334/CDIAC/ffe.MonthlyMass.2016.
- Anselmo, D., Moran, M., Ménard, S., Talbot, D., Bouchet, V., Makar, P., Gong, W., Kallaur, A., Beaulieu, P., Landry, H., et al. (2010). A new Canadian air quality forecast model: GEM–MACH 15. Technical report, American Meteorological Society.
- Arneth, A., Kelliher, F., McSeveny, T., and Byers, J. (1998). Net ecosystem productivity, net primary productivity and ecosystem carbon sequestration in a *Pinus radiata* plantation subject to soil water deficit. *Tree Physiol.*, 18(12):785–793, doi:10.1093/treephys/18.12.785.
- Arora, V. K. and Boer, G. J. (2005). A parameterization of leaf phenology for the terrestrial ecosystem component of climate models. *Global Change Biol.*, 11(1):39–59, doi:10.1111/j.1365-2486.2004.00890.x.
- Bacastow, R. (1976). Modulation of atmospheric carbon dioxide by the southern oscillation. *Nature*, 261(5556):116–118, doi:10.1038/261116a0.

- Bacastow, R., Adams, J., Keeling, C., Moss, D., Whorf, T., and Wong, C. (1980). Atmospheric carbon dioxide, the Southern Oscillation, and the weak 1975 El Niño. *Science*, 210(4465):66–68, doi:10.1126/science.210.4465.66.
- Bacastow, R., Keeling, C., and Whorf, T. (1985). Seasonal amplitude increase in atmospheric CO₂ concentration at Mauna Loa, Hawaii, 1959–1982. *J. Geophys. Res.-Atmos.*, 90(D6):10529–10540, doi:10.1029/JD090iD06p10529.
- Badawy, B., Polavarapu, S., Jones, D. B. A., Deng, F., Neish, M., Melton, J. R., Nassar, R., and Arora, V. K. (2018). Coupling the Canadian Terrestrial Ecosystem Model (CTEM v. 2.0) to Environment and Climate Change Canada’s greenhouse gas forecast model (v.107-glb). *Geosci. Model Dev.*, 11(2):631–663, doi:10.5194/gmd-11-631-2018.
- Baker, D., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., et al. (2006a). TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO₂ fluxes, 1988–2003. *Global Biogeochem. Cy.*, 20(GB1002), doi:10.1029/2004GB002439.
- Baker, D. F., Doney, S. C., and Schimel, D. S. (2006b). Variational data assimilation for atmospheric CO₂. *Tellus B*, 58(5):359–365, doi:10.1111/j.1600-0889.2006.00218.x.
- Baker, I., Denning, A. S., Hanan, N., Prihodko, L., Uliasz, M., Vidale, P.-L., Davis, K., and Bakwin, P. (2003). Simulated and observed fluxes of sensible and latent heat and CO₂ at the WLEF–TV tower using SiB2. 5. *Glob. Change Biol.*, 9(9):1262–1277.
- Baker, I., Prihodko, L., Denning, A., Goulden, M., Miller, S., and Da Rocha, H. (2008). Seasonal drought stress in the Amazon: Reconciling models and observations. *J. Geophys. Res.-Biogeo.*, 113(G00B01), doi:10.1029/2007JG000644.
- Baker, I. T., Denning, A. S., and Stöckli, R. (2010). North American gross primary productivity: regional characterization and interannual variability. *Tellus B*, 62(5):533–549, doi:10.1111/j.1600-0889.2010.00492.x.
- Baldocchi, D., Chu, H., and Reichstein, M. (2018). Inter-annual variability of net and gross ecosystem carbon fluxes: A review. *Agr. Forest Meteorol.*, 249(Supplement C):520–533, doi:10.1016/j.agrformet.2017.05.015.
- Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthony, P., Bernhofer, C., Davis, K., Evans, R., 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. *B. Am. Meteorol. Soc.*, 82(11):2415–2434.

- Bares, R., Lin, J. C., Hoch, S. W., Baasandorj, M., Mendoza, D. L., Fasoli, B., Mitchell, L., Catharine, D., and Stephens, B. B. (2018). The wintertime covariation of CO₂ and criteria pollutants in an urban valley of the western United States. *J. Geophys. Res.-Atmos.*, 123(5):2684–2703, doi:10.1002/2017JD027917.
- Barnes, E. A., Parazoo, N., Orbe, C., and Denning, A. S. (2016). Isentropic transport and the seasonal cycle amplitude of CO₂. *J. Geophys. Res.-Atmos.*, 121(13):8106–8124, doi:10.1002/2016JD025109.
- Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., et al. (2013). Global CO₂ fluxes estimated from GOSAT retrievals of total column CO₂. *Atmos. Chem. Phys.*, 13(17):8695–8717, doi:10.5194/acp-13-8695-2013.
- Basu, S., Houweling, S., Peters, W., Sweeney, C., Machida, T., Maksyutov, S., Patra, P. K., Saito, R., Chevallier, F., Niwa, Y., et al. (2011). The seasonal cycle amplitude of total column CO₂: Factors behind the model-observation mismatch. *J. Geophys. Res.-Atmos.*, 116(D23306), doi:10.1029/2011JD016124.
- Basu, S., Krol, M., Butz, A., Clerbaux, C., Sawa, Y., Machida, T., Matsueda, H., Frankenberg, C., Hasekamp, O., and Aben, I. (2014). The seasonal variation of the CO₂ flux over Tropical Asia estimated from GOSAT, CONTRAIL, and IASI. *Geophys. Res. Lett.*, 41(5):1809–1815, doi:10.1002/2013GL059105.
- Becker, K., Lörzer, J., Kurtenbach, R., Wiesen, P., Jensen, T., and Wallington, T. (1999). Nitrous oxide (N₂O) emissions from vehicles. *Environ. Sci. Technol.*, 33(22):4134–4139, doi:10.1021/es9903330.
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., et al. (2018). Redefine statistical significance. *Nat. Hum. Behav.*, 2(1):6, doi:10.1038/s41562-017-0189-z.
- Berry, J. and Bjorkman, O. (1980). Photosynthetic response and adaptation to temperature in higher plants. *Ann. Rev. Plant Physio.*, 31(1):491–543.
- Blumenstock, T., Hase, F., Schneider, M., Garcia, O. E., and Sepulveda, E. (2014). TCCON data from Izana (ES), Release GGG2014R0. doi:10.14291/tccon.ggg2014.izana01.R0/1149295.

- Bolin, B. and Bischof, W. (1970). Variations of the carbon dioxide content of the atmosphere in the northern hemisphere. *Tellus*, 22(4):431–442, doi:10.3402/tellusa.v22i4.10236.
- Bolin, B. and Eriksson, E. (1959). Changes in the carbon dioxide content of the atmosphere and sea due to fossil fuel combustion. *The atmosphere and the sea in motion*, 1:30–142.
- Bolin, B. and Keeling, C. (1963). Large-scale atmospheric mixing as deduced from the seasonal and meridional variations of carbon dioxide. *J. Geophys. Res.*, 68(13):3899–3920, doi:0.1029/JZ068i013p03899.
- Bousquet, P., Peylin, P., Ciais, P., Le Quéré, C., Friedlingstein, P., and Tans, P. P. (2000). Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science*, 290(5495):1342–1346.
- Bousquet, P., Peylin, P., Ciais, P., Ramonet, M., and Monfray, P. (1999). Inverse modeling of annual atmospheric CO₂ sources and sinks: 2. sensitivity study. *J. Geophys. Res.-Atmos.*, 104(D21):26179–26193.
- Bowman, K., Liu, J., Bloom, A., Parazoo, N., Lee, M., Jiang, Z., Menemenlis, D., Gierach, M., Collatz, G., Gurney, K., et al. (2017). Global and Brazilian carbon response to El Niño Modoki 2011–2010. *Earth and Space Sci.*, 4(10):637–660, doi:10.1002/2016EA000204.
- Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P. (2015). An attempt at estimating Paris area CO₂ emissions from atmospheric concentration measurements. *Atmos. Chem. Phys.*, 15(4):1707–1724, doi:10.5194/acp-15-1707-2015.
- Bruhwiller, L., Michalak, A., Peters, W., Baker, D., and Tans, P. (2005). An improved Kalman Smoother for atmospheric inversions. *Atmos. Chem. Phys.*, 5(10):2691–2702, doi:10.5194/acp-5-2691-2005.
- Bruhwiller, L., Michalak, A., and Tans, P. (2011). Spatial and temporal resolution of carbon flux estimates for 1983–2002. *Biogeosciences*, 8(5):1309–1331, doi:10.5194/bg-8-1309-2011.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.*, 16(5):1190–1208, doi:10.1137/0916069.

- Cambaliza, M., Shepson, P., Bogner, J., Caulton, D., Stirm, B., Sweeney, C., Montzka, S., Gurney, K., Spokas, K., Salmon, O., et al. (2015). Quantification and source apportionment of the methane emission flux from the city of Indianapolis. *Elem Sci Anth*, 3, doi:10.12952/journal.elementa.000037.
- Channan, S., Collins, K., and Emanuel, W. (2014). Global mosaics of the standard MODIS land cover type data. *University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA*, 30.
- Chapin, F. S., Matson, P. A., and Mooney, H. A. (2002). *Principles of Terrestrial Ecosystem Ecology*. Springer.
- Charney, J. G., Arakawa, A., Baker, D. J., Bolin, B., Dickinson, R. E., Goody, R. M., Leith, C. E., Stommel, H. M., and Wunsch, C. I. (1979). *Carbon dioxide and climate: a scientific assessment*. National Academy of Sciences, Washington, DC.
- Chatterjee, A., Gierach, M., Sutton, A., Feely, R., Crisp, D., Eldering, A., Gunson, M., O'dell, C., Stephens, B., and Schimel, D. (2017). Influence of El Niño on atmospheric CO₂ over the tropical Pacific Ocean: Findings from NASA's OCO-2 mission. *Science*, 358(6360):eaam5776, doi:10.1126/science.aam5776.
- Chen, J., Viatte, C., Hedelius, J. K., Jones, T., Franklin, J. E., Parker, H., Gottlieb, E. W., Wennberg, P. O., Dubey, M. K., and Wofsy, S. C. (2016). Differential column measurements using compact solar-tracking spectrometers. *Atmospheric Chemistry and Physics*, 16(13):8479–8498.
- Chen, Z., Chen, J. M., Zhang, S., Zheng, X., Ju, W., Mo, G., and Lu, X. (2017). Optimization of terrestrial ecosystem model parameters using atmospheric CO₂ concentration data with the global carbon assimilation system (GCAS). *Journal of Geophysical Research: Biogeosciences*, 122(12):3218–3237.
- Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P. (2005). Inferring CO₂ sources and sinks from satellite observations: Method and application to TOVS data. *J. Geophys. Res.-Atmos.*, 110(D24), doi:10.1029/2005JD006390.
- Chevallier, F., Palmer, P. I., Feng, L., Boesch, H., O'Dell, C. W., and Bousquet, P. (2014). Toward robust and consistent regional CO₂ flux estimates from in situ and spaceborne measurements of atmospheric CO₂. *Geophys. Res. Lett.*, 41(3):1065–1070, doi:10.1002/2013GL058772. 2013GL058772.

- Ciais, P., Borges, A., Abril, G., Meybeck, M., Folberth, G., Hauglustaine, D., and Janssens, I. (2008). The impact of lateral carbon fluxes on the European carbon balance. *Biogeosciences*, 5(5):1259–1271, doi:10.5194/bg-5-1259-2008.
- Ciais, P., Bousquet, P., Freibauer, A., and Naegler, T. (2007). Horizontal displacement of carbon associated with agriculture and its impacts on atmospheric CO₂. *Global Biogeochem. Cy.*, 21(GB2014), doi:10.1029/2006GB002741.
- Ciais, P., Dolman, A., Bombelli, A., Duren, R., Peregon, A., Rayner, P., Miller, C., Gobron, N., Kinderman, G., Marland, G., et al. (2014). Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system. *Biogeosciences*, 11:3547–3602, doi:10.5194/bg-11-3547-2014.
- Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., et al. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058):529–533, doi:https://doi.org/10.1038/nature03972.
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R., Piao, S., and Thornton, P. (2013). *Carbon and Other Biogeochemical Cycles*, pages 465–570. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Clark, D., Mercado, L., Sitch, S., Jones, C., Gedney, N., Best, M., Pryor, M., Rooney, G., Essery, R., Blyth, E., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description-part 2: carbon fluxes and vegetation dynamics. *Geosci. Model Dev.*, 4(3):701–722, doi:10.1029/93GB02725.
- Colorado, A., McDonell, V., and Samuelsen, S. (2017). Direct emissions of nitrous oxide from combustion of gaseous fuels. *Int. J. Hydrogen Energ.*, 42(1):711–719, doi:10.1016/j.ijhydene.2016.09.202.
- Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y.-W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., et al. (2017). Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra. *Proc. Natl. Acad. Sci.*, 114(21):5361–5366, doi:10.1073/pnas.1618567114.
- Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D. (2008). Orbiting Carbon Observatory: Inverse method and prospective error analysis. *J. Geophys. Res.-Atmos.*, 113(D5), doi:10.1029/2006JD008336.

- Conway, T. J., Tans, P. P., Waterman, L. S., Thoning, K. W., Kitzis, D. R., Masarie, K. A., and Zhang, N. (1994). Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory global air sampling network. *J. Geophys. Res.-Atmos.*, 99(D11):22831–22855.
- Crisp, D. (2015). Measuring atmospheric carbon dioxide from space with the orbiting carbon observatory-2 (OCO-2). In *Earth observing systems xx*, volume 9607, page 960702. International Society for Optics and Photonics.
- Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L. (2012). The ACOS CO₂ retrieval algorithm-Part II: Global XCO₂ data characterization. *Atmos. Meas. Tech.*, 5(4):687–707, doi:10.5194/amt-5-687-2012.
- Croft, H., Chen, J., Froelich, N., Chen, B., and Staebler, R. (2015). Seasonal controls of canopy chlorophyll content on forest carbon uptake: Implications for GPP modeling. *Journal of Geophysical Research: Biogeosciences*, 120(8):1576–1586.
- Dai, A. (2011). Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. *J. Geophys. Res.-Atmos.*, 116(D12115), doi:10.1029/2010JD015541. D12115.
- Dai, A. (2017). Dai Global Palmer Drought Severity Index (PDSI). Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/D6QF8R93>. Accessed 15 Aug 2017.
- Dai, A., Trenberth, K. E., and Qian, T. (2004). A global dataset of Palmer Drought Severity Index for 1870–2002: relationship with soil moisture and effects of surface warming. *J. Hydrometeorol.*, 5(6):1117–1130.
- Damm, A., Guanter, L., Paul-Limoges, E., Van der Tol, C., Hueni, A., Buchmann, N., Eugster, W., Ammann, C., and Schaepman, M. (2015). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An

- assessment based on observational and modeling approaches. *Remote Sens. Environ.*, 166:91–105.
- Davis, K. J., Deng, A., Lauvaux, T., Miles, N. L., Richardson, S. J., Sarmiento, D. P., Gurney, K. R., Hardesty, R. M., Bonin, T. A., Brewer, W. A., et al. (2017). The indianapolis flux experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements. *Elem Sci Anth*, 5, doi:10.1525/elementa.188.
- De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Dufflot, V., and Cammas, J.-P. (2014). TCCON data from Reunion Island (RE), Release GGG2014R0. doi:10.14291/tccon.ggg2014.reunion01.R0/1149288.
- DeLucia, E., Drake, J. E., Thomas, R. B., and Gonzalez-Meler, M. (2007). Forest carbon use efficiency: is respiration a constant fraction of gross primary production? *Glob. Change Biol.*, 13(6):1157–1167.
- Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., Wu, K., Sarmiento, D. P., Hardesty, R. M., Bonin, T. A., et al. (2017). Toward reduced transport errors in a high resolution urban CO₂ inversion system. *Elem. Sci. Anth.*, 5, doi:10.1525/elementa.133.
- Deng, F., Jones, D., Henze, D., Bousserez, N., Bowman, K., Fisher, J., Nassar, R., O'Dell, C., Wunch, D., Wennberg, P., et al. (2014). Inferring regional sources and sinks of atmospheric CO₂ from GOSAT XCO₂ data. *Atmos. Chem. Phys.*, 14(7):3703–3727.
- Deng, F., Jones, D., O'Dell, C. W., Nassar, R., and Parazoo, N. C. (2016). Combining gosat XCO₂ observations over land and ocean to improve regional CO₂ flux estimates. *J. Geophys. Res.-Atmos.*, 121(4):1896–1913, doi:10.1002/2015JD024157.
- Deng, F., Jones, D. B. A., Walker, T. W., Keller, M., Bowman, K. W., Henze, D. K., Nassar, R., Kort, E. A., Wofsy, S. C., Walker, K. A., Bourassa, A. E., and Degenstein, D. A. (2015). Sensitivity analysis of the potential impact of discrepancies in stratosphere–troposphere exchange on inferred sources and sinks of CO₂. *Atmos. Chem. Phys.*, 15(20):11773–11788, doi:10.5194/acp-15-11773-2015.
- Denning, A. S., Collatz, G. J., Zhang, C., Randall, D. A., Berry, J. A., Sellers, P. J., Colell, G. D., and Dazlich, D. A. (1996). Simulations of terrestrial carbon metabolism and atmospheric CO₂ in a general circulation model. Part 1: Surface carbon fluxes. *Tellus B*, 48:521–542, doi:10.1034/j.1600-0889.1996.t01-2-00009.x.

- Denning, A. S., Fung, I. Y., and Randall, D. (1995). Latitudinal gradient of atmospheric CO₂ due to seasonal exchange with land biota. *Nature*, 376(6537):240.
- Denning, A. S., Holzer, M., Gurney, K. R., Heimann, M., Law, R. M., Rayner, P. J., Fung, I. Y., Fan, S.-M., Taguchi, S., Friedlingstein, P., et al. (1999a). Three-dimensional transport and concentration of SF₂ A model intercomparison study (TransCom 2). *Tellus B*, 51(2):266–297.
- Denning, A. S., Takahashi, T., and Friedlingstein, P. (1999b). Can a strong atmospheric CO₂ rectifier effect be reconciled with a “reasonable” carbon budget? *Tellus B*, 51(2):249–253.
- Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., Grupe, P., and Katrynski, K. (2014). TCCON data from Bialystok (PL), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Dubey, M., Henderson, B., Green, D., Butterfield, Z., Keppel-Aleks, G., Allen, N., Blavier, J.-F., Roehl, C., Wunch, D., and Lindenmaier, R. (2014). TCCON data from Manaus (BR), Release GGG2014R0. doi:10.14291/tccon.ggg2014.manaus01.R0/1149274.
- Enting, I. and Mansbridge, J. (1989). Seasonal sources and sinks of atmospheric CO₂ direct inversion of filtered data. *Tellus B*, 41(2):111–126.
- Enting, I., Trudinger, C., and Francey, R. (1995). A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO₂. *Tellus B*, 47(1-2):35–52.
- Fan, S., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., and Tans, P. (1998). A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science*, 282(5388):442–446.
- Fang, Y., Michalak, A. M., Schwalm, C. R., Huntzinger, D. N., Berry, J. A., Ciais, P., Piao, S., Poulter, B., Fisher, J. B., Cook, R. B., Hayes, D., Huang, M., Ito, A., Jain, A., Lei, H., Lu, C., Mao, J., Parazoo, N. C., Peng, S., Ricciuto, D. M., Shi, X., Tao, B., Tian, H., Wang, W., Wei, Y., and Yang, J. (2017). Global land carbon sink response to temperature and precipitation varies with enso phase. *Environmental Research Letters*, 12(6):064007, doi:10.1088/1748-9326/aa6e8e.
- Feist, D. G., Arnold, S. G., John, N., and Geibel, M. C. (2014). TCCON data from Ascension Island (SH), Release GGG2014R0. doi:10.14291/tccon.ggg2014.ascension01.R0/1149285.

- Feng, L., Palmer, P. I., Bösch, H., and Dance, S. (2009). Estimating surface CO₂ fluxes from space-borne CO₂ dry air mole fraction observations using an ensemble Kalman Filter. *Atmos. Chem. Phys.*, 9(8):2619–2633, doi:10.5194/acp-9-2619-2009.
- Feng, L., Palmer, P. I., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R. (2016). Estimates of European uptake of CO₂ inferred from GOSAT XCO₂ retrievals: sensitivity to measurement bias inside and outside Europe. *Atmos. Chem. Phys.*, 16(3):1289–1302, doi:10.5194/acp-16-1289-2016.
- Field, C. B., Randerson, J. T., and Malmström, C. M. (1995). Global net primary production: combining ecology and remote sensing. *Remote Sens. Environ.*, 51(1):74–88.
- Flesch, T. K., Baron, V. S., Wilson, J. D., Griffith, D. W., Basarab, J. A., and Carlson, P. J. (2016). Agricultural gas emissions during the spring thaw: Applying a new measurement technique. *Agr. Forest Meteorol.*, 221:111–121.
- Frank, D., Reichstein, M., Bahn, M., Thonicke, K., Frank, D., Mahecha, M. D., Smith, P., Van der Velde, M., Vicca, S., Babst, F., et al. (2015). Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts. *Global Change Biology*, 21(8):2861–2880.
- Frankenberg, C., Butz, A., and Toon, G. C. (2011a). Disentangling chlorophyll fluorescence from atmospheric scattering effects in O₂ A-band spectra of reflected sun-light. *Geophys. Res. Lett.*, 38(3), doi:10.1029/2010GL045896. L03801.
- Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., et al. (2011b). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophys. Res. Lett.*, 38(17706), doi:10.1029/2011GL048738.
- Frankenberg, C., Kulawik, S. S., Wofsy, S. C., Chevallier, F., Daube, B., Kort, E. A., O’Dell, C., Olsen, E. T., and Osterman, G. (2016). Using airborne HIAPER pole-to-pole observations (HIPPO) to evaluate model and remote sensing estimates of atmospheric carbon dioxide. *Atmos. Chem. Phys.*, 16(12):7867–7878, doi:10.5194/acp-16-7867-2016.
- Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refine-

- ments and characterization of new datasets. *Remote Sens. Environ.*, 114(1):168–182, doi:10.1016/j.rse.2009.08.016.
- Fu, D., Pongetti, T., Blavier, J.-F., Crawford, T., Manatt, K., Toon, G., Wong, K., and Sander, S. (2014). Near-infrared remote sensing of Los Angeles trace gas distributions from a mountaintop site. *Atmospheric Measurement Techniques*, 7(3):713–729.
- Fung, I., Prentice, K., Matthews, E., Lerner, J., and Russell, G. (1983). Three-dimensional tracer model study of atmospheric CO₂: Response to seasonal exchanges with the terrestrial biosphere. *J. Geophys. Res.-Oceans*, 88(C2):1281–1294.
- Fung, I., Tucker, C., and Prentice, K. (1987). Application of advanced very high resolution radiometer vegetation index to study atmosphere-biosphere exchange of CO₂. *J. Geophys. Res.-Atmos.*, 92(D3):2999–3015.
- Fung, I. Y. (1986). Analysis of the seasonal and geographical patterns of atmospheric CO₂ distributions with a three-dimensional tracer model. In *The Changing Carbon Cycle*, pages 459–473. Springer.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Climate*, 30(14):5419–5454.
- Gerbig, C., Lin, J., Wofsy, S., Daube, B., Andrews, A., Stephens, B., Bakwin, P., and Grainger, C. (2003). Toward constraining regional-scale fluxes of CO₂ with atmospheric observations over a continent: 2. analysis of COBRA data using a receptor-oriented framework. *J. Geophys. Res.-Atmos.*, 108(D24), doi:10.1029/2003JD003770.
- GFZ-SIF (2016). GFZ postdam GOME-2 Solar Induced Fluorescence product. <ftp://ftp.gfz-potsdam.de>. Accessed: 2016-03-08.
- Gilmanov, T., Soussana, J., Aires, L., Allard, V., Ammann, C., Balzarolo, M., Barcza, Z., Bernhofer, C., Campbell, C., Cernusca, A., et al. (2007). Partitioning European grassland net ecosystem CO₂ exchange into gross primary productivity and ecosystem respiration using light response function analysis. *Agr. Ecosyst. Environ.*, 121(1-2):93–120.
- Goody, R. and Yung, Y. (1995). *Atmospheric radiation: theoretical basis*. Oxford University Press.

- Graven, H., Keeling, R., Piper, S., Patra, P., Stephens, B., Wofsy, S., Welp, L., Sweeney, C., Tans, P., Kelley, J., et al. (2013). Enhanced seasonal exchange of CO₂ by northern ecosystems since 1960. *Science*, 341(6150):1085–1089, doi:10.1126/science.1239207.
- Griffith, D., Deutscher, N., Caldw, C., Kettlewell, G., Riggenbach, M., and Hammer, S. (2012). A Fourier transform infrared trace gas and isotope analyser for atmospheric applications. *Atmos. Meas. Tech.*, 5(10):2481–2498.
- Griffith, D. W. (1996). Synthetic calibration and quantitative analysis of gas-phase FT-IR spectra. *Appl. Spectrosc.*, 50(1):59–70.
- Griffith, D. W. (2003). Non-linear least squares: High precision quantitative analysis of gas phase FTIR spectra. 2nd Intl. Conference on Advanced Vibrational Spectroscopy, Nottingham.
- Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Aleks, G. K., Washenfelder, R. a., Toon, G. C., Blavier, J.-F., Murphy, C., Jones, N., Kettlewell, G., Connor, B. J., Macatangay, R., Roehl, C., Ryzcek, M., Glowacki, J., Culgan, T., and Bryant, G. (2014a). TCCON data from Darwin (AU), Release GGG2014R0. doi:10.14291/tccon.ggg2014.darwin01.R0/1149290.
- Griffith, D. W., Pöhler, D., Schmitt, S., Hammer, S., Vardag, S. N., and Platt, U. (2018). Long open-path measurements of greenhouse gases in air using near-infrared Fourier transform spectroscopy. *Atmos. Meas. Tech.*, 11(3):1549.
- Griffith, D. W., Velazco, V. A., Deutscher, N. M., Murphy, C., Jones, N., Wilson, S., Macatangay, R., Kettlewell, G., Buchholz, R. R., and Riggenbach, M. (2014b). TCCON data from Wollongong (AU), Release GGG2014R0. doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291.
- Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G. (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sens. Environ.*, 121:236–251, doi:10.1016/j.rse.2012.02.006.
- Guerlet, S., Basu, S., Butz, A., Krol, M., Hahne, P., Houweling, S., Hasekamp, O., and Aben, I. (2013). Reduced carbon uptake during the 2010 Northern Hemisphere summer from GOSAT. *Geophys. Res. Lett.*, 40(10):2378–2383.
- Gurney, K. R., Baker, D., Rayner, P., and Denning, S. (2008). Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated

- from atmospheric CO₂ inversions for the period 1980 to 2005. *Global Biogeochem. Cy.*, 22(3), doi:10.1029/2007GB003082. GB3025.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., et al. (2002). Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models. *Nature*, 415(6872):626–630.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., et al. (2004). Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochem. Cy.*, 18(1).
- Harper, A. B., Wiltshire, A. J., Cox, P. M., Friedlingstein, P., Jones, C. D., Mercado, L. M., Sitch, S., Williams, K., and Duran-Rojas, C. (2018). Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4.6 with new plant functional types. *Geosci. Model Dev.*, 11(7):2857–2873, doi:10.5194/gmd-11-2857-2018.
- Hase, F., Blumenstock, T., Dohe, S., Gross, J., and Kiel, M. (2014). TCCON data from Karlsruhe (DE), Release GGG2014R0. doi:10.14291/tccon.ggg2014.karlsruhe01.R0/1149270.
- Haszpra, L., Barcza, Z., Haszpra, T., Pátkai, Z., and Davis, K. (2015). How well do tall-tower measurements characterize the CO₂ mole fraction distribution in the planetary boundary layer? *Atmos. Meas. Tech.*, 8(4):1657–1671.
- He, Y., Piao, S., Li, X., Chen, A., and Qin, D. (2018). Global patterns of vegetation carbon use efficiency and their climate drivers deduced from modis satellite data and process-based models. *Agr. Forest Meteorol.*, 256:150–158.
- Hedelius, J. K., Feng, S., Roehl, C. M., Wunch, D., Hillyard, P. W., Podolske, J. R., Iraci, L. T., Patarasuk, R., Rao, P., O’Keeffe, D., et al. (2017). Emissions and topographic effects on column CO₂ (x_{CO_2}) variations, with a focus on the Southern California Megacity. *Journal of Geophysical Research: Atmospheres*, 122(13):7200–7215.
- Hedelius, J. K., Liu, J., Oda, T., Maksyutov, S., Roehl, C. M., Iraci, L. T., Podolske, J. R., Hillyard, P. W., Wunch, D., and Wennberg, P. O. (2018). Southern california megacity CO₂, CH₄, and CO flux estimates using remote sensing and a lagrangian model. *Atmos. Chem. Phys. Discussions*, 2018:1–32, doi:10.5194/acp-2018-517.

- Henze, D. K., Hakami, A., and Seinfeld, J. H. (2007). Development of the adjoint of GEOS-Chem. *Atmos. Chem. Phys.*, 7(9):2413–2433.
- Heskel, M. A., Atkin, O. K., Turnbull, M. H., and Griffin, K. L. (2013). Bringing the Kok effect to light: a review on the integration of daytime respiration and net ecosystem exchange. *Ecosphere*, 4(8):1–14.
- Holtslag, A. and Boville, B. (1993). Local versus nonlocal boundary-layer diffusion in a global climate model. *J. Climate*, 6(10):1825–1842.
- Hopkins, F. M., Kort, E. A., Bush, S. E., Ehleringer, J. R., Lai, C.-T., Blake, D. R., and Randerson, J. T. (2016). Spatial patterns and source attribution of urban methane in the los angeles basin. *Journal of Geophysical Research: Atmospheres*, 121(5):2490–2507.
- Houghton, R. (1987). Biotic changes consistent with the increased seasonal amplitude of atmospheric CO₂ concentrations. *J. Geophys. Res.-Atmos.*, 92(D4):4223–4230.
- Houweling, S., Aben, I., Breon, F.-M., Chevallier, F., Deutscher, N., Engelen, R., Gerbig, C., Griffith, D., Hungershofer, K., Macatangay, R., et al. (2010). The importance of transport model uncertainties for the estimation of CO₂ sources and sinks using satellite measurements. *Atmos. Chem. Phys.*, 10(20):9981–9992.
- Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E., Feng, L., Ganshin, A., et al. (2015). An intercomparison of inverse models for estimating sources and sinks of CO₂ using GOSAT measurements. *J. Geophys. Res.-Atmos.*, 120(10):5253–5266.
- Huntzinger, D., Michalak, A., Schwalm, C., Ciais, P., King, A., Fang, Y., Schaefer, K., Wei, Y., Cook, R., Fisher, J., et al. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, 7.
- Huntzinger, D., Post, W. M., Wei, Y., Michalak, A., West, T. O., Jacobson, A., Baker, I., Chen, J. M., Davis, K., Hayes, D., et al. (2012). North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison. *Ecol. Model.*, 232:144–157.
- IPCC (2013). *WGI Summary for Policymakers*, book section SPM, page 1–30. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

- IPCC (2014). *WGII Summary for Policymakers*, book section SPM, page 1–32. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Iraci, L. T., Podolske, J., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., Allen, N., Wunch, D., Osterman, G. B., and Albertson, R. (2016). TCCON data from Edwards (US), Release GGG2014R1. doi:10.14291/tccon.ggg2014.edwards01.R1/1255068.
- Ishizawa, M., Mabuchi, K., Shirai, T., Inoue, M., Morino, I., Uchino, O., Yoshida, Y., Belikov, D., and Maksyutov, S. (2016). Inter-annual variability of summertime CO₂ exchange in Northern Eurasia inferred from GOSAT XCO₂. *Environ. Res. Lett.*, 11(10):105001.
- Ito, A. (2010). Changing ecophysiological processes and carbon budget in east asian ecosystems under near-future changes in climate: implications for long-term monitoring from a process-based model. *J. Plant Res.*, 123(4):577–588.
- Jassal, R. S., Black, T. A., Cai, T., Morgenstern, K., Li, Z., Gaumont-Guay, D., and Nesic, Z. (2007). Components of ecosystem respiration and an estimate of net primary productivity of an intermediate-aged douglas-fir stand. *Agr. Forest Meteorol.*, 144(1-2):44–57.
- Jeong, S.-J., Schimel, D., Frankenberg, C., Drewry, D. T., Fisher, J. B., Verma, M., Berry, J. A., Lee, J.-E., and Joiner, J. (2017). Application of satellite solar-induced chlorophyll fluorescence to understanding large-scale variations in vegetation phenology and function over northern high latitude forests. *Remote Sens. Environ.*, 190:178–187, doi:10.1016/j.rse.2016.11.021.
- Jöckel, P., von Kuhlmann, R., Lawrence, M. G., Steil, B., Brenninkmeijer, C. A., Crutzen, P. J., Rasch, P. J., and Eaton, B. (2001). On a fundamental problem in implementing flux-form advection schemes for tracer transport in 3-dimensional general circulation and chemistry transport models. *Quarterly Journal of the Royal Meteorological Society*, 127(573):1035–1052.
- Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A., Middleton, E., Huemmrich, K., Yoshida, Y., and Frankenberg, C. (2013). Global monitoring of terrestrial chlorophyll fluorescence from moderate spectral resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2. *Atmos. Meas. Tech.*, 6(2):2803–2823, doi:10.5194/amt-6-2803-2013.

- Joiner, J., Yoshida, Y., Guanter, L., and Middleton, E. M. (2016). New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY. *Atmos. Meas. Tech.*, 9(8):3939–3967, doi:10.5194/amt-9-3939-2016.
- Joiner, J., Yoshida, Y., Vasilkov, A., Middleton, E., et al. (2011). First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences*, 8(3):637–651.
- Jones, C. D. and Cox, P. M. (2005). On the significance of atmospheric CO₂ growth rate anomalies in 2002–2003. *Geophys. Res. Lett.*, 32(14):n/a–n/a, doi:10.1029/2005GL023027. L14816.
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.-Biogeo.*, 116(G3).
- Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., et al. (2017). Compensatory water effects link yearly global land CO₂ sink changes to temperature. *Nature*, 541(7638):516–520.
- Junge, C. and Czeplak, G. (1968). Some aspects of the seasonal variation of carbon dioxide and ozone. *Tellus*, 20(3):422–434.
- Kaminski, T., Heimann, M., and Giering, R. (1999). A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. inversion of the transport of CO₂ in the 1980s. *J. Geophys. Res.-Atmos.*, 104(D15):18555–18581.
- Kaminski, T., Knorr, W., Rayner, P., and Heimann, M. (2002). Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle. *Global Biogeochemical Cycles*, 16(4):14–1.
- Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P. (2012). Consistent assimilation of MERIS FAPAR and atmospheric CO₂ into a terrestrial vegetation model and interactive mission benefit analysis. *Biogeosciences*, 9(8):3173–3184, doi:10.5194/bg-9-3173-2012.

- Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., et al. (2013). The BETHY/JSBACH carbon cycle data assimilation system: Experiences and challenges. *Journal of Geophysical Research: Biogeosciences*, 118(4):1414–1426.
- Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G. (2001). On aggregation errors in atmospheric transport inversions. *J. Geophys. Res.-Atmos.*, 106(D5):4703–4715.
- Kaminski, T., Scholze, M., and Houweling, S. (2010). Quantifying the benefit of A-SCOPE data for reducing uncertainties in terrestrial carbon fluxes in CCDAS. *Tellus B: Chemical and Physical Meteorology*, 62(5):784–796.
- Karion, A., Sweeney, C., Tans, P., and Newberger, T. (2010). AirCore: An innovative atmospheric sampling system. *J. Atmos Ocean Tech.*, 27(11):1839–1853.
- Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M. (2014). TCCON data from Saga (JP), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Keeling, C. D. (1960). The concentration and isotopic abundances of carbon dioxide in the atmosphere. *Tellus*, 12(2):200–203.
- Keeling, C. D. (1998). Rewards and penalties of monitoring the earth. *Annu. Rev. Energ. Env.*, 23(1):25–82.
- Keeling, C. D., Adams Jr, J. A., Ekdahl Jr, C. A., and Guenther, P. R. (1976a). Atmospheric carbon dioxide variations at the South Pole. *Tellus*, 28(6):552–564.
- Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl Jr, C. A., Guenther, P. R., Waterman, L. S., and Chin, J. F. (1976b). Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus*, 28(6):538–551.
- Keeling, C. D., Bacastow, R. B., Carter, A., Piper, S. C., Whorf, T. P., Heimann, M., Mook, W. G., and Roeloffzen, H. (1989a). A three-dimensional model of atmospheric CO₂ transport based on observed winds: 1. analysis of observational data. *Aspects of climate variability in the Pacific and the Western Americas*, pages 165–236.
- Keeling, C. D., Chin, J., and Whorf, T. (1996). Increased activity of northern vegetation inferred from atmospheric CO₂ measurements. *Nature*, 382(6587):146–149.

- Keeling, C. D., Piper, S. C., and Heimann, M. (1989b). A three-dimensional model of atmospheric CO₂ transport based on observed winds: 4. mean annual gradients and interannual variations. *Aspects of climate variability in the Pacific and the Western Americas*, pages 305–363.
- Keppel-Aleks, G., Wennberg, P., and Schneider, T. (2011). Sources of variations in total column carbon dioxide. *Atmos. Chem. Phys.*, 11(8):3581–3593.
- Keppel-Aleks, G., Wennberg, P., Washenfelder, R., Wunch, D., Schneider, T., Toon, G., Andres, R. J., Blavier, J.-F., Connor, B., Davis, K., et al. (2012). The imprint of surface fluxes and transport on variations in total column carbon dioxide. *Biogeosciences*, 9(3):875.
- Keppel-Aleks, G., Wolf, A. S., Mu, M., Doney, S. C., Morton, D. C., Kasibhatla, P. S., Miller, J. B., Dlugokencky, E. J., and Randerson, J. T. (2014a). Separating the influence of temperature, drought, and fire on interannual variability in atmospheric CO₂. *Global Biogeochem. Bi.*, 28(11):1295–1310.
- Keppel-Aleks, G., Wolf, A. S., Mu, M., Doney, S. C., Morton, D. C., Kasibhatla, P. S., Miller, J. B., Dlugokencky, E. J., and Randerson, J. T. (2014b). Separating the influence of temperature, drought, and fire on interannual variability in atmospheric CO₂. *Global Biogeochem. Cy.*, 28(11):1295–1310.
- Kivi, R. and Heikkinen, P. (2016). Fourier transform spectrometer measurements of column CO₂ at Sodankylä, Finland. *Geosci. Instrum. Method. Data Syst.*, 5(2):271–279.
- Kivi, R., Heikkinen, P., and Kyrö, E. (2014). TCCON data from Sodankylä (FI), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Koffi, E., Rayner, P., Scholze, M., and Beer, C. (2012). Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system. *Global Biogeochemical Cycles*, 26(1).
- Koffi, E., Rayner, P., Scholze, M., Chevallier, F., and Kaminski, T. (2013). Quantifying the constraint of biospheric process parameters by CO₂ concentration and flux measurement networks through a carbon cycle data assimilation system. *Atmospheric Chemistry and Physics*, 13(21):10555–10572.

- Köhler, P., Guanter, L., and Joiner, J. (2015). A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. *Atmos. Meas. Tech.*, 8(6):2589–2608, doi:10.5194/amt-8-2589-2015.
- Kohlmaier, G. H., Siré, E.-O., Janecek, A., Keeling, C. D., Piper, S. C., and Revelle, R. (1989). Modelling the seasonal contribution of a CO₂ fertilization effect of the terrestrial vegetation to the amplitude increase in atmospheric CO₂ at Mauna Loa Observatory. *Tellus B*, 41(5):487–510.
- Kondo, M., Ichii, K., Takagi, H., and Sasakawa, M. (2015). Comparison of the data-driven top-down and bottom-up global terrestrial CO₂ exchanges: GOSAT CO₂ inversion and empirical eddy flux upscaling. *J. Geophys. Res.-Biogeo*, 120(7):1226–1245.
- Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E. (2013). Surface observations for monitoring urban fossil fuel CO₂ emissions: Minimum site location requirements for the Los Angeles megacity. *J. Geophys. Res.-Atmos.*, 118(3):1577–1584.
- Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 1. model structure. *J. Geophys. Res.-Atmos.*, 105(D20):24809–24822.
- Krol, M., Houweling, S., Bregman, B., Broek, M., Segers, A., Velthoven, P. v., Peters, W., Dentener, F., and Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model tm5: algorithm and applications. *Atmos. Chem. Phys.*, 5(2):417–432, doi:10.5194/acp-5-417-2005.
- Kulawik, S., Wunch, D., O’Dell, C., Frankenberg, C., Reuter, M., Oda, T., Chevallier, F., Sherlock, V., Buchwitz, M., Osterman, G., Miller, C. E., Wennberg, P. O., Griffith, D., Morino, I., Dubey, M. K., Deutscher, N. M., Notholt, J., Hase, F., Warneke, T., Sussmann, R., Robinson, J., Strong, K., Schneider, M., De Mazière, M., Shiomi, K., Feist, D. G., Iraci, L. T., and Wolf, J. (2016). Consistent evaluation of ACOS-GOSAT, BESD-SCIAMACHY, carbontracker, and MACC through comparisons to TCCON. *Atmos. Meas. Tech.*, 9(2):683–709, doi:10.5194/amt-9-683-2016.
- Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T. (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(35):6716–6733.
- Lakens, D., Adolphi, F. G., Albers, C. J., Anvari, F., Apps, M. A., Argamon, S. E.,

- Baguley, T., Becker, R. B., Benning, S. D., Bradford, D. E., et al. (2018). Justify your alpha. *Nat. Hum. Behav.*, pages 168–171.
- Landschuetzer, P., Gruber, N., and Bakker, D. C. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochem. Cy.*, 30(10):1396–1417.
- Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O’Keefe, D., et al. (2016). High-resolution atmospheric inversion of urban CO₂ emissions during the dormant season of the indianapolis flux experiment (influx). *J. Geophys. Res.-Atmos.*, 121(10):5213–5236.
- Lauvaux, T., Miles, N. L., Richardson, S. J., Deng, A., Stauffer, D. R., Davis, K. J., Jacobson, G., Rella, C., Calonder, G.-P., and DeCola, P. L. (2013). Urban emissions of CO₂ from Davos, Switzerland: The first real-time monitoring system using an atmospheric inversion technique. *J. Appl. Meteorol.*, 52(12):2654–2668.
- Law, R., Rayner, P., Denning, A., Erickson, D., Fung, I., Heimann, M., Piper, S., Ramonet, M., Taguchi, S., Taylor, J., et al. (1996). Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO₂ inversions. *Global Biogeochem. Cy.*, 10(4):783–796.
- Law, R. and Simmonds, I. (1996). The sensitivity of deduced CO₂ sources and sinks to variations in transport and imposed surface concentrations. *Tellus B*, 48(5):613–625.
- Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Nojiri, Y., Padin, X. A., Peregón, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D. (2018). Global carbon budget 2017. *Earth Syst. Sci. Data*, 10(1):405–448, doi:10.5194/essd-10-405-2018.

- Lee, J.-E., Berry, J. A., van der Tol, C., Yang, X., Guanter, L., Damm, A., Baker, I., and Frankenberg, C. (2015). Simulations of chlorophyll fluorescence incorporated into the Community Land Model version 4. *Global change biology*, 21(9):3469–3477.
- Lin, S.-J. and Rood, R. B. (1996). Multidimensional flux-form semi-lagrangian transport schemes. *Monthly Weather Review*, 124(9):2046–2070.
- Liu, J., Bowman, K. W., and Henze, D. K. (2015). Source-receptor relationships of column-average CO₂ and implications for the impact of observations on flux inversions. *J. Geophys. Res.-Atmos.*, 120(10):5214–5236.
- Liu, J., Bowman, K. W., Lee, M., Henze, D. K., Bousserez, N., Brix, H., Collatz, G. J., Menemenlis, D., Ott, L., Pawson, S., et al. (2014). Carbon monitoring system flux estimation and attribution: impact of ACOS-GOSAT XCO₂ sampling on the inference of terrestrial biospheric sources and sinks. *Tellus B*, 66(1):22486, doi:10.3402/tellusb.v66.22486.
- Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., Bloom, A. A., Wunch, D., Frankenberg, C., Sun, Y., O’Dell, C. W., Gurney, K. R., Menemenlis, D., Gierach, M., Crisp, D., and Eldering, A. (2017). Contrasting carbon cycle responses of the tropical continents to the 2015–2016 el niño. *Science*, 358(6360), doi:10.1126/science.aam5690.
- Luus, K., Commane, R., Parazoo, N., Benmergui, J., Euskirchen, E., Frankenberg, C., Joiner, J., Lindaas, J., Miller, C., Oechel, W., et al. (2017). Tundra photosynthesis captured by satellite-observed solar-induced chlorophyll fluorescence. *Geophys. Res. Lett.*, 44(3):1564–1573, doi:10.1002/2016GL070842.
- Machta, L. (1972). Mauna loa and global trends in air quality. *B. Am. Meteorol. Soc.*, 53(5):402–420.
- Makar, P., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R., et al. (2015). Feedbacks between air pollution and weather, part 2: Effects on chemistry. *Atmos. Environ.*, 115:499–526, doi:10.1016/j.atmosenv.2014.10.021.
- Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R., Ito, A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T. (2013). Regional CO₂ flux estimates for 2009–2010 based on GOSAT and ground-based

- CO₂ observations. *Atmos. Chem. Phys.*, 13(18):9351–9373, doi:10.5194/acp-13-9351-2013.
- Mandrake, L., Frankenberg, C., O'Dell, C. W., Osterman, G., Wennberg, P., and Wunch, D. (2013). Semi-autonomous sounding selection for OCO-2. *Atmos. Meas. Tech.*, 6(10):2851–2864, doi:10.5194/amt-6-2851-2013.
- Marcolla, B., Rödenbeck, C., and Cescatti, A. (2017). Patterns and controls of inter-annual variability in the terrestrial carbon budget. *Biogeosciences*, 14(16):3815–3829, doi:10.5194/bg-14-3815-2017.
- Masarie, K., Peters, W., Jacobson, A., and Tans, P. (2014). Obspack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements. *Earth System Science Data*, 6(2):375–384.
- Mays, K. L., Shepson, P. B., Stirm, B. H., Karion, A., Sweeney, C., and Gurney, K. R. (2009). Aircraft-based measurements of the carbon footprint of Indianapolis. *Environ. Sci. Technol.*, 43(20):7816–7823, doi:10.1021/es901326b. PMID: 19921899.
- McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyla, L. R., Floerchinger, C., Herdon, S. C., Nehrkorn, T., Zahniser, M. S., Jackson, R. B., et al. (2015). Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proceedings of the National Academy of Sciences*, 112(7):1941–1946.
- McKinley, G. A., Fay, A. R., Lovenduski, N. S., and Pilcher, D. J. (2017). Natural variability and anthropogenic trends in the ocean carbon sink. *Annual review of marine science*, 9:125–150.
- Melton, J. and Arora, V. (2016). Competition between plant functional types in the canadian terrestrial ecosystem model (CTEM) v. 2.0. *Geosci. Model Dev.*, 9(1):323–361, doi:10.5194/gmd-9-323-2016.
- Messerschmidt, J., Parazoo, N., Wunch, D., Deutscher, N. M., Roehl, C., Warneke, T., and Wennberg, P. O. (2013). Evaluation of seasonal atmosphere–biosphere exchange estimations with TCCON measurements. *Atmos. Chem. Phys.*, 13(10):5103–5115, doi:10.5194/acp-13-5103-2013.
- Michalak, A. M. (2008). Technical note: Adapting a fixed-lag Kalman smoother to a geostatistical atmospheric inversion framework. *Atmos. Chem. Phys.*, 8(22):6789–6799, doi:10.5194/acp-8-6789-2008.

- Miles, N. L., Richardson, S. J., Lauvaux, T., Davis, K. J., Balashov, N. V., Deng, A., Turnbull, J. C., Sweeney, C., Gurney, K. R., Patarasuk, R., et al. (2017). Quantification of urban atmospheric boundary layer greenhouse gas dry mole fraction enhancements in the dormant season: Results from the Indianapolis flux experiment (INFLUX). *Elem Sci Anth*, 5, doi:10.1525/elementa.127.
- Miller, C., Crisp, D., DeCola, P., Olsen, S., Randerson, J. T., Michalak, A. M., Alkhaled, A., Rayner, P., Jacob, D. J., Suntharalingam, P., et al. (2007). Precision requirements for space-based data. *J. Geophys. Res.-Atmos.*, 112(D10).
- Miller, S. M., Michalak, A. M., Yadav, V., and Tadić, J. M. (2018). Characterizing biospheric carbon balance using CO₂ observations from the OCO-2 satellite. *Atmos. Chem. Phys.*, 18(9):6785–6799, doi:10.5194/acp-18-6785-2018.
- Mitchell, L. E., Crosman, E. T., Jacques, A. A., Fasoli, B., Leclair-Marzolf, L., Horel, J., Bowling, D. R., Ehleringer, J. R., and Lin, J. C. (2018a). Monitoring of greenhouse gases and pollutants across an urban area using a light-rail public transit platform. *Atmos. Environ.*, 187:9–23, doi:10.1016/j.atmosenv.2018.05.044.
- Mitchell, L. E., Lin, J. C., Bowling, D. R., Pataki, D. E., Strong, C., Schauer, A. J., Bares, R., Bush, S. E., Stephens, B. B., Mendoza, D., et al. (2018b). Long-term urban carbon dioxide observations reveal spatial and temporal dynamics related to urban characteristics and growth. *Proc. Natl. Acad. Sci.*, 115(12):2912–2917.
- Molina, L., Broquet, G., Imbach, P., Chevallier, F., Poulter, B., Bonal, D., Burban, B., Ramonet, M., Gatti, L. V., Wofsy, S. C., Munger, J. W., Dlugokencky, E., and Ciais, P. (2015). On the ability of a global atmospheric inversion to constrain variations of CO₂ fluxes over Amazonia. *Atmos. Chem. Phys.*, 15(14):8423–8438, doi:10.5194/acp-15-8423-2015.
- Monsi, M. and Saeki, T. (1953). Über den lichtfaktor in den pflanzengesellschaften und seine bedeutung für die stoffproduktion. *Jpn. J. Bot.*, 14:22–52.
- Monson, R. K., Lipson, D. L., Burns, S. P., Turnipseed, A. A., Delany, A. C., Williams, M. W., and Schmidt, S. K. (2006). Winter forest soil respiration controlled by climate and microbial community composition. *Nature*, 439(7077):711–714.
- Morino, I., Matsuzaki, T., and Shishime, A. (2014a). TCCON data from Tsukuba (JP), 125HR, Release GGG2014R1. TCCON data archive, hosted by CaltechDATA.

- Morino, I., Yokozeki, N., Matzuzaki, T., and Horikawa, M. (2014b). TCCON data from Rikubetsu (JP), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA.
- Myneni, R. B., Keeling, C., Tucker, C. J., Asrar, G., and Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, 386(6626):698–702.
- NASA-SIF (2016). NASA GOME-2 Solar Induced Fluorescence product. <http://avdc.gsfc.nasa.gov/>. Accessed: 2016-07-25.
- Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D., and Crisp, D. (2017). Quantifying CO₂ emissions from individual power plants from space. *Geophys. Res. Lett.*, 44(19).
- Nassar, R., Jones, D., Kulawik, S., Worden, J., Bowman, K., Andres, R. J., Suntharalingam, P., Chen, J., Brenninkmeijer, C., Schuck, T., et al. (2011). Inverse modeling of CO₂ sources and sinks using satellite observations of CO₂ from TES and surface flask measurements. *Atmos. Chem. Phys.*, 11(12):6029–6047.
- Nassar, R., Jones, D. B., Suntharalingam, P., Chen, J. M., Andres, R. J., Wecht, K. J., Yantosca, R. M., Kulawik, S. S., Bowman, K. W., Worden, J. R., et al. (2010). Modeling global atmospheric CO₂ with improved emission inventories and CO₂ production from the oxidation of other carbon species. *Geosci. Model Dev.*, 3(2):689.
- Newman, S., Jeong, S., Fischer, M., Xu, X., Haman, C., Lefer, B., Alvarez, S., Rappenglueck, B., Kort, E., Andrews, A., et al. (2013). Diurnal tracking of anthropogenic CO₂ emissions in the Los Angeles basin megacity during spring 2010. *Atmospheric Chemistry and Physics*, 13(8):4359–4372.
- Newman, S., Xu, X., Gurney, K. R., Hsu, Y. K., Li, K. F., Jiang, X., Keeling, R., Feng, S., O’Keefe, D., Patarasuk, R., Wong, K. W., Rao, P., Fischer, M. L., and Yung, Y. L. (2016). Toward consistency between trends in bottom-up CO₂ emissions and top-down atmospheric measurements in the los angeles megacity. *Atmos. Chem. Phys.*, 16(6):3843–3863, doi:10.5194/acp-16-3843-2016.
- Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Buschmann, M., Weinzierl, C., Macatangay, R., and Grupe, P. (2014). TCCON data from Bremen (DE), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.

- O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D. (2012). The ACOS CO₂ retrieval algorithm – part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*, 5(1):99–121, doi:10.5194/amt-5-99-2012.
- Olsen, S. C. and Randerson, J. T. (2004). Differences between surface and column atmospheric CO₂ and implications for carbon cycle research. *J. Geophys. Res.-Atmos.*, 109(D2), doi:10.1029/2003JD003968. D02301.
- Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., et al. (2007). The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan. Ser. II*, 85(3):369–432.
- Papageorgiou, G. C. and Govindjee (2007). *Chlorophyll a fluorescence: a signature of photosynthesis*, volume 19. Springer Science & Business Media.
- Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D., Cescatti, A., Pérez-Priego, Ó., Wohlfahrt, G., and Montagnani, L. (2014). Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. *Glob. Change Biol.*, 20(10):3103–3121.
- Pataki, D., Bowling, D., and Ehleringer, J. (2003). Seasonal cycle of carbon dioxide and its isotopic composition in an urban atmosphere: Anthropogenic and biogenic effects. *J. Geophys. Res.-Atmos.*, 108(D23), doi:10.1029/2003JD003865.
- Pataki, D., Bowling, D., Ehleringer, J., and Zobitz, J. (2006). High resolution atmospheric monitoring of urban carbon dioxide sources. *Geophys. Res. Lett.*, 33(3), doi:10.1029/2005GL024822.
- Paton-Walsh, C., Smith, T., Young, E., Griffith, D. W., and Guérette, É.-A. (2014). New emission factors for Australian vegetation fires measured using open-path Fourier transform infrared spectroscopy—Part 1: Methods and Australian temperate forest fires. *Atmos. Chem. Phys.*, 14(20):11313–11333.
- Pearman, G. and Hyson, P. (1980). Activities of the global biosphere as reflected in atmospheric CO₂ records. *J. Geophys. Res.-Oceans*, 85(C8):4457–4467.

- Pearman, G. I., Hyson, P., and Fraser, P. (1983). The global distribution of atmospheric carbon dioxide: 1. aspects of observations and modeling. *J. Geophys. Res.-Oceans*, 88(C6):3581–3590.
- Peischl, J., Ryerson, T., Aikin, K., Gouw, J., Gilman, J., Holloway, J., Lerner, B., Nadkarni, R., Neuman, J., Nowak, J., et al. (2015). Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. *J. Geophys. Res.-Atmos.*, 120(5):2119–2139.
- Peng, S., Ciais, P., Chevallier, F., Peylin, P., Cadule, P., Sitch, S., Piao, S., Ahlström, A., Huntingford, C., Levy, P., et al. (2015). Benchmarking the seasonal cycle of CO₂ fluxes simulated by terrestrial ecosystem models. *Global Biogeochem. Cy.*, 29(1):46–64.
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M., Pétron, G., Hirsch, A. I., et al. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proc. Natl. Acad. Sci.*, 104(48):18925–18930, doi:10.1073/pnas.0708986104.
- Peters, W., Miller, J., Whitaker, J., Denning, A., Hirsch, A., Krol, M., Zupanski, D., Bruhwiler, L., and Tans, P. (2005). An ensemble data assimilation system to estimate CO₂ surface fluxes from atmospheric trace gas observations. *J. Geophys. Res.-Atmos.*, 110(D24), doi:10.1029/2005JD006157.
- Petty, G. (2006). *A first course in atmospheric radiation*. Sundog Pub.
- Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P. (2002). Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO₂ data. *J. Geophys. Res.-Atmos.*, 107(D19), doi:10.1029/2001JD000857.
- Peylin, P., Law, R., Gurney, K., Chevallier, F., Jacobson, A., Maki, T., Niwa, Y., Patra, P., Peters, W., Rayner, P., et al. (2013). Global atmospheric carbon budget: results from an ensemble of atmospheric CO₂ inversions. *Biogeosciences*, 10:6699–6720.
- Pierrehumbert, R. (2010). *Principles of planetary climate*. Cambridge Univ Pr.
- Polavarapu, S. M., Deng, F., Byrne, B., Jones, D. B. A., and Neish, M. (2018). A comparison of atmospheric CO₂ flux signals obtained from GEOS-Chem flux inversions constrained by in situ or GOSAT observations. *Atmos. Chem. Phys. Discussions*, 2018:1–55, doi:10.5194/acp-2017-1235.

- Polavarapu, S. M., Neish, M., Tanguay, M., Girard, C., Grandpré, J. d., Semeniuk, K., Gravel, S., Ren, S., Roche, S., Chan, D., et al. (2016). Greenhouse gas simulations with a coupled meteorological and transport model: the predictability of CO₂. *Atmos. Chem. Phys.*, 16(18):12005–12038, doi:10.5194/acp-16-12005-2016.
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A. (1993). Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochem. Cy.*, 7(4):811–841, doi:10.1029/93GB02725.
- Poulter, B., Frank, D., Hodson, E., and Zimmermann, N. (2011). Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO₂ airborne fraction. *Biogeosciences*, 8(8):2027.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in C++*. Oxford University Press.
- Project, C. G. A. D. I. (2013). NOAA. 2013, updated annually. *Cooperative Global Atmospheric Data Integration Project. Multi-laboratory compilation of atmospheric carbon dioxide data for the period 2000-2012 (obspack_CO2_1_PROTOTYPE_v1.0.4b_2014-02-13)*. Compiled by NOAA Global Monitoring Division: Boulder, Colorado, U.S.A. Data product accessed at <http://dx.doi.org/10.3334/OBSPACK/1001>.
- Pugliese, S. C., Murphy, J. G., Vogel, F. R., Moran, M. D., Zhang, J., Zheng, Q., Stroud, C. A., Ren, S., Worthy, D., and Broquet, G. (2018). High-resolution quantification of atmospheric CO₂ mixing ratios in the Greater Toronto Area, Canada. *Atmos. Chem. and Phys.*, 18(5):3387–3401, doi:10.5194/acp-18-3387-2018.
- Randerson, J., Chapin, F., Harden, J., Neff, J., and Harmon, M. (2002). Net ecosystem production: a comprehensive measure of net carbon accumulation by ecosystems. *Ecological applications*, 12(4):937–947.
- Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y., and Field, C. B. (1997). The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochem. Cy.*, 11(4):535–560.
- Randerson, J. T., Thompson, M. V., Malmstrom, C. M., Field, C. B., and Fung, I. Y. (1996). Substrate limitations for heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric CO₂. *Global Biogeochem. Cy.*, 10(4):585–602, doi:10.1029/96GB01981.

- Rayner, P., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, 19(2).
- Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P., Koster, R. D., and De Lannoy, G. J. (2017). Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, 30(8):2937–2960.
- Reichle, R. H., Koster, R. D., De Lannoy, G. J., Forman, B. A., Liu, Q., Mahanama, S. P., and Touré, A. (2011). Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, 24(24):6322–6338.
- Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J., Boesch, H., Parker, R., et al. (2014). Satellite-inferred european carbon sink larger than expected. *Atmos. Chem. Phys.*, 14(24):13739–13753.
- Revelle, R. and Suess, H. E. (1957). Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO₂ during the past decades. *Tellus*, 9(1):18–27.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J. (2011). MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, 24(14):3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Robichaud, A. and Ménard, R. (2014). Multi-year objective analyses of warm season ground-level ozone and PM 2.5 over North America using real-time observations and canadian operational air quality models. *Atmos. Chem. Phys.*, 14(4):1769–1800, doi:10.5194/acp-14-1769-2014.
- Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M. (2003). CO₂ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmos. Chem. Phys.*, 3(6):1919–1964.
- Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M. (2018). How does the terrestrial carbon exchange respond to inter-annual climatic variations? a quantification based on atmospheric co₂ data. *Biogeosciences*, 15(8):2481–2498, doi:10.5194/bg-15-2481-2018.

- Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice*, volume 2. World Scientific.
- Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coudert, L. H., Devi, V. M., Drouin, B. J., Fayt, A., Flaud, J. M., Gamache, R. R., Harrison, J. J., Hartmann, J. M., Hill, C., Hodges, J. T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R. J., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T., Mikhailenko, S., Mueller, H. S. P., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E. R., Richard, C., Smith, M. A. H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G. C., Tyuterev, V. G., and Wagner, G. (2013). The HITRAN2012 molecular spectroscopic database. *J. Quant. Spectrosc. Ra.*, 130(SI):4–50, doi:10.1016/j.jqsrt.2013.07.002.
- Ryan, M. G. and Law, B. E. (2005). Interpreting, measuring, and modeling soil respiration. *Biogeochemistry*, 73(1):3–27.
- Saeki, T., Maksyutov, S., Saito, M., Valsala, V., Oda, T., RJ, A., Belikov, D., Tans, P., Dlugokencky, E., Yoshida, Y., et al. (2013). Inverse modeling of CO₂ fluxes using gosat data and multi-year ground-based observations. *Sola*, 9:45–50.
- Saito, M., Ito, A., and Maksyutov, S. (2011). Evaluation of biases in JRA-25/JCDAS precipitation and their impact on the global terrestrial carbon balance. *J. Climate*, 24(15):4109–4125.
- Sargent, M., Barrera, Y., Nehrkorn, T., Hutyrá, L. R., Gately, C. K., Jones, T., McKain, K., Sweeney, C., Hegarty, J., Hardiman, B., et al. (2018). Anthropogenic and biogenic CO₂ fluxes in the Boston urban region. *Proceedings of the National Academy of Sciences*, 115(29):7491–7496, doi:10.1073/pnas.1803715115.
- Sasakawa, M., Machida, T., Tsuda, N., Arshinov, M., Davydov, D., Fofonov, A., and Krasnov, O. (2013). Aircraft and tower measurements of CO₂ concentration in the planetary boundary layer and the lower free troposphere over southern taiga in West Siberia: Long-term records from 2002 to 2011. *J. Geophys. Res.-Atmos.*, 118(16):9489–9498, doi:10.1002/jgrd.50755.
- Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., Davydov, D., Fofonov, A., Krasnov, O., Saeki, T., et al. (2010). Continuous measurements of methane from a tower network over Siberia. *Tellus B*, 62(5):403–416, doi:10.1111/j.1600-0889.2010.00494.x.

- Schuh, A., Denning, A., Corbin, K., Baker, I., Uliasz, M., Parazoo, N., Andrews, A., and Worthy, D. (2010). A regional high-resolution carbon flux inversion of North America for 2004. *Biogeosciences*, 7(5):1625–1644.
- Schuh, A. E., Denning, S., Baker, I., Haynes, K., Geyer, N., and McKeown, R. (2016). Global Atmospheric Inversions of XCO₂ and Solar-Induced Fluorescence (SIF) from OCO-2. Abstract A32D-02 presented at 2016 Fall Meeting, AGU, San Francisco, Calif., 12-16 Dec.
- Schuh, A. E., Lauvaux, T., West, T. O., Denning, A. S., Davis, K. J., Miles, N., Richardson, S., Uliasz, M., Lokupitiya, E., Cooley, D., et al. (2013). Evaluating atmospheric CO₂ inversions at multiple scales over a highly inventoried agricultural landscape. *Glob. Change Biol.*, 19(5):1424–1439.
- Schwalm, C. R., Anderegg, W. R., Michalak, A. M., Fisher, J. B., Biondi, F., Koch, G., Litvak, M., Ogle, K., Shaw, J. D., Wolf, A., et al. (2017). Global patterns of drought recovery. *Nature*, 548(7666):202.
- Schwandner, F. M., Gunson, M. R., Miller, C. E., Carn, S. A., Eldering, A., Krings, T., Verhulst, K. R., Schimel, D. S., Nguyen, H. M., Crisp, D., et al. (2017). Spaceborne detection of localized carbon dioxide sources. *Science*, 358(6360):eaam5782.
- Sellers, P., Randall, D., Collatz, G., Berry, J., Field, C., Dazlich, D., Zhang, C., Collelo, G., and Bounoua, L. (1996a). A revised land surface parameterization (SiB2) for atmospheric GCMs. part I: Model formulation. *J. Climate*, 9(4):676–705.
- Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., and Randall, D. A. (1996b). A revised land surface parameterization (SiB2) for atmospheric GCMs. part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate*, 9(4):706–737.
- Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., et al. (2014). *Human settlements, infrastructure and spatial planning*. Cambridge University Press.
- Sherlock, V., Connor, B. J., Robinson, J., Shiona, H., Smale, D., and Pollard, D. (2014). TCCON data from Lauder (NZ), 125HR, Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Shiga, Y. P., Tadić, J. M., Qiu, X., Yadav, V., Andrews, A. E., Berry, J. A., and Michalak, A. M. (2018). Atmospheric CO₂ observations reveal strong correlation between regional

- net biospheric carbon uptake and solar-induced chlorophyll fluorescence. *Geophys. Res. Lett.*, 45(2):1122–1132, doi:10.1002/2017GL076630.
- Shusterman, A. A., Teige, V. E., Turner, A. J., Newman, C., Kim, J., and Cohen, R. C. (2016). The BERkeley Atmospheric CO₂ Observation Network: initial evaluation. *Atmos. Chem. Phys.*, 16(21):13449–13463, doi:10.5194/acp-16-13449-2016.
- Siegenthaler, U. and Sarmiento, J. (1993). Atmospheric carbon dioxide and the ocean. *Nature*, 365(6442):119.
- Sippel, S., Reichstein, M., Ma, X., Mahecha, M. D., Lange, H., Flach, M., and Frank, D. (2018). Drought, heat, and the carbon cycle: a review. *Current Climate Change Reports*, pages 1–21.
- Stanevich, I. (2018a). Characterizing model errors in chemical transport modelling of methane: Impact of coarse model resolution. *manuscript in preparation*.
- Stanevich, I. (2018b). Diagnosing model biases using gosat xch 4 retrievals with weak constraint four dimensional variational data assimilation. *manuscript in preparation*.
- Stauffer, J., Broquet, G., Bréon, F.-M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P. (2016). The first 1-year-long estimate of the Paris region fossil fuel CO₂ emissions based on atmospheric inversion. *Atmos. Chem. Phys.*, 16(22):14703–14726, doi:10.5194/acp-16-14703-2016.
- Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., et al. (2007). Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO₂. *Science*, 316(5832):1732–1735.
- Stöckli, R., Rutishauser, T., Dragoni, D., O’keefe, J., Thornton, P., Jolly, M., Lu, L., and Denning, A. (2008). Remote sensing data assimilation for a prognostic phenology model. *J. Geophys. Res.-Biogeo*, 113(G4), doi:10.1029/2008JG000781.
- Strong, K., Mendonca, J., Weaver, D., Fogal, P., Drummond, J., Batchelor, R., and Lindenmaier, R. (2017). TCCON data from Eureka (CA), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA.

- Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D., Jung, M., Guanter, L., Drewry, D., Verma, M., Porcar-Castell, A., Griffis, T. J., et al. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, 358(6360):eaam5747.
- Suntharalingam, P., Jacob, D. J., Palmer, P. I., Logan, J. A., Yantosca, R. M., Xiao, Y., Evans, M. J., Streets, D. G., Vay, S. L., and Sachse, G. W. (2004). Improved quantification of Chinese carbon fluxes using CO₂/CO correlations in Asian outflow. *J. Geophys. Res.-Atmos.*, 109(D18).
- Super, I., van der Gon, H. D., Visschedijk, A., Moerman, M., Chen, H., Van der Molen, M., and Peters, W. (2017). Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam. *Atmos. Pollut. Res.*, 8(1):174–187.
- Sussmann, R. and Rettinger, M. (2014). TCCON data from Garmisch (DE), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Taguchi, S. (2000). Synthesis inversion of atmospheric CO₂ using the NIRE chemical transport model. *Geoph. Monog. Series*, 114:239–253, doi:10.1029/GM114p0239.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., et al. (2009). Climatological mean and decadal change in surface ocean pCO₂, and net sea–air CO₂ flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(8):554–577.
- Tans, P. P., Conway, T. J., and Nakazawa, T. (1989). Latitudinal distribution of the sources and sinks of atmospheric carbon dioxide derived from surface observations and an atmospheric transport model. *J. Geophys. Res.-Atmos.*, 94(D4):5151–5172.
- Tans, P. P., Fung, I. Y., and Takahashi, T. (1990). Observational constraints on the global atmospheric CO₂ budget. *Science*, 247(4949):1431–1438.
- Taylor, J. (1989). A stochastic lagrangian atmospheric transport model to determine global CO₂ sources and sinks—a preliminary discussion. *Tellus B*, 41(3):272–285.
- Tcherkez, G., Gauthier, P., Buckley, T. N., Busch, F. A., Barbour, M. M., Bruhn, D., Heskell, M. A., Gong, X. Y., Crous, K. Y., Griffin, K., et al. (2017). Leaf day respiration: low CO₂ flux but high significance for metabolism and carbon balance. *New Phytol.*, 216(4):986–1001.

- Te, Y., Jeseck, P., and Janssen, C. (2014). TCCON data from Paris (FR), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Thoning, K. W., Tans, P. P., and Komhyr, W. D. (1989). Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985. *J. Geophys. Res.-Atmos.*, 94(D6):8549–8565.
- Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, 13(14):4291–4313, doi:10.5194/bg-13-4291-2016.
- Trumbore, S. (2000). Age of soil organic matter and soil respiration: radiocarbon constraints on belowground c dynamics. *Ecological Applications*, 10(2):399–411.
- Van der Tol, C., Berry, J., Campbell, P., and Rascher, U. (2014). Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. *Journal of Geophysical Research: Biogeosciences*, 119(12):2312–2327.
- van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr, A. F. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmos. Chem. Phys.*, 6(11):3423–3441.
- van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S. (2017). Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data*, 9(2):697–720, doi:10.5194/essd-9-697-2017.
- Verhulst, K. R., Karion, A., Kim, J., Salameh, P. K., Keeling, R. F., Newman, S., Miller, J., Sloop, C., Pongetti, T., Rao, P., Wong, C., Hopkins, F. M., Yadav, V., Weiss, R. F., Duren, R. M., and Miller, C. E. (2017). Carbon dioxide and methane measurements from the Los Angeles megacity carbon project – part 1: calibration, urban enhancements, and uncertainty estimates. *Atmos. Chem. Phys.*, 17(13):8313–8341, doi:10.5194/acp-17-8313-2017.
- Viatte, C., Lauvaux, T., Hedelius, J. K., Parker, H., Chen, J., Jones, T., Franklin, J. E., Deng, A. J., Gaudet, B., Verhulst, K., et al. (2017). Methane emissions from dairies in the Los Angeles Basin. *Atmos. Chem. Phys.*, 17(12):7509–7528.

- Vidale, P. and Stöckli, R. (2005). Prognostic canopy air space solutions for land surface exchanges. *Theor. Appl. Climatol.*, 80(2-4):245–257.
- Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., Jung, M., Varlagin, A., and Guanter, L. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global Change Biol.*, 22(9):2979–2996, doi:10.1111/gcb.13200.
- Wang, W., Ciais, P., Nemani, R. R., Canadell, J. G., Piao, S., Sitch, S., White, M. A., Hashimoto, H., Milesi, C., and Myneni, R. B. (2013). Variations in atmospheric CO₂ growth rates coupled with tropical temperature. *P. Natl. Acad. Sci.*, 110(32):13061–13066.
- Wang, Y.-P., Leuning, R., Cleugh, H. A., and Coppin, P. A. (2001). Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? *Global Change Biology*, 7(5):495–510.
- Waring, R. H. and Running, S. W. (2010). *Forest ecosystems: analysis at multiple scales*. Elsevier.
- Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., Grupe, P., Vuillemin, C., Truong, F., Schmidt, M., Ramonet, M., and Parmentier, E. (2014). TCCON data from Orléans (FR), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Waxman, E. M., Cossel, K. C., Truong, G.-W., Giorgetta, F. R., Swann, W. C., Coburn, S., Wright, R. J., Rieker, G. B., Coddington, I., and Newbury, N. R. (2017). Intercomparison of open-path trace gas measurements with two dual-frequency-comb spectrometers. *Atmospheric Measurement Techniques*, 10(9):3295–3311, doi:10.5194/amt-10-3295-2017.
- Wehr, R., Munger, J., McManus, J., Nelson, D., Zahniser, M., Davidson, E., Wofsy, S., and Saleska, S. (2016). Seasonality of temperate forest photosynthesis and daytime respiration. *Nature*, 534(7609):680.
- Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R., Schaefer, K., Jacobson, A. R., Lu, C., et al. (2014). The North American carbon program multi-scale synthesis and terrestrial model intercomparison project—Part 2:

- environmental driver data. *Geosci. Model Dev.*, 7(6):2875–2893, doi:10.5194/gmd-7-2875-2014.
- Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., et al. (2012). On the sources of methane to the Los Angeles atmosphere. *Environmental science & technology*, 46(17):9282–9289.
- Wennberg, P. O., Roehl, C., Blavier, J.-F., Wunch, D., Landeros, J., and Allen, N. (2014a). TCCON data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2014R0. <http://dx.doi.org/10.14291/tccon.ggg2014.jpl02.R0/1149297>. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.
- Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R. a., Keppel-Aleks, G., Allen, N., and Ayers, J. (2014b). TCCON data from Park Falls (US), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA.
- Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., Dowell, P., Teske, K., Martin, C., and Martin, J. (2016). TCCON data from Lamont (US), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA.
- Wiacek, A., Li, L., Tobin, K., and Mitchell, M. (2018). Characterization of trace gas emissions at an intermediate port. *Atmos. Chem. Phys. Discussions*, 2018:1–54, doi:10.5194/acp-2017-1153.
- Wohlfahrt, G., Bahn, M., Haslwanter, A., Newesely, C., and Cernusca, A. (2005). Estimation of daytime ecosystem respiration to determine gross primary production of a mountain meadow. *Agr. Forest Meteorol.*, 130(1-2):13–25.
- Wong, C. K., Pongetti, T. J., Oda, T., Rao, P., Gurney, K. R., Newman, S., Duren, R. M., Miller, C. E., Yung, Y. L., and Sander, S. P. (2016). Monthly trends of methane emissions in Los Angeles from 2011 to 2015 inferred by CLARS-FTS observations. *Atmospheric Chemistry and Physics*, 16(20):13121–13130.
- Wong, K., Fu, D., Pongetti, T., Newman, S., Kort, E., Duren, R., Hsu, Y.-K., Miller, C., Yung, Y., and Sander, S. (2015). Mapping $\text{ch}_4:\text{co}_2$ ratios in Los Angeles with CLARS-FTS from Mount Wilson, California. *Atmospheric Chemistry and Physics*, 15(1):241–252.

- Wood, J. D., Griffis, T. J., Baker, J. M., Frankenberg, C., Verma, M., and Yuen, K. (2017). Multiscale analyses of solar-induced fluorescence and gross primary production. *Geophys. Res. Lett.*, 44(1):533–541, doi:10.1002/2016GL070775. 2016GL070775.
- Wu, S., Mickley, L. J., Jacob, D. J., Logan, J. A., Yantosca, R. M., and Rind, D. (2007). Why are there large differences between models in global budgets of tropospheric ozone? *J. Geophys. Res.-Atmos.*, 112(D5).
- Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W., Sherlock, V., and Wennberg, P. O. (2011). The Total Carbon Column Observing Network. *Philos. T. Roy. Soc. A*, 369(1943):2087–2112, doi:10.1098/rsta.2010.0240.
- Wunch, D., Toon, G. C., Hedelius, J. K., Vizenor, N., Roehl, C. M., Saad, K. M., Blavier, J.-F. L., Blake, D. R., and Wennberg, P. O. (2016). Quantifying the loss of processed natural gas within California’s South Coast Air Basin using long-term measurements of ethane and methane. *Atmospheric Chemistry and Physics*, 16(22):14091.
- Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A. (2010). Calibration of the total carbon column observing network using aircraft profile data. *Atmospheric Measurement Techniques*, 3(5):1351–1362, doi:10.5194/amt-3-1351-2010.
- Wunch, D., Wennberg, P., Messerschmidt, J., Parazoo, N., Toon, G., Deutscher, N., Keppel-Aleks, G., Roehl, C., Randerson, J., Warneke, T., et al. (2013). The covariation of Northern Hemisphere summertime CO₂ with surface temperature in boreal regions. *Atmos. Chem. Phys.*, 13(18):9447–9459.
- Wunch, D., Wennberg, P., Toon, G., Keppel-Aleks, G., and Yavin, Y. (2009). Emissions of greenhouse gases from a North American megacity. *Geophys. Res. Lett.*, 36(15).
- Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O’Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velasco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R.,

- Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A. (2017). Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) x_{CO_2} measurements with TCCON. *Atmos. Meas. Tech.*, 10(6):2209–2238, doi:10.5194/amt-10-2209-2017.
- Xueref-Remy, I., Dieudonné, E., Vuillemin, C., Lopez, M., Lac, C., Schmidt, M., Delmotte, M., Chevallier, F., Ravetta, F., Perrussel, O., Ciais, P., Bréon, F.-M., Broquet, G., Ramonet, M., Spain, T. G., and Ampe, C. (2018). Diurnal, synoptic and seasonal variability of atmospheric CO₂ in the paris megacity area. *Atmos. Chem. Phys.*, 18(5):3335–3362, doi:10.5194/acp-18-3335-2018.
- Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A., and Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophys. Res. Lett.*, 42(8):2977–2987.
- Yang, Z., Washenfelder, R., Keppel-Aleks, G., Krakauer, N., Randerson, J., Tans, P., Sweeney, C., and Wennberg, P. (2007). New constraints on northern hemisphere growing season net flux. *Geophysical Research Letters*, 34(12).
- Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G., Wunch, D., et al. (2013). Improvement of the retrieval algorithm for GOSAT SWIR XCO₂ and XCH₄ and their validation using TCCON data.
- You, Y., Staebler, R. M., Moussa, S. G., Su, Y., Munoz, T., Stroud, C., Zhang, J., and Moran, M. D. (2017). Long-path measurements of pollutants and micrometeorology over Highway 401 in Toronto. *Atmos. Chem. Phys.*, 17(22):14119–14143.
- Yu, K., Keller, C. A., Jacob, D. J., Molod, A. M., Eastham, S. D., and Long, M. S. (2018). Errors and improvements in the use of archived meteorological data for chemical transport modeling: an analysis using GEOS-Chem v11-01 driven by GEOS-5 meteorology. *Geosci. Model Dev.*, 11(1):305–319, doi:10.5194/gmd-11-305-2018.
- Zeng, N., Zhao, F., Collatz, G. J., Kalnay, E., Salawitch, R. J., West, T. O., and Guanter, L. (2014). Agricultural green revolution as a driver of increasing atmospheric CO₂ seasonal amplitude. *Nature*, 515(7527):394–397.

- Zhang, Y., Guanter, L., Berry, J. A., van der Tol, C., Yang, X., Tang, J., and Zhang, F. (2016a). Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications. *Remote Sens. Environ.*, 187:145–155.
- Zhang, Y., Joiner, J., Gentine, P., and Zhou, S. (2018). Reduced solar-induced chlorophyll fluorescence from GOME-2 during Amazon drought caused by dataset artifacts. *Glob. Change Biol.*, 24(6):2229–2230.
- Zhang, Y., Xiao, X., Jin, C., Dong, J., Zhou, S., Wagle, P., Joiner, J., Guanter, L., Zhang, Y., Zhang, G., et al. (2016b). Consistency between sun-induced chlorophyll fluorescence and gross primary production of vegetation in North America. *Remote Sens. Environ.*, 183:154–169.
- Ziehn, T., Scholze, M., and Knorr, W. (2011). Development of an ensemble-adjoint optimization approach to derive uncertainties in net carbon fluxes. *Geoscientific Model Development*, 4(4):1011–1018.
- Ziehn, T., Scholze, M., and Knorr, W. (2012). On the capability of Monte Carlo and adjoint inversion techniques to derive posterior parameter uncertainties in terrestrial ecosystem models. *Global Biogeochemical Cycles*, 26(3).
- Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., et al. (2016). Cold season emissions dominate the Arctic tundra methane budget. *Proc. Natl. Acad. Sci.*, 113(1):40–45.
- Zscheischler, J., Mahecha, M. D., Von Buttler, J., Harmeling, S., Jung, M., Rammig, A., Randerson, J. T., Schölkopf, B., Seneviratne, S. I., Tomelleri, E., et al. (2014). A few extreme events dominate global interannual variability in gross primary production. *Environ. Res. Lett.*, 9(3):035001.