

DISSERTATION

**Moist Synoptic Transport of CO<sub>2</sub> Along Midlatitude Storm  
Tracks, Transport Uncertainty, and Implications for Flux  
Estimation**

Submitted by

Nicholas C. Parazoo

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2011

COLORADO STATE UNIVERSITY

June 20, 2011

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY NICHOLAS C. PARAZOO ENTITLED Moist Synoptic Transport of CO<sub>2</sub> Along Midlatitude Storm Tracks, Transport Uncertainty, and Implications for Flux Estimation BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

---

---

---

---

Advisor

---

Department Head

## ABSTRACT OF DISSERTATION

### Moist Synoptic Transport of CO<sub>2</sub> Along Midlatitude Storm Tracks, Transport Uncertainty, and Implications for Flux Estimation

Mass transport along moist isentropic surfaces on baroclinic waves represents an important component of the atmospheric heat engine that operates between the equator and poles. This is also an important vehicle for tracer transport, and is correlated with ecosystem metabolism because large-scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. In this research, I pursue a dynamical framework for explaining atmospheric transport of CO<sub>2</sub> by synoptic weather systems at middle and high latitudes. A global model of atmospheric tracer transport, driven by meteorological analysis in combination with a detailed description of surface fluxes, is used to create time varying CO<sub>2</sub> distributions in the atmosphere. Simulated mass fluxes of CO<sub>2</sub> are then decomposed into a zonal monthly mean component and deviations from the monthly mean in space and time. Mass fluxes of CO<sub>2</sub> are described on moist isentropic surfaces in order to include transport along frontal systems in the eddy-terms rather than in the mean.

Synoptic weather systems transport large amounts of CO<sub>2</sub> north and south in northern mid-latitudes, up to 1 PgC month<sup>-1</sup> during winter when baroclinic waves ramp up. During boreal winter when northern plants respire, warm moist air, high in CO<sub>2</sub>, is swept upward and poleward along the east side of baroclinic waves and dumped into the polar vortex, while cold dry air, low in CO<sub>2</sub>, that had been transported into the polar vortex earlier in the year is swept equatorward. Synoptic eddies strongly reduce seasonality in the

biologically active mid-latitudes by 50% of that implied by net ecosystem exchange while amplifying seasonality in the Arctic. Transport along stormtracks is correlated with rising, moist, cloudy air, which systematically hides this CO<sub>2</sub> transport from satellites. Meridional fluxes of CO<sub>2</sub> are of comparable magnitude as surface exchange of CO<sub>2</sub> in mid-latitudes, and thus require careful consideration in (inverse) modeling of the carbon cycle.

Because synoptic transport of CO<sub>2</sub> by frontal systems and moist processes is unobserved and poorly represented in global models, it may be a source of error for inverse flux estimates. Uncertainty in CO<sub>2</sub> transport by synoptic eddies is investigated using a global model driven by four reanalysis products from the Goddard EOS Data Assimilation System for 2005. Eddy transport is found to be highly variable between models, with significant seasonal biases of up to 0.2 PgC, representing up to 50% of fossil fuel emissions, caused primarily by differences in grid spacing and vertical mixing by moist convection and PBL turbulence. To test for aliasing of transport bias into inverse flux estimates, synthetic satellite data is generated using a model at global mesoscale resolution and inverted using a global model run with coarse transport. An ensemble filtering method called the Maximum Likelihood Ensemble Filter (MLEF) is used to optimize fluxes. Flux estimates are found to be highly sensitive to transport biases at pixel and continental scale, with errors of up to 0.5 PgC year<sup>-1</sup> in Europe and North America.

Nicholas C. Parazoo  
Department of Atmospheric Science  
Colorado State University  
Fort Collins, Colorado 80523-1371  
Summer 2011



## ACKNOWLEDGEMENTS

# CONTENTS

<b>1 Background and Objectives</b>	<b>1</b>
1.1 Background . . . . .	1
1.1.1 Atmospheric Observations . . . . .	4
1.1.2 Transport Models . . . . .	9
1.1.3 Optimization Technique . . . . .	17
1.2 Objectives of this Study . . . . .	20
<b>2 Experimental Methods</b>	<b>24</b>
2.1 Model Description . . . . .	24
2.1.1 PCTM . . . . .	24
2.2 Surface Boundary Conditions . . . . .	26
2.2.1 Terrestrial Fluxes . . . . .	26
2.2.2 Fossil Fuel Fluxes . . . . .	28
2.2.3 Oceanic Fluxes . . . . .	28
2.3 Transport Driver Data . . . . .	30
2.3.1 GEOS4-1.25x1 . . . . .	30
2.3.2 GEOS4-2.5x2 . . . . .	31
2.3.3 GEOS5-0.67x0.5 . . . . .	31
2.3.4 GEOS5-1.25x1 . . . . .	32
2.4 Transport Simulations for Parts 1 and 2 . . . . .	32
2.5 Methodology for Transport Analysis . . . . .	34
2.5.1 Tendency Equation for Atmospheric $CO_2$ . . . . .	35
2.5.2 Eddy Decomposition . . . . .	37
2.5.3 Vertical Coordinate . . . . .	40
2.6 Methodology for Inversion used in Part 3 . . . . .	44
2.6.1 Assimilation Scheme . . . . .	44
2.6.2 Synthetic Satellite Data . . . . .	46
2.6.3 MLEF . . . . .	50
2.6.4 Covariance Smoothing and Localization . . . . .	53
2.6.5 Defining Priors . . . . .	53
2.6.6 Terminology . . . . .	54
2.6.7 Flux and Uncertainty Calculation . . . . .	55

<b>3</b>	<b>Moist Synoptic Transport of CO<sub>2</sub> Along the Midlatitude Storm Track</b>	<b>57</b>
3.1	Introduction . . . . .	57
3.2	The Zonal Mean Circulation - Streamfunctions . . . . .	59
3.2.1	Streamfunction on Terrain Following Surfaces ( $\eta$ ) . . . . .	60
3.2.2	Streamfunction on $\theta$ . . . . .	62
3.2.3	Streamfunction on $\theta_e$ . . . . .	64
3.3	Energy Transport . . . . .	66
3.4	Seasonal Covariance of Baroclinicity and Ecosystem Metabolism . . . . .	68
3.5	Meridional Transport of CO <sub>2</sub> by Mean and Eddy Circulations . . . . .	70
3.5.1	CO <sub>2</sub> Transport by the MMC . . . . .	71
3.5.2	CO <sub>2</sub> Transport by Synoptic Eddies . . . . .	74
3.5.3	Vertically Integrated Transport . . . . .	82
3.5.4	Dependence on Vertical Coordinate . . . . .	84
3.6	Seasonal Variations . . . . .	89
3.7	Conceptual Diagram . . . . .	89
3.8	Discussion and Conclusions from Part 1 . . . . .	92
<b>4</b>	<b>Uncertainty in Moist Synoptic Transport of CO<sub>2</sub></b>	<b>95</b>
4.1	Introduction . . . . .	96
4.2	Comparison to Observations . . . . .	99
4.3	Estimates of Transport Uncertainty . . . . .	101
4.4	Causes of Transport Uncertainty . . . . .	107
4.5	Discussion and Conclusions from Part 2 . . . . .	115
<b>5</b>	<b>Implications of Transport Bias for Inverse Flux Estimation</b>	<b>119</b>
5.1	Introduction . . . . .	119
5.2	Control Experiments - Perfect Transport . . . . .	123
5.2.1	Experiment 1: Perfect Transport, Unbiased Truth, and no Cloud Screening . . . . .	123
5.2.2	Experiment 2: Perfect Transport, Unbiased Truth, and Cloud Screening	131
5.2.3	Aggregation of Experiments 1 and 2 up to TransCom Regions . . . . .	133
5.3	Bias Recovery and Perfect Transport . . . . .	140
5.3.1	Experiment 3: Perfect Transport, Constant and Unperturbed Biases, and Cloud Screening . . . . .	143
5.3.2	Experiment 4: Perfect Transport, Seasonal and Noisy Biases, and Cloud Screening . . . . .	149
5.4	Biased Transport . . . . .	150
5.4.1	Experiment 5: Biased Transport, Unbiased Truth, and Cloud Screening	153
5.4.2	Experiment 6: Biased Transport, Biased Truth, and Cloud Screening	156
5.5	Discussion and Conclusions from Part 3 . . . . .	161
<b>6</b>	<b>Conclusions and Future Work</b>	<b>165</b>
	<b>Bibliography</b>	<b>170</b>

## FIGURES

1.1	Components of the global CO <sub>2</sub> budget [taken from Le Quere et al., 2009]. . .	2
1.2	World map showing the NOAA ESRL/GMD CCGG Cooperative Air Sampling Network . . . . .	5
1.3	Earth's net radiation budget [courtesy of Pearson Prentice Hall, Inc]. . . . .	12
1.4	Illustration of frontal transports associated with a baroclinic wave. . . . .	13
1.5	Model representation of a cold front and associated divergence / convergence couplet (courtesy The COMET Program, <a href="http://www.meted.ucar.edu/">http://www.meted.ucar.edu/</a> ). . .	16
2.1	SiB3 Biome Map for GiMMsG NDVI. Vegetation Type by Number on Color Bar: 1) C3 Tall Broadleaf-Evergreen Trees, 2) C3 Tall Broadleaf-Deciduous Trees, 3) C3 Tall Broadleaf and Needleleaf Trees, 4) C3 Tall Needleleaf Trees, 5) C3 Tall Needleleaf-Deciduous Trees, 6) C4 Short Vegetation, Same as Types 6, 7, 8, 11 7) C4 Short Vegetation (Maize Optical Properties), 8) Same as 7, 9) Short Broadleaf Shrubs with Bare Soil, 10) C3 Short Ground Cover (Tundra), 11) C4 No Vegetation (Low Lat Desert), 12) Agriculture (Wheat) and C3 Grasslands, 13) Ice . . . . .	29
2.2	Number of points sampled by pixels in GEOS4-2.5x2 along the GOSAT orbit in one year for all days (a) and clear days (b). . . . .	47
2.3	Percent of points discarded due to cloud screening for winter (top left), spring (top right), summer (bottom left), and fall (bottom right). . . . .	49
3.1	The global mean circulation computed on Eta surfaces from GEOS4-1.25x1 for 2005. Contour interval is $2.5 \times 10^{10} \text{ kg s}^{-1}$ . Solid red contours are positive values of $\Psi$ and correspond counterclockwise circulations, with northward flow at low levels and southward flow at high levels. Dashed blue contours are negative values of $\Psi$ and correspond to clockwise circulations, with southward flow at low levels and northward flow at high levels. . . . .	61
3.2	Same as Figure 3.1, except computed on potential temperature surfaces. . .	63
3.3	Same as Figure 3.2, except computed on equivalent potential temperature surfaces. . . . .	65
3.4	Annual mean transport of moist static energy by various components of the mid-latitude circulation. Energy has units of PW month <sup>-1</sup> . . . . .	67
3.5	Seasonal covariance of moist static energy transport ( $h$ , black line) and zonally-monthly averaged terrestrial net ecosystem exchange (NEE, green line) for boreal winter (A, Dec-Jan-Feb) and boreal summer (B, Jun-Jul-Aug). NEE is plotted as a zonal mean in units of $\mu\text{mol m}^{-2} \text{ s}^{-1}$ on the right y-axis. . . . .	69

3.6	Diagnosis of vertical distribution of atmospheric CO <sub>2</sub> transport by the MMC on $\theta_e$ in the annual mean (A) and during boreal winter (B) and summer (C). Red (blue) contours indicate poleward (equatorward) transport. Units are PgC K <sup>-1</sup> month <sup>-1</sup> . . . . .	72
3.7	Zonal and monthly mean distribution of CO <sub>2</sub> (shaded) and mass flux (contours) in the annual mean (A) and during boreal winter (B) and summer (C). Shaded values are represent by the colorbar in (C), where the mean value in the column has been subtracted (hence “CO <sub>2</sub> anomaly”). Contour intervals of mass fluxes are 0.05 Pg s <sup>-1</sup> K <sup>-1</sup> . Red solid (blue dashed) contours correspond to poleward (equatorward) flow. . . . .	73
3.8	Diagnosis of vertical distribution of atmospheric CO <sub>2</sub> transport by synoptic eddies (stationary and transient waves combined) on $\theta_e$ in the annual mean (A) and during boreal winter (B) and summer (C). Red (blue) contours indicate poleward (equatorward) transport. Units are PgC K <sup>-1</sup> month <sup>-1</sup> . . . . .	75
3.9	Case study from 13 Jan 2005 showing an example of eddy CO <sub>2</sub> transport associated with a typical mid-latitude cyclone. The surface composite map in (A) is reproduced with permission of Unisys Corporation (c) 2011. Model output of anomalous CO <sub>2</sub> (ppm, shaded) and mass flux (1x10 <sup>-4</sup> kg m <sup>-1</sup> K <sup>-1</sup> , contour) along $\theta_e = 300$ K is plotted in (B). . . . .	77
3.10	Annual mean correlation coefficient of mass flux and CO <sub>2</sub> deviations for (A) transient eddies and (B) stationary eddies (bottom plot), plotted as a function of $\theta_e$ . Colorbar is the correlation coefficient. . . . .	79
3.11	The annual mean spatial distribution of column integrated transport by (A) transient and stationary eddies, (B) transient eddies and (C) stationary eddies. The colorbar for each plot is shown at the bottom in units of PgC month <sup>-1</sup> . . . . .	81
3.12	Column integrated transport in the annual mean (A) and during boreal winter (B) and summer (C). Total transport (black line) is parsed into transport by total eddy (red, transient + stationary), transient eddy (magenta), stationary eddy (cyan) and mean (blue) components. . . . .	83
3.13	Column integrated transport in the annual mean as a function of vertical coordinate. Transport is calculated on moist isentropic surfaces ( $\theta_e$ , solid line), dry isentropic surfaces ( $\theta$ , dashed line), and terrain following surfaces ( $\eta$ , dotted line). Total transport (black) is partitioned to transport by eddy (red) and mean (blue) components. . . . .	84
3.14	Zonal and month mean distribution of CO <sub>2</sub> (shaded) and meridional mass flux (contours) in the annual mean on $\theta_e$ (A) and $\eta$ (B). The colorbar for CO <sub>2</sub> (in ppm) is shown on the bottom. CO <sub>2</sub> is plotted on the same scale. . . . .	87
3.15	Diagram of isentropic surfaces and air streams along polar front (courtesy <a href="http://www-das.uwyo.edu/geerts/cwx/notes/chap01/tropo.html">http://www-das.uwyo.edu/geerts/cwx/notes/chap01/tropo.html</a> ). . . . .	88

3.16	Column integrated seasonal CO <sub>2</sub> budget tendencies (ppm month <sup>-1</sup> ) on $\theta_e$ binned into (A) high latitudes and (B) mid-latitudes. Transport by transient and stationary eddies is plotted in blue, mean meridional transport in red, the total CO <sub>2</sub> tendency in black, and the total surface flux of carbon from land, ocean, and fossil fuels in green. The fossil fuel tendency is plotted as a dashed green line. The sum of individual tendencies (red, blue and green lines) is equal to the total tendency (black line). Error bars represent the root mean squared error for the two years of averaging used. . . . .	90
3.17	Illustration of important CO <sub>2</sub> transport mechanisms associated with a baroclinic wave. . . . .	91
4.1	Lag Correlations between model and observations at mid-latitude sites in North America and Europe. The four line styles refers to the model used in the correlation, where GEOS5-0.67x0.5 is solid, GEOS5-1.25x1 as dashed, GEOS4-1.25x1 is dotted, and GEOS4-2.5x2 is dash-dotted. These same line styles are used in the remaining figures in the Chapter. . . . .	100
4.2	Modeled and observed surface CO <sub>2</sub> mixing ratio from January 1-30, 2005 at a site in northern Wisconsin (LEF). Observations are plotted at 76 m above the surface in black. Line styles for models are the same as in Figure 4.1 . .	101
4.3	Total column integrated CO <sub>2</sub> transport (black) partitioned into transport by the mean circulation (blue) and synoptic eddies (baroclinic + stationary waves, red) in the annual mean (A) and during boreal winter (Jan-Feb, B) and summer (Jun-Jul-Aug, C). . . . .	102
4.4	Column integrated seasonal CO <sub>2</sub> budget tendencies (ppm month <sup>-1</sup> ) for eddies binned into high latitudes (70-90°N, A) and mid-latitudes (30-70°N, B). Tendencies due to total surface flux (green solid) and fossil fuel emissions (green dashed) are shown in (B). . . . .	104
4.5	CO <sub>2</sub> transport by stationary waves (cyan) and baroclinic waves (magenta) along $\theta_e$ in the annual mean (A) and during boreal winter (B) and boreal summer (C). . . . .	106
4.6	Spatial structure of annual mean column integrated CO <sub>2</sub> transport by synoptic eddies along $\theta_e$ for GEOS4-2.5x2 (A), GEOS4-1.25x1 (B), GEOS5-1.25x1 (C) and GEOS5-0.67x0.5 (D). The colorbar is shown in units of PgC month <sup>-1</sup> . . . . .	108
4.7	Column CO <sub>2</sub> snapshots on 10 January 2005 for GEOS5-0.67x0.5 (a) and GEOS4-2.5x2 (b) . . . . .	110
4.8	Zonal-annual mean values for vertical diffusion (A) and cumulus mass flux (B) binned into mid-latitudes (30-70°N) and plotted as a function of the terrain following coordinate $\eta$ for each reanalysis product. These values are taken directly from the driver files used to force vertical mixing in PCTM. GEOS4 lines are identical and lie directly over one another. . . . .	111
4.9	Zonal-annual mean CO <sub>2</sub> mixing ratio (ppm) binned into mid-latitudes (30-70°N) and plotted as a function of $\eta$ . . . . .	112
4.10	Eddy CO <sub>2</sub> transport by GEOS5-1.25x1 (A) and GEOS4-1.25x1 (B). Difference in eddy transport between models are shown in (C). . . . .	113

4.11	Mass transport of low-level cold dry air (blue line), low-level warm moist air (red line), and upper-level warm dry air (black line) as define in Eqs. (5a)-(5c) of Pauluis et al. [2009]. . . . .	114
5.1	Annual mean errors in terrestrial NEE in units of $\mu\text{moles } m^{-2} s^{-1}$ at pixel scale ( $2.5^\circ \times 2^\circ$ ) for Experiment 1. . . . .	124
5.2	Similar to Figure 5.1 except plotted as a function of season for Experiment 1: (a) boreal winter, or December-January-February (DJF), (b) boreal spring, or March-April-May (MAM), (c) boreal summer, or June-July-August (JJA), and (d) boreal fall, or September-October-November (SON). The colorbar has the same scale as in Figure 5.1. . . . .	125
5.3	Seasonal distribution of prior estimate of terrestrial NEE, plotted on the same scale as Figure 5.1. These priors are used for all experiments in this chapter. . . . .	126
5.4	Seasonal flux errors for Experiment 1 plotted as a fraction of the prior. . . .	127
5.5	Annual mean uncertainties in Experiment 1 for (a) prior and (b) analyzed flux. Uncertainty reduction for terrestrial NEE is shown in the bottom plot (c). . . . .	128
5.6	Same as Figure 5.1 but for air-sea exchange for Experiment 1. Note difference in scale between land and ocean. . . . .	129
5.7	Same as Figure 5.6 but for prior air-sea exchange. . . . .	130
5.8	Similar to Figure 5.6 for Experiment 1, except plotted as a function of season (a) DJF, (b) MAM, (c) JJA, and (d) SON. The colorbar has the same scale as in Figure 5.6. . . . .	131
5.9	Same as Figure 5.8 except for air-sea exchange priors. . . . .	132
5.10	Seasonal errors in air-sea exchange in Experiment 1 plotted as a fraction of the prior. . . . .	133
5.11	Annual mean uncertainties in Experiment 1 for (a) prior and (b) analyzed air-seas exchange. Uncertainty reduction is shown in the bottom plot (c). . .	134
5.12	Same as Figure 5.1 except for Experiment 2. . . . .	135
5.13	Flux errors plotted as a function of season for Experiment 2: (a) boreal winter, or December-January-February (DJF), (b) boreal spring, or March-April-May (MAM), (c) boreal summer, or June-July-August (JJA), and (d) boreal fall, or September-October-November (SON). The colorbar has the same scale as in Figure 5.12. . . . .	136
5.14	Annual mean uncertainties in Experiment 2 for (a) prior and (b) analyzed flux. Uncertainty reduction for terrestrial NEE is shown in the bottom plot (c). . . . .	137
5.15	TransCom basis functions. . . . .	138
5.16	Bar plot of total annual fluxes from prior (dark blue), analysis from Experiment 1 (green), and analysis from Experiment 2 (red) with corresponding posterior uncertainty estimates (standard deviation), aggregated up to TransCom Regions (see Figure 5.15), in $\text{GtC year}^{-1}$ . The top 11 regions correspond to land, the next 11 regions correspond to ocean, and the bottom two regions correspond to total land and ocean. . . . .	139

5.17	Seasonal flux errors for TransCom land regions 1-11 for Experiment 1 (green) and Experiment 2 (red), plotted in GtC month <sup>-1</sup> . . . . .	140
5.18	Multiplicative correction factors $\beta_{GPP}$ (a), $\beta_{RESP}$ (b), and $\beta_{OCEAN}$ (c). . .	142
5.19	True (a) and recovered (b) annual mean terrestrial NEE in $\mu moles\ m^2\ s^{-1}$ for Experiment 3. . . . .	144
5.20	Seasonal cycle of prior (black line), true (red line), and recovered (blue line) terrestrial NEE in $\mu moles\ m^2\ s^{-1}$ at a forested location in Temperate N. America with coordinates [80°W, 40°N]. . . . .	145
5.21	Difference between true and recovered annual mean flux, in $\mu moles\ m^2\ s^{-1}$ , for Experiment 3. . . . .	146
5.22	Bar plot of annual prior flux (dark blue), true flux (green), and recovered flux (red), aggregated up to TransCom regions like Figure 5.16, and plotted in GtC year <sup>-1</sup> for Experiment 3. Posterior uncertainty is plotted in red. . .	147
5.23	Seasonal cycle of prior (black), true (red), and recovered (blue) TransCom fluxes for Experiment 3 (biased truth) for land regions 1-11, plotted in GtC month <sup>-1</sup> . Posterior (red) uncertainties are plotted as error bars. . . . .	148
5.24	True (a) and recovered (b) annual mean terrestrial NEE in $\mu moles\ m^2\ s^{-1}$ for Experiment 4. . . . .	150
5.25	Bar plot of annual prior flux (dark blue), true flux (green), and recovered fluxes (red) for Experiment 4. . . . .	151
5.26	Annual mean flux errors over land for Experiment 5 (unbiased truth), plotted in $\mu moles\ m^{-2}\ s^{-1}$ at pixel scale (2.5°x2°). . . . .	153
5.27	Seasonal flux errors over land for Experiment 5. . . . .	154
5.28	Bar plot of annual prior (dark blue), flux errors from Experiment 1 (light blue, perfect transport and no cloud screening), Experiment 2 (yellow, perfect transport and cloud screening), and Experiment 5 (red, biased transport and cloud screening), aggregated up to TransCom land regions and plotted in GtC year <sup>-1</sup> . . . . .	155
5.29	Seasonal cycle of true (black) and recovered (red) TransCom fluxes for Experiment 5 (unbiased truth) for land regions 1-11, plotted in GtC month <sup>-1</sup> . Prior (blue) and posterior (red) uncertainties are plotted as error bars. . . .	157
5.30	Bar plot of annual prior (dark blue), flux errors from Experiment 4 (green, perfect transport and monthly varying sinks), and Experiment 6 (red, biased transport and monthly varying sinks), aggregated up to TransCom land regions and plotted in GtC year <sup>-1</sup> . . . . .	158
5.31	Seasonal cycle of true (black) and recovered (red) TransCom fluxes for Experiment 6 (biased truth) for land regions 1-11, plotted in GtC month <sup>-1</sup> . Prior (blue) and posterior (red) uncertainties are plotted as error bars. . . .	159
5.32	Seasonal flux errors for TransCom land regions 1-11 for Experiment 5 (blue) and Experiment 6 (green), plotted in GtC month <sup>-1</sup> . . . . .	160



## TABLES

## Chapter 1

### BACKGROUND AND OBJECTIVES

#### 1.1 Background

CO<sub>2</sub> emissions from fossil fuel combustion, including small contributions from cement production and gas flaring, were about  $8.7 \pm 0.5$  PgC yr<sup>-1</sup> in 2008, with emissions increasing at a rate of 3.4% per year between 2000 and 2008 [Le Quere et al., 2009, see Figure 1.1]. Fossil fuel emissions are expected to double over the next 50 years under business as usual conditions [Pacala and Socolow, 2004], and likely more as new economies emerge [Le Quere et al., 2009], causing atmospheric CO<sub>2</sub> concentrations to nearly triple from pre-industrial levels. It is even worse considering land use change (LUC) emissions due to anthropogenic activities such as deforestation, logging, and intensive cultivation of cropland soils (see Figure 1.1b). Although there is much interannual variability in LUC fluxes, especially in tropical regions, the 2008 LUC emissions were about 1.2 PgC yr<sup>-1</sup>, bringing total anthropogenic CO<sub>2</sub> emissions to  $9.9 \pm 0.9$  PgC yr<sup>-1</sup> in 2008.

Rising CO<sub>2</sub> is expected to alter radiative forcing, warm climate, and cause significant global climatic changes in the near future, including natural and socioeconomic impacts. Projections of future CO<sub>2</sub> levels and the associated climate forcing depends strongly on our scientific understanding of both the anthropogenic and natural components of the carbon cycle and our ability to accurately predict future sources and sinks of carbon [Cox et al., 2000; Friedlingstein et al., 2006; IPCC, 2007]. The existing network of surface in situ CO<sub>2</sub> measurement stations [GLOBALVIEW-CO<sub>2</sub>, 2009] suggest that on average 43% of total CO<sub>2</sub> emissions each year between 1959 and 2008 remain in the atmosphere (see Figure 1.1a,

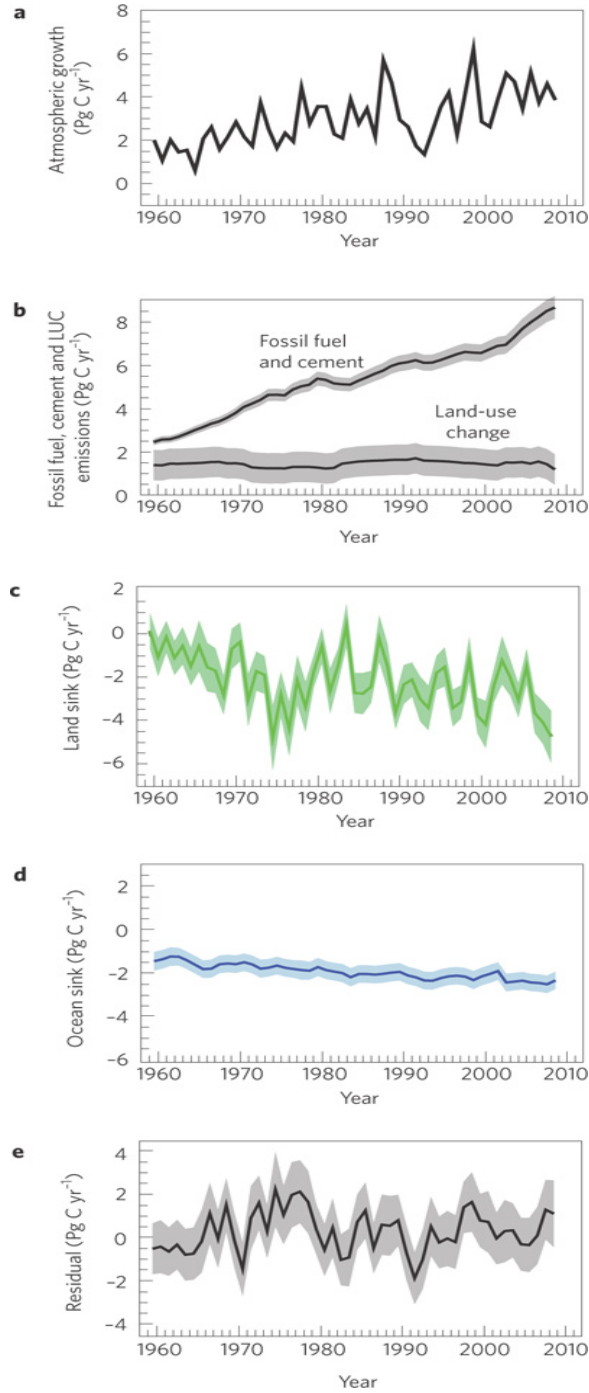


Figure 1.1: Components of the global CO<sub>2</sub> budget [taken from Le Quere et al., 2009].

Le Quere et al. [2009]). The remaining fraction, roughly 4-5 PgC yr<sup>-1</sup>, is absorbed naturally by land and ocean sinks (see Figure 1.1 c and d, respectively). Atmospheric and ocean observations constrain the global land sink to 2.6 $\pm$ 0.7 and the ocean sink to 2.2 $\pm$ 0.4

PgC yr<sup>-1</sup> for 1990-2000 [McNeil et al., 2003; Manning and Keeling, 2006; Canadell et al., 2007; Denman et al., 2007; Gruber et al., 2009]. This natural sink, however, is not guaranteed to grow or even sustain its current capacity [Fung et al., 2005; Le Quere et al., 2009]. Natural emissions might actually increase, for example, as a result of already observable rapid warming in parts of the Arctic [Hinzman et al., 2005] due to mobilization of carbon currently stored in permafrost [Zimov et al., 2006; Khvorostyanov et al., 2008].

One of the first studies to determine that a combination of terrestrial and oceanic processes are responsible for net removal of CO<sub>2</sub> from that atmosphere was Tans et al. [1990], who compared relatively well known fossil fuel emissions and the corresponding north-south gradient in atmospheric CO<sub>2</sub> as determined by a three-dimensional atmospheric general circulation model to the north-south gradient observed from a global network of flask measurements. Combined with pCO<sub>2</sub> measurements to constrain the ocean sink, this technique implied a Northern Hemisphere (NH) terrestrial sink, but was unable to reveal information about regional processes that determine the sink. It is understanding regional processes, and how they determine the time and spatially varying source/sink distribution on synoptic, seasonal, interannual (e.g., Figure 1.1c) and climatic time scales, that has motivated more recent ( $\sim$  mid 1990's to present) research.

Variations of atmospheric CO<sub>2</sub> contain information about sources and sinks which air interacts with as it is transported from place to place. Atmospheric inverse modelers combine tracer transport models with atmospheric measurements of CO<sub>2</sub> mixing ratio to optimize estimates of the strength and spatiotemporal distribution of sources and sinks [Gurney et al., 2002; Rödenbeck et al., 2003; Baker et al., 2006]. Results are limited by (1) data availability and precision; (2) model fidelity; and (3) optimization technique. The following sections describe in more detail these three important components of inversion systems.

### 1.1.1 *Atmospheric Observations*

Atmospheric observations of CO<sub>2</sub> mixing ratio include in situ and remotely sensed measurements. The in situ network consists of surface observations by flask (weekly) and continuous (hourly) instruments, which measure at a fixed location continuously in time, as well as airborne measurements taken from aircraft, which can sample most parts of the atmosphere but tend to have fair weather bias. Remotely sensed measurements are typically column integrated and taken from aircraft or satellite. Data accuracy and density are important constraints for inversions, as illustrated by the inversion intercomparison experiment, TransCom. For example, in an annual mean inversion of flask data for 22 source regions, TransCom found uncertainties in surface flux estimates related to data sparsity ranging from 0.2-1.1 GtC yr<sup>-1</sup> [Gurney et al., 2002], which represents as much as 25% of global sink estimates.

The TransCom 3 experiment, described in Gurney et al. [2002, 2003, 2004], was based on flask measurements, which are taken under “clean-air” or “baseline” conditions in order to sample air representative of large-scale air masses. While flasks are ideally suited for making precise measurements of background concentrations, they are not designed to measure continental or any kind of fine scale variations due to nearby terrestrial sources and sinks. Continuous measurements are therefore a nice complement to the flask network, as can be seen in Figure 1.2,. Continuous measurement sites have existed for a number of years and the network continues to grow, especially in northern middle latitudes, and are extremely desirable for carbon cycle studies in that they are (1) collected on site and can sample more frequently (i.e., hourly) in time; (2) well-calibrated and high precision ( $\sim 0.1$  ppm); (3) typically located over continents close to terrestrial source and sinks; and (4) influenced by strong synoptic signals that contain upstream surface flux information. Because of continuous on-site monitoring, they are especially useful for sampling CO<sub>2</sub> mixing ratios contained in air masses advecting across continents at synoptic scales ( $\sim 1000$ km), with total upstream footprint proportional to measurement height above the surface [Bakwin et al.,

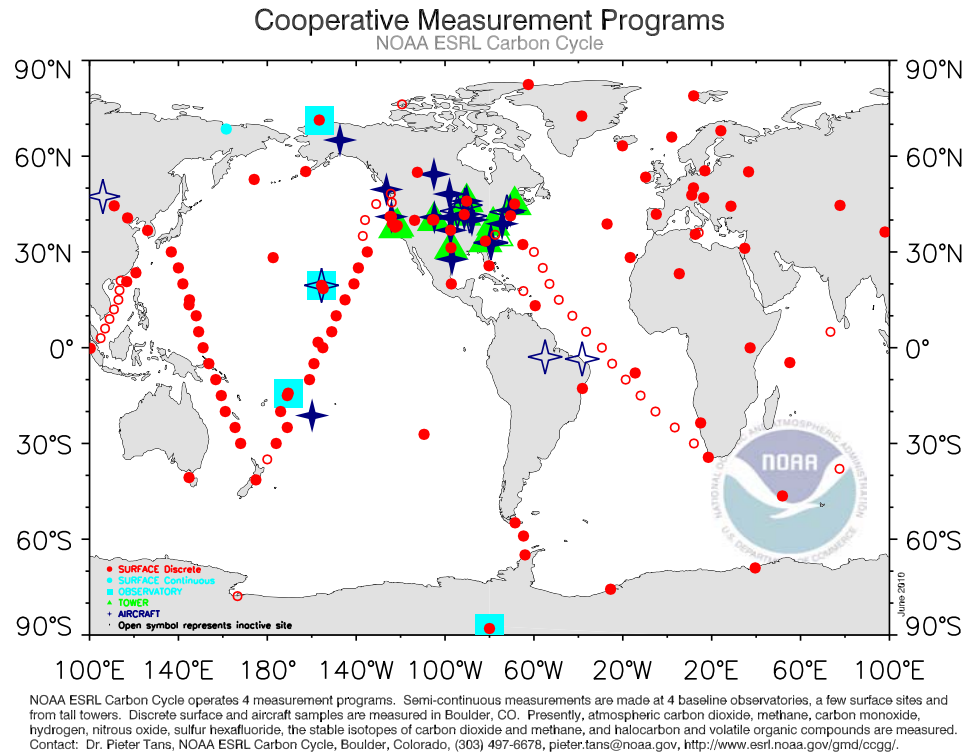


Figure 1.2: World map showing the NOAA ESRL/GMD CCGG Cooperative Air Sampling Network

1998]. Tall tower measurements from National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) Global Monitoring Division (GMD), for example, can measure continuously at up to 400 m above the surface and can “feel” the influence of surface fluxes for hundreds of kilometers, while surface measurements are typically much closer to the ground (i.e., < 30 m above ground level) and more strongly influenced by local processes.

There is an interesting relationship between atmospheric measurements and surface fluxes. It is desirable for measurements to contain as much information about upstream sources and sinks as possible. Other than process based land surface models, mixing ratio measurements are our primary source of information regarding ecosystem scale fluxes of

CO<sub>2</sub>. Measurements have a higher probability of influence from larger scales the higher a measurement is above the surface. This is where the tradeoff begins. In general, the more information we want to know about surface fluxes (i.e., the higher up and more frequent a measurement is) the better we (i.e., modelers) have to be at representing transport between flux and concentration.

Nevertheless, use of high frequency data in CO<sub>2</sub> inversions has been shown to reduce overall flux uncertainties [Law et al., 2002], even more so when combined with monthly mean flask observations [e.g., Rödenbeck et al., 2003]. While continuous measurements improve our process-based understanding of local and upstream terrestrial sources and sinks [Gloor et al., 2001; Worthy et al., 2003; Geels et al., 2004; Hurwitz et al., 2004], they contain signals that result from complex atmospheric processes. At first this may seem a bad thing, because this presents an interesting challenge for models of the atmosphere to correctly represent these processes (more detail in Section 1.1.2); that being said, much has been learned about the carbon cycle and numerical models of the carbon cycle as a result of these measurements. It is also worth mentioning here that aircraft measurements are very useful for profiling vertical distributions and have helped further constrain carbon budgets at regional scales when combined with tower measurements [Lauvaux et al., 2008].

A major shortcoming of the in-situ network, as seen in Figure 1.2, is that large geographic data gaps exist, especially over the oceans, in the Southern Hemisphere (SH) and in the tropics. The surface network, despite continuous monitoring at high precision, is insufficient to correct surface fluxes at regional scales. Satellites measure column averaged CO<sub>2</sub> mixing ratio with good spatial coverage, thus complementing data poor regions of surface measurements, but have strict precision requirements for top-down inversion because source and sink information is contained in column averaged variations, which are small (and diluted) in contrast to Planetary Boundary Layer (PBL) variations where surface networks reside. In an effort to evaluate the potential usefulness of remotely sensed mixing ratios, synthesis inversion models indicate that regional surface flux uncertainties could be

reduced substantially if in-situ data are combined with column-integrated measurements with precision of  $<1\%$ , or  $\sim 3$  ppm [Rayner and O'Brien, 2001; Houweling et al., 2004]. Miller et al. [2007] estimate that precisions of 1-2 ppm are necessary to monitor carbon fluxes at regional scales.

The first space-based measurements of  $\text{CO}_2$  were performed by NOAA's Television Infrared Observation Satellite Operational Vertical Sounder (TOVS, Smith et al. [1979]), which is able to retrieve seasonality and growth rate with some reliability [Chedin et al., 2002, 2003] but was not originally designed to monitor  $\text{CO}_2$ . Later studies found that TOVS was not ideally suited for surface flux inversion [Chevallier et al., 2005]. Like TOVS, the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY, Bovensmann et al., [1999]) was not designed specifically for  $\text{CO}_2$  retrieval. Unlike TOVS, SCIAMACHY uses short wave near infrared spectrometers to measure  $\text{CO}_2$ . SCIAMACHY is thus more sensitive to surface variations and ultimately found to better constrain surface fluxes given certain levels of precision [Houweling et al., 2004]. Observations from the Atmospheric Infrared Sounder (AIRS, Aumann and Pagano [1994]), which uses a thermal infrared spectrometer, are higher precision and high in number, but are mostly sensitive to the upper troposphere, making flux retrieval difficult.

High-precision with sensitivity down to the surface is possible when taking advantage of reflected sunlight and clear-sky conditions [O'Brien, 1990; Houweling et al., 2004]. Recently, two satellites have been designed specifically to measure the column-averaged dry air mole fraction of  $\text{CO}_2$  ( $X_{\text{CO}_2}$ ) at high (1) precision, (2) surface sensitivity, (3) global coverage and (4) regional scale resolution, all of which are necessary to characterize  $\text{CO}_2$  sources and sinks on regional scales and thus provide the most realistic opportunity to supplement data poor regions: Japan's Greenhouse Gases Observing Satellite (GOSAT, Kuze et al., [2009]) and the National Aeronautics and Space Administrations (NASA) Orbiting Carbon Observatory [OCO, Rayner et al., 2002; Crisp et al., 2004; Houweling et al., 2004; Baker et al., 2006b; Miller et al., 2007; Chevallier et al., 2007]. GOSAT was launched successfully



in January 2009. OCO failed to launch in February 2009; however, the OCO design had certain strong points that have led to a scheduled relaunch (OCO-2) as soon as 2012. In particular, OCO-2 would measure more frequently than GOSAT (180 vs. 13.4 cross-scans per minute) with a smaller field of view ( $\sim 2 \text{ km}^2$  vs.  $\sim 100 \text{ km}^2$ ) and thus ought to find more cloud free scenes [Crisp et al., 2004].

The previous instruments discussed use passive measurements based on shortwave or thermal infrared sensing, with shortwave (SCIAMACHY, GOSAT, and OCO) providing more information than thermal (AIRS) and OCO providing more information than GOSAT [Hungerschofer et al., 2010]. Another instrument being explored that may overcome certain issues with OCO and GOSAT and provide even better surface flux constraint is a new satellite-based laser-sounder. Such an instrument would use Light Detection and Ranging (or LIDAR) technology to measure  $\text{CO}_2$  absorption at nadir in laser pulses sent from the satellite to the surface and back. In this way  $\text{CO}_2$  is measured actively (carries own light source in laser), allowing continuous measurements during day and night, over land and ocean, throughout all latitudes and seasons, with less sensitivity to aerosol and thin clouds yet strong sensitivity to the lower troposphere. NASA’s Active Sensing of  $\text{CO}_2$  Emissions over Nights, Days, and Seasons (ASCENDS, Abshire et al. [2008]) and the European Space Agency (ESA) Advanced Space Carbon and Climate Observation of Planet Earth (A-SCOPE, Ingmann, [2009]) are two promising missions under consideration.

Nevertheless, cloud contamination does represent a problem for remote measurements from space, and using column satellite measurements to represent a transport model grid column introduces temporal, spatial, and local-clear sky sampling errors into inversions. Since satellite measurement require clear-sky conditions, systematic differences in atmospheric  $\text{CO}_2$  concentration between clear and cloudy conditions introduces sampling biases of up to -0.2 to -0.4 ppm into tracer transport inversions that use satellite  $\text{CO}_2$  products to represent temporal averages [Corbin and Denning, 2006]. Spatially coherent biases as small as a few tenths of ppm alters flux estimates by a few tenths of a gigaton of carbon

[Chevallier et al., 2007; Miller et al., 2007]. Spatial representation errors of mixed layer average CO<sub>2</sub> can reach 1-2 ppm for a typical 200-400 km horizontal resolution grid cell, due to unaccounted for horizontal spatial heterogeneity [Gerbig et al., 2003], while column CO<sub>2</sub> spatial representation errors can reach  $\sim 0.6$ - $0.7$  ppm over the land and  $\sim 0.2$ - $0.3$  ppm over the ocean [Lin et al., 2004], with much seasonal and geographical sensitivity [Miller et al., 2007]. Spatial and local clear-sky errors tend to increase when a single satellite track is used to represent a coarse (450 x 450 km) versus fine (100 x 100 km) grid column, with local clear-sky errors biased and large ( $> 2$ ppm) in North America [Corbin et al., 2008].

Spatial and clear-sky sampling errors are attributed partly to differences in NEE on clear and cloudy days as well as surface heterogeneity, but mostly to differential advection by synoptic systems on clear and cloudy days, leading to the largest clear-sky biases in mid-latitude regions [Corbin et al., 2008]. This is not surprising considering CO<sub>2</sub> observations show large day-to-day CO<sub>2</sub> variations associated with passing weather disturbances and horizontal and vertical mixing along fronts [Parazoo et al., 2008]. Since frontal systems create large gradients of CO<sub>2</sub> that are masked by clouds and thus more prone to reductions of sample size, inversions that use satellite measurements to represent coarse regions may incur large and biased spatial and local clear-sky errors. Considering that inversions are influenced by biases as small as a tenth of a ppm in the total column [Chevallier et al., 2007; Miller et al., 2007] and sampling errors grow as large as 2 ppm with a negative bias in coarse grid columns of NH mid-latitudes [Corbin et al., 2008], it is recommended that transport models be run at the finest possible horizontal grid spacing to avoid introducing large errors and bias. Improved grid spacing, however, is likely just one of several factors to account for when considering errors due to transport bias.

### *1.1.2 Transport Models*

Although improvements to the global measurement network reduce uncertainty in inverse flux estimates, accounting for other unknowns such as fossil fuel emissions, wild

fires, deforestation, land use change, interannual climate variability, and other unaccounted for carbon reservoirs and processes in the carbon budget contribute additional uncertainty. Variable inversion setup, including surface flux prior estimates, transport model choice and optimization technique (discussed in Section 1.1.3), has made it difficult to reconcile differences in inverse estimates [Peylin et al., 2002]. The model intercomparison experiment widely known as TransCom3 was designed specifically to isolate the effect of differences in transport using common surface flux “prior” estimates and atmospheric observations. TransCom3 revealed a terrestrial carbon sink distributed evenly amount NH continents. The magnitude of the sink, however, was sensitive to transport differences among 16 different models, and led to the conclusion that simulated transport is one of the greatest sources of uncertainty in inversions [Gurney et al., 2002, 2003, 2004; Baker et al., 2006]. Since transport models are the primary link between atmospheric measurements and surface fluxes, transport uncertainty reduces the ability to fully utilize surface and satellite data [Houweling et al., 2010]. The most common tracer transport modeling framework to date as been global chemistry transport models (CTMs), which use offline stored meteorology fields derived from other atmospheric models and/or data analysis systems.

Factors that can degrade CTM performance include (1) horizontal and vertical grid spacing; (2) parameterization of “subgrid-scale” processes; (3) numerical errors due to finite-differencing algorithms; and (4) data assimilation techniques. An example from (2) above is fine-scale vertical mixing processes such as moist convection and turbulent mixing in the PBL, which typically occur at scales smaller ( $<10$  km) than most global transport models can resolve and thus cannot be solved explicitly. These processes must instead be solved through parameterization - semiempirical statistical theory governing how subgrid-scale processes manifest themselves on the resolved grid - which is highly variable between models, not as easily constrained in reanalysis products as resolvable flows, and therefore highly uncertain.

Covariance between subgrid-scale atmospheric mixing and biological processes is par-

ticularly troubling. It has long been recognized, for example, that seasonal and diurnal covariance between terrestrial ecosystem metabolism and fine-scale vertical transport in the atmosphere is a strong determinant of vertical structure in  $\text{CO}_2$  [“ $\text{CO}_2$  Rectifier,” Denning et al, 1995, 1996, 1999]. Numerical treatment of subgrid-scale vertical mixing continues to be a leading source of uncertainty in  $\text{CO}_2$  inverse models [Denning et al, 1999; Gurney et al, 2003; Yi et al, 2004; Helliker et al, 2004; Baker et al, 2005; Stephens et al, 2007; Yang et al., 2007].

Another process which controls the distribution of  $\text{CO}_2$  on synoptic to seasonal timescales is transport by baroclinic waves along the mid-latitude storm track. Like the PBL-modulated  $\text{CO}_2$  rectifier, synoptic transport of  $\text{CO}_2$  involves strong vertical motion by turbulent mixing and moist convection within frontal cloud bands, and is correlated with ecosystem metabolism because large scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. Baroclinic wave activity is enhanced in winter when respiration and decomposition dominates ecosystem metabolism, and is suppressed in summer when photosynthesis dominates. Unlike the  $\text{CO}_2$  rectifier, transport by baroclinic waves involves a strong meridional component of motion. Tracer transport by baroclinic waves thus follows the path of “slantwise convection” or “slantwise ascent” [e.g., Emanuel, 1988] and is intimately tied to condensation and precipitation processes [Emanuel, 1988; Kuo, 1991].

Before analysis of the role of baroclinic waves in the carbon cycle, background describing the origin of baroclinic waves is provided. Net tropical heating (incoming solar radiation exceeds outgoing longwave radiation (OLR)) and polar cooling (OLR exceeds incoming solar) leads to an energy surplus at low latitudes and deficit at high latitudes (see Figure 1.3). Net heating of the atmosphere in low latitudes and cooling in high latitudes generates meridional temperature gradients and a quantity referred to as zonal available potential energy, or energy available within some distribution of atmospheric mass for conversion to kinetic energy under adiabatic flow. Earth’s rapid rotation inhibits meridional

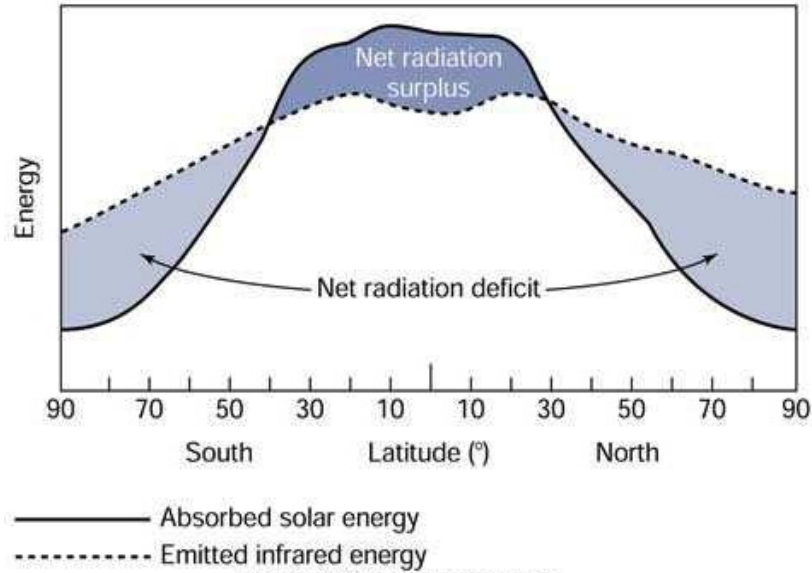


Figure 1.3: Earth's net radiation budget [courtesy of Pearson Prentice Hall, Inc].

flow in the atmosphere in mid-latitudes such that temperature gradients cannot be efficiently weakened by a steady symmetric circulation. The symmetric circulation is unstable, and nature chooses baroclinic waves as a more efficient way of transporting energy poleward, resulting in a weaker meridional temperature gradient.

Baroclinic waves, as discussed by Lorenz [1955] and Peixoto and Oort [1992], result from conversion of zonal available potential energy into eddy available potential energy by growing baroclinic disturbances, which enhance temperature variance around a latitude circle. Eddy available potential energy is then converted into eddy kinetic energy as unstable baroclinic waves cause the growth of quasi-horizontal, quasi-geostrophic eddies of cyclone scale through sinking of colder portions of eddies and rising of the warmer portions. Baroclinic instability is the dominant mechanism for generating eddy transport, through cyclones and anticyclones, in mid-latitudes where vertical wind shear is strongest. Stationary eddies, forced by temperature contrasts due to zonal inhomogeneities along mountains and land-sea boundaries, also transport energy poleward, mostly in the NH.

Baroclinic waves are most common poleward of tropical latitudes where the Coriolis force is large and a westerly thermal wind can be established between warm low latitudes

and cool high latitudes. The embedded cyclones and anticyclones are responsible for most of the synoptic scale weather in NH and SH mid-latitudes [Wallace et al., 1998; Trenberth, 1991], as well as precipitation along surface fronts and moist conveyors [Stewart et al., 1998; Eckhardt et al., 2004]. Mid-latitude cyclones also provide the energy necessary for trace gas transport. Several case studies have reported transport events of chemical trace constituents associated with moving cyclonic systems [Berkowitz et al., 1996; Merrill and Moody, 1996; Moody et al., 1996; Wang and Shallcross, 2000]. Some of these events consist of long range transport ( $\sim 1000$ 's km) between continents within warm sector flow ahead of advancing cold fronts.

Following Carlson [1998], the classic mid-latitude cyclone is roughly 1 km deep, 100's km wide, and composed of three main airstreams: (1) warm conveyor belt (WCB), (2) cold conveyor belt (CCB), and (3) dry intrusion (DI). WCBs and DIs, depicted in Figure 1.4, have strong meridional component of motion necessary for meridional tracer transport. The DI consists of cold dry air that descends behind the cold front, generally equatorward, from the upper troposphere ( $\sim 300$  mb) to the middle to lower troposphere ( $\sim 600$  mb). The



Figure 1.4: Illustration of frontal transports associated with a baroclinic wave.

CCB flows ahead of the warm front and splits into two separate anticyclonic and cyclonic

branches [Schultz, 2001]. WCBs consist of a poleward flow of warm moist air that originates in subtropical latitudes near the surface ( $\sim 800$  mb) in the warm sector ahead of the surface cold front and ascends above the surface warm front into the upper troposphere ( $\sim 300$  mb) at higher latitudes. WCBs are responsible for most of the cyclones meridional energy transport and have the strongest components of vertical motion of the three air streams. Eckhardt et al. [2004] constructed a 15 year climatology of WCBs from reanalysis data and found that WCBs tend to originate between  $25\text{--}45^\circ$  latitude in either hemisphere, occur more frequently during winter than summer, and are almost always found within 1000 km of a cyclone center (quasi-stationary cyclones (i.e. stationary waves) as well as traveling cyclones (i.e., baroclinic or synoptic waves)) and in particular cyclones with strong diabatic heating.

It is not surprising then that WCBs are recognized as the primary mechanism for rapid air pollution transport throughout the NH mid-latitudes [Bethan et al., 1998; Prados et al., 1999; Parrish et al., 2000; Fischer et al., 2002; Miyazaki et al., 2003; Stohl et al., 2003; Cooper et al., 2004]. Poleward transport of pollutants by rising warm air following WCBs is compensated by equatorward transport by sinking of cold dry air along DIs. Transport of chemical trace constituents by DIs almost always has a stratospheric component, and has been shown to advect ozone and dry, thermodynamically stable air from the stratosphere into the upper, middle, and sometime lower troposphere [Parrish et al., 2000; Cooper et al., 2002]. Since WCBs and DIs tend to form in different parts of their world, the histories of their respective air masses are different, which means they almost certainly have different  $\text{CO}_2$  concentrations.

An increasingly important issue for inversion modelers is the ability of CTMs to simulate synoptic variations associated with fronts, WCBs and DIs as they are confronted more with high frequency continuous observations. Patra et al. [2008] analyzed synoptic variations in 25 transport models. Several of the models were actually run at two different horizontal grid spacings with the same primary meteorology, such that grid spacing effects

were isolated from other factors. The authors found that simulations at finer grid spacing produced a better match with observed CO<sub>2</sub> variations at most of the 25 sites in the study, mostly because finer grid models were able to sample closer to the measurement site location. One test indicated that the better match for finer grids was independent of surface flux resolution, although Geels et al. [2007] showed that traditional large scale transport models are not sufficient to resolve fine-scale features associated with fossil fuel emissions. The overall conclusion was that first-order transport mechanisms are fairly well represented in CTMs, but that finer grid spacing should be employed when possible.

For example, a typical synoptic weather pattern has a horizontal scale of approximately 1000 km. A “coarse” grid model run at 250 km resolution will sample a synoptic system 16 times every time step, which is more than enough to capture the large scale behavior. A “fine” grid model run at 50 km resolution will sample the same synoptic system 400 times! This means a fine grid model is 25 times more likely to capture important variations and gradients that might otherwise average out inside of a coarse grid column, while the coarse model will tend to smooth out important spatial gradients. Coarse grid models may simulate overall structure but miss the strength, timing, and location of important variations. The timing issue is of particular relevance for interpreting observed synoptic variations, since gradients of 10 ppm or more may occur at sub-pixel scale [e.g., Chen et al., 2004; Parazoo et al., 2008; Wang et al., 2008].

Figure 1.5 demonstrates the importance of grid spacing for a synoptic system with embedded mesoscale features. Shown is a cold front with a divergence / convergence couplet (purple and red lines) and clouds (depicted by white shading). Overlaid are grid points with 50 km grid spacing. Certain features, such as the convergence zone and pre-frontal clouds, are fairly well resolved. The divergence zone is very weakly resolved. Other features, such as post-frontal clouds and the location of the cold front, are only represented by a couple grid points, and will probably not show up in the model circulation. This demonstrates that overall storm structure is resolved, but smaller features associated with the storm are



unresolved at certain grid spacing.

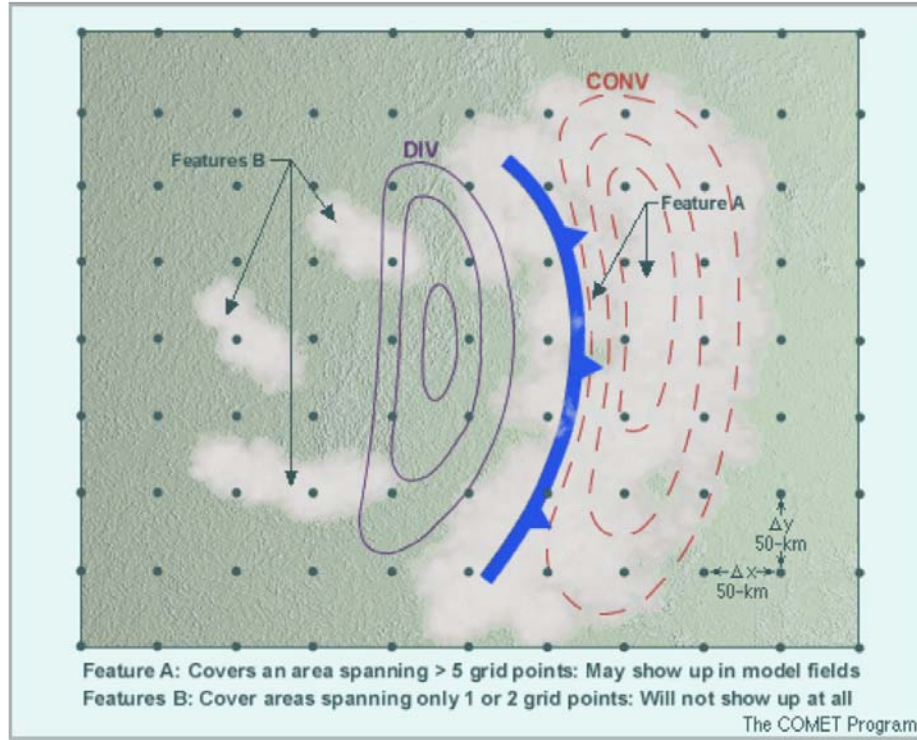


Figure 1.5: Model representation of a cold front and associated divergence / convergence couplet (courtesy The COMET Program, <http://www.meted.ucar.edu/>).

As demonstrated in Figure 1.5, many important processes are not resolved even at what is considered fine grid spacing for global models. While cold fronts may be hundreds of km's long, they are typically less than 100 km wide, so many features are unlikely to show up in global models. Donnell et al. [2001] show, for example, a case study where frontal ascent of tracer is confined to a 60 km wide region. Even at 50 km's such circulations are not resolvable. Fine scale vertical transport by PBL turbulence and moist convection influence frontal lifting and slantwise ascent. Since these processes must be parameterized, they contribute additional uncertainty. Turbulent mixing is essential for boundary layer ventilation of tracer to levels that can be penetrated by conveyor belts [Donnell et al., 2001; Sinclair et al., 2008]. WCBs are heavily influenced by moist processes due to the influx of warm moist air out of the subtropics. Moist processes are found to enhance ascent along WCBs [Sinclair et al., 2008], an unsurprising result given the large amount of latent heat

release associated with condensation. Small inaccuracies in convective representation that specify total latent heat release can lead to large inaccuracies in numerical simulation of dynamical systems where latent heat release is important [Emanuel, 1988]. While it is difficult to test the specific influence of moist convective parameterization on meridional transport of tracer along WCBs without a full blown case study, it is important to consider model differences in meridional transport of moisture, since they may lead to large errors in transport by baroclinic waves.

### 1.1.3 *Optimization Technique*

Spatial and temporal gradients of observed CO<sub>2</sub> mixing ratio are directly related to surface carbon fluxes. In-situ and satellite observations can exploit surface flux information contained in synoptic signals if they are modeled well. Unfortunately, there is no standard way of quantifying transport uncertainty, so even if processes aren't modeled well, inversions must assume perfect transport. Nevertheless, methods are needed to optimally combine models and observations to provide estimates of carbon fluxes and their uncertainties. The scale at which surface fluxes may be optimally inferred therefore depends on the grid spacing of the transport model and prior information, the number and frequency of observations, and the optimization technique, the latter of which has tended to progress along with (or adapt to) the former components.

Early inversions solved for CO<sub>2</sub> fluxes at monthly and annual time scales over continental scale regions using monthly mean in-situ data from networks such as GLOBALVIEW-CO<sub>2</sub> [e.g., Rayner and Law, 1999; Bousquet et al., 2000; Peylin et al., 2002; Gurney et al., 2002; Law et al., 2003]. “Bayesian synthesis” or “batch mode” approaches were used to solve for a single vector of unknowns (spatio-temporal distribution of surface CO<sub>2</sub> flux) that maximized consistency between atmospheric observations and a priori (or “background”) surface fluxes estimated from “bottom-up” approaches. In Bayesian synthesis, a cost function (defined in Section 2.6.3) consisting of two components is optimized. The first term of

the cost function involves the observations and controls the difference between the observed and predicted CO<sub>2</sub> mixing ratios. The second term constrains the solutions by an a priori flux distribution, which is necessary to stabilize the solution when the inversion problem is under-constrained. A linear system of equations, representing the relationship between unknown (CO<sub>2</sub> flux) and known (atmospheric CO<sub>2</sub> observations) quantities, is solved in one step using singular value decomposition to invert large matrices such that fluxes for all source regions are estimated at all times simultaneously using all observations. The TransCom3 experiment is an example of a large-region inversion in which the globe is divided into 22 regions, consisting of 11 land and 11 ocean regions [Gurney et al., 2002]. Such large regions were necessary at the time because of the sparseness of the in situ network. This technique is mathematically overdetermined because the number of unknowns is much less than the available observations.

Large-region inversions work well for solving for a set of relatively few unknowns because they are computationally efficient, but likely lead to aggregation errors [Kaminski et al., 2001; Engelen et al., 2002] because atmospheric CO<sub>2</sub> is sensitive to the distribution of sources and sinks within large basis regions. New approaches were needed to supplement large-region estimates to solve for more unknowns with the same number of observations to improve the detail of fluxes. Grid scale inversions are an example of such an approach. Large-region inversions like TransCom3 are essentially grid point inversions, except that grid points are perfectly correlated in space in a region and the number of regions is order of magnitudes smaller than the number of grid points. In “pixel-scale” grid scale inversions, if grid boxes are assumed to be uncorrelated, the problem becomes mathematically under-determined because the number of unknowns exceeds the available observations, but can be solved by the priors. Since it is not desirable for inversions to be too heavily weighted by priors, grid-scale inversions must compromise by correlating grid cells.

Grid scale inversions have become a very popular technique for estimating CO<sub>2</sub> fluxes [Kaminski et al., 1999; Houweling et al., 1999; Rödenbeck et al., 2003; Peylin et al., 2005].

Rödenbeck et al. [2003] performed a grid scale inversion at  $10^\circ$  longitude by  $8^\circ$  latitude by accounting for unique spatial covariance of flux uncertainties over land and ocean. Geostatistical techniques use Bayesian methods that replace prior flux estimates with prespecified covariance structure from vegetation cover, leaf area index, greenness fraction, etc., that varies with the mean behavior of the fluxes [Michalak et al., 2004]. This technique acts as a good independent estimate to bottom-up estimates. Peylin et al [2005] used continuous measurements to estimate daily  $\text{CO}_2$  fluxes at the model grid scale over Europe.

Bayesian techniques worked well with the relatively sparse observation network. Each new year brings with it more surface measurements at hourly resolution and huge arrays of satellites measurements at sub-hourly resolution, which together increases the observation vector beyond the capacity of bayesian techniques. Given that enhanced observation density will improve optimization at finer scales, alternative assimilation techniques to batch inversions are needed. Bruhwiler et al. [2005] introduced a fixed-lag Kalman smoother, used extensively in Numerical Weather Prediction, which steps through observations sequentially in time to optimize a subset of observations and update forecast error covariance. This method assumes that information contained in measurements smoothes out over time due to atmospheric mixing of spatial gradients such that fewer observations (6-9 months worth) need be used compared to large Bayesian systems (1 year or longer). Although the Kalman smoother is not ideally suited for continuous observations, it was extremely innovative for inverse studies of  $\text{CO}_2$ .

Peters et al [2005] expanded on the fixed lag Kalman smoother by introducing an ensemble based approach in which the covariance matrix was approximated by ensemble members. An advantage of ensemble approaches is that they incorporate uncertainty in estimates of flux distributions that arise from errors in the observations and transport model. These approaches attempt to estimate the probability distribution functions (PDFs) of the analysis and forecasts using a relatively small ensemble of random realizations from the distributions, and use the ensemble estimate of a short term forecast PDF when assimilating

ing observations. Ensemble Kalman filters such as that proposed by Peters et al. [2005] use ensemble estimates of error covariance such that surface flux estimates are those that best fit the assimilated observations. The so called “fixed-lag ensemble Kalman smoother” replaces bottom-up flux estimates with background fluxes that result from the assimilation process and therefore contain information from previous analysis cycles. While this technique can ingest large amounts of observations without the need to precalculate observation operators, it is still too computationally expensive for large observation vectors that occur in the satellite era. Another ensemble-based data assimilation approach involves the Maximum Likelihood Ensemble Filter, or MLEF [Zupanski, 2005; Zupanski and Zupanski, 2006; Zupanski et al., 2006; Zupanski et al., 2007; Lokupitiya et al., 2008], which is designed to be able to handle large observational datasets.

Baker et al. [2006] and Chevallier et al [2005] introduced variational data assimilation schemes as an alternative to pure matrix-based schemes in order to cope with the large amount of satellite data. Four-dimensional variational data assimilation uses an adjoint version of the transport model to estimate that atmospheric state that best fits assimilated observations. Calculation of the model adjoint can be a bit disconcerting, especially if frequent improvements are introduced to models and maintenance of the adjoint is necessary. Because the “atmospheric memory” can be quite long, variational approaches are desirable in that it is possible to optimize fluxes using a year or more’s worth of observations that contain information at multiple mixing time scales.

## 1.2 Objectives of this Study

A major science objective of most carbon cycle studies is quantification of carbon sources and sinks at the finest possible resolution with the highest possible accuracy. CO<sub>2</sub> inversion studies attempt to map and quantify source and sink distributions using atmospheric transport models to communicate observed signals to the surface through increasingly sophisticated optimization techniques. The literature review has examined each of

the three major components of this top-down approach, and revealed several well known weaknesses. While the other components are important, the transport model is the primary tool by which the inversion determines at what geographic location a given parcel of air containing some mixing ratio of  $\text{CO}_2$  originates. Given that the distribution of  $\text{CO}_2$  mixing ratios is ultimately determined by winds between the surface flux and the observation, it is crucial to get the winds correct. Inability of atmospheric models to properly represent important dynamical processes associated with weather systems is a leading cause of incorrect winds and hence uncertain source and sink maps. This study is motivated by (1) getting the winds correct and (2) quantifying the effect of incorrect winds on inverse estimates of  $\text{CO}_2$  flux from satellite data.

To account for all winds in the atmosphere would be well beyond the scope of any single study. This research therefore focuses on a certain set of winds that have received relatively little attention in the carbon science literature: winds from extratropical cyclones. Extratropical cyclones involve processes such as frontal ascent, cumulus convection, and turbulence mixing in the PBL that can transport tracer rapidly and over long distances but occur at scales that are unresolved and therefore poorly represented in today's weather prediction, climate, and inversion models. Extratropical cyclones are a crucial component of the moist circulation and are responsible for the bulk of weather in mid-latitudes. They are also responsible for a great deal of variance in  $\text{CO}_2$  at continental observing stations in northern middle latitudes. Such highly variable  $\text{CO}_2$  distributions are actually signals that contain important upstream surface flux information at synoptic and seasonal scales, so it is important to represent them well. Unfortunately, these weather systems and resulting  $\text{CO}_2$  variations are difficult to represent numerically, and as it turns out are difficult to observe. To complicate matters further, extratropical cyclones represent only a single component of a system of processes including baroclinic waves, extratropical cyclones, and frontal weather systems. This system of processes will be referred to from this point on as synoptic processes. Seasonal covariance between synoptic processes and terrestrial ecosystem metabolism, and

resulting meridional CO<sub>2</sub> transport within the mid-latitude storm track, has received very little attention in the literature. This research seeks a comprehensive description of CO<sub>2</sub> transport by synoptic processes and an understanding of possible implications of synoptic transport (and underlying uncertainties) for surface flux estimates and carbon science. This study consists of three parts:

- (1) Part 1 seeks a dynamical framework for explaining synoptic transport processes for CO<sub>2</sub>. In particular, Part 1 explores how the moist component of the mid-latitude atmospheric circulation acts as a vehicle for transport of CO<sub>2</sub> between middle and polar latitudes. This part of the study is motivated by recent findings [Pauluis et al., 2008] that mass transport along moist isentropic surfaces on baroclinic waves represents an important component of the atmospheric heat engine that operates between the equator and poles, and that moist transport accounts for as much as half the air in polar regions. Additionally, moist synoptic transport along the mid-latitude storm track is correlated with ecosystem metabolism because large-scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. This seasonal covariance between meridional heat transport and surface fluxes of CO<sub>2</sub> should therefore have interesting implications for the seasonal cycle of CO<sub>2</sub> in northern latitudes. Results from Part 1 are discussed in Chapter 3.
- (2) Part 2 seeks to quantify uncertainty associated with transport by synoptic waves. As discussed in Section 1.1.2, there are many inherent uncertainties associated with synoptic transport, including moist processes within warm conveyors, frontal circulations, and fine-scale vertical transport by moist convection and PBL turbulence. These uncertainties are linked in part to grid spacing and parameterization of sub-grid scale vertical mixing. These uncertainties are investigated using a self-consistent modeling framework driven by reanalysis products that are similar numerically, but have important differences in grid spacing and vertical mixing schemes. Many carbon cycle studies conclude that it is important to improve the

representation of important transport processes within tracer transport models. Part 2 of this study seeks to identify and explain key differences in modeled synoptic transport. Results from Part 2 are discussed in Chapter 4.

- (3) Part 3 seeks to quantify errors in estimates of surface fluxes of  $\text{CO}_2$  that result from transport biases. Synoptic transport of  $\text{CO}_2$  occurs on scales that are poorly represented in global models, so it may be a source of inversion errors. Observing System Simulation Experiments (OSSEs) are designed to quantitatively assess the value of atmospheric observing systems to numerical forecasting. They have been applied by the  $\text{CO}_2$  inversion community in a similar manner to assess assimilation techniques and the potential of atmospheric  $\text{CO}_2$  observations to provide constraint for surface flux estimation. Part 3 uses an OSSE framework to test (1) the ability of an ensemble technique for flux estimation to retrieve realistic looking sinks hidden within satellite data and (2) whether sink estimates are substantially degraded in the presence of imperfect transport. Results from Part 3 are discussed in Chapter 5.

Each results chapter also includes an introduction, motivation, and conclusions section. Conclusions from all parts of this research are tied together in Chapter 6. Transport simulations are performed with the Parameterized Chemistry and Transport Model (PCTM). Chapter 2 includes description of PCTM (Chapter 2.1), along with prescribed ocean and terrestrial surface fluxes (Chapter 2.2), various meteorological driver data sets (Chapter 2.3), and forward simulation experiments (Chapter 2.4). Parts 1 and 2 of this study require special analysis techniques to characterize moist synoptic transport at global scales, which are described in Chapter 2.5. The inversion methodology used for Part 3 is described in Chapter 2.6.



## Chapter 2

### EXPERIMENTAL METHODS

#### 2.1 Model Description

##### 2.1.1 *PCTM*

The chemistry transport model used for CO<sub>2</sub> forward simulations, PCTM, has been adapted from a full-chemistry and transport model [e.g., Douglass and Kawa, 1999], technical aspects of which are discussed in Nielson [2000]. Three-dimensional tracer advection in PCTM is based on the transport code of Lin and Rood [1996], which uses a flux form semi-Lagrangian formulation (FFSL). Several modifications were made to PCTM for mass conservation [Kawa et al., 2004]. Transport in PCTM can be driven by simulated or analyzed weather fields.

All transport is calculated on the hybrid sigma-pressure ( $\sigma - p$ ) terrain following vertical coordinate, which can be defined according to Simmons and Burridge [1981] by  $\eta = \frac{p}{p_s} + \left(\frac{p}{p_s} - 1\right) \left(\frac{p}{p_s} - \frac{p}{p_o}\right)$ , where  $p$  is pressure,  $p_s$  is surface pressure, and  $p_o$  is the reference pressure equal to 1000.0 mb. This equation forces  $\eta$  to be monotonic with height, where  $\eta = 1$  when  $p = p_s$  and  $\eta \rightarrow 0$  as  $p \rightarrow 0$ , and is designed such that  $\eta$  is terrain following near the surface and tends uniformly to pressure at upper levels. Simmons and Strüfing [1983] approximate  $\eta$  by specifying the pressure of each model level as a function of surface pressure and vertically dependent coefficients  $a$  and  $b$ . The  $\eta$  coordinate can be generalized such that pressure is given by

$$p(i, j, k, t) = a(k) + b(k) p_s(i, j, t) \quad (2.1)$$

where for a given time  $t$ ,  $p$  is the pressure at the coordinates  $i, j, k$  and  $a(k)$  and  $b(k)$  are components of the hybrid coordinate at the layer edge  $k$ .  $p_s(i, j, k)$  is the surface pressure at the horizontal grid point  $(i, j)$  and time  $t$ .  $a(k)$  and  $b(k)$  at the layer edges are stored as global metadata values in each  $\eta$  product file. This study uses transport fields saved on  $\eta$  from four versions of the NASA Goddard Earth Observation System (GEOS, discussed below). Surface  $\text{CO}_2$  boundary conditions include terrestrial, oceanic, and anthropogenic fluxes, also discussed below.

PCTM solves the constituent continuity equation

$$\frac{\partial \chi}{\partial t} = -\tilde{\mathbf{V}} \cdot \nabla \chi + \text{L}(\chi) + \text{P}(\chi) \quad (2.2)$$

where  $\chi$  is the constituent mixing ratio concentration,  $t$  is time,  $\tilde{\mathbf{V}}$  is the total velocity vector, and  $\text{P}$  and  $\text{L}$  are the production and loss rates, respectively using a technique known as process splitting [Nielson, 2000]. The Lin and Rood scheme offers several major advantages of FFSL necessary to maintain the statistics of advected tracers. One includes independence of stability on time step [Lin and Rood, 1996]. In addition, the scheme meets the physical constraints of tracer advection and accounts for the problem of consistency between the tracer continuity equation and the underlying equation of continuity of air due to process splitting techniques [Lin and Rood, 1996]. The accuracy of the code for large-scale transport is well documented [Kawa et al., 2004].

PCTM simulations that use analyzed meteorology face the problem of local non-conservation of mass that arises during the assimilation process, which leads to inconsistencies between surface pressure tendencies and mass flux convergence. This, in turn, introduces errors in the advected tracer field. Kawa et al. [2004] add a pressure fixer [Rotman et al., 2001] to the model, which acts to remove zonally distributed pressure errors without inducing a vertical wind change, and find that inconsistencies are removed.

Subgrid-scale vertical mixing processes such as cumulus convection and PBL turbulence are also important for atmospheric tracer transport. Convective mixing is implemented in PCTM using a mass conserving, semi-implicit convective transport module [Kawa et al.,

2004], formulated to be consistent with convective mass fluxes provided by the Zhang and McFarlane [1995] deep convection scheme used in the finite volume GCM (FVGCM). Mixing by PBL turbulence is implemented using diffusion between adjacent layers.

## 2.2 Surface Boundary Conditions

This section briefly discusses surface flux data, which act as sources and sinks of  $\text{CO}_2$  to the atmosphere. Surface fluxes used in this study include terrestrial biological fluxes, anthropogenic sources, and oceanic fluxes. Fluxes due to land use change and fire emissions are not included.

### 2.2.1 *Terrestrial Fluxes*

The Simple Biosphere model (SiB) was developed by Sellers et al. [1986] to calculate surface energy budgets in climate models and after substantial modification [Sellers et al., 1996 a,b] became SiB2. Vegetation parameters are derived directly from processed satellite data. The parameterization of stomatal and canopy conductance used in the calculation of the surface terrestrial energy budget involves the direct calculation of carbon assimilation by photosynthesis [Farquhar et al., 1980], making possible the calculation of  $\text{CO}_2$  exchange between the land and atmosphere [Denning et al., 1996, 2003]. Photosynthetic carbon assimilation is linked to stomatal conductance and thence to the surface energy budget and atmospheric climate by the Ball-Berry equation [Collatz et al., 1991, 1992]. SiB models include a prognostic canopy airspace of temperature, moisture, and  $\text{CO}_2$ . Time invariant biophysical parameters include canopy height, leaf angle distribution, leaf transmittance, photosynthetic parameters, and soil properties. Time and spatially varying biophysical parameters are calculated from satellite observations (see below) and include leaf area index (LAI) and fraction of absorbed photosynthetic active radiation (fPAR).

Additional modifications have occurred since SiB2. For example, the ability to accumulate up to five layers of snow, each with unique thermodynamic properties, has been

added, improving the treatment of soil insulation and thermal properties in the winter. A more realistic root profile is used, with 10 soil layers and an initial soil column depth of 3.5 meters, along with better treatment of soil water stress and frozen soil. Respiration includes an autotrophic component, accounting for maintenance and growth. The revised model is referred to as SiB3 [Baker et al., 2008].

Soil respiration is modeled as a function of temperature and moisture of each layer of soil, and is scaled with annual gross primary production (GPP) to achieve carbon balance over an annual cycle [Denning et al., 1996a]. Running SiB3 in this steady state mode, such that ecosystem respiration balance GPP over one year at every grid point, is necessary because processes governing long term sources and sinks of CO<sub>2</sub> such as forest regrowth and soil carbon pools are not simulated in SiB3. Processes that govern biospheric exchange over diurnal, synoptic, and seasonal time scales, which are controlled primarily by climate and weather, are modeled very well by SiB3. For example, SiB3 has been evaluated against eddy covariance measurements at a number of sites [Baker et al., 2003; Hanan et al., 2005; Vidale and Stockli, 2005].

Net ecosystem exchange (NEE, representing the net terrestrial biogenic flux of CO<sub>2</sub> and calculated as Respiration - GPP) estimates from SiB3 are provided at hourly resolution. In this experiment, initial conditions for soil moisture and temperature are specified using an 18-year global spinup of SiB3 (January 1, 1982 to December 31, 1999) driven by the National Center for Environmental Prediction Department of Energy reanalysis (NCEP2) 1° x 1° (longitude x latitude) global dataset [Kanamitsu et al., 2002]. A different reanalysis product is used to drive SiB3 for the PCTM simulation period (discussed below).

### *Land Surface Parameters*

Vegetation greenness is prescribed using LAI and fPAR estimates from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the NASA EOS satellite TERRA. LAI and fPAR products (MOD15A2) are produced as 8-day composites at 1-km

resolution; this study uses a version regridded to  $0.25^\circ \times 0.25^\circ$  and corrected for cloud and aerosol contamination produced jointly by Boston University and the University of Montana (Myneni et al. [2003]; data courtesy <http://cybele.bu.edu>). The use of satellite derived LAI and fPAR parameters is necessary for calculation of surface photosynthesis, evapotranspiration, and annual net primary production, and provides adequate resolution in time and space for studies of terrestrial carbon processes at synoptic and seasonal scales.

The soil map for SiB3 is provided by the International Geosphere-Biosphere Programme (IGBP) at a resolution of 10 km and then modified to correspond to SiB3 classes. The biome map is a satellite data product of the University of Maryland with a resolution of 1 km, converted to  $1^\circ$  by  $1^\circ$  resolution (see Figure 2.1). Biome types are also converted to SiB3 classes.

SiB3 combines information about MODIS, biome, and soil type to determine surface characteristics, which are used in  $\text{CO}_2$  flux calculations.

### 2.2.2 *Fossil Fuel Fluxes*

Fossil fuel fluxes are prescribed from anthropogenic  $\text{CO}_2$  emission ( $\text{kg C m}^{-2} \text{ s}^{-1}$ ) estimates based on calculations done by Andres et al. [1996] at  $1^\circ$  by  $1^\circ$  resolution for the entire globe using fossil fuel consumption, cement manufacture, and population density data from 1998. Fossil fuel fluxes represent a source of  $\text{CO}_2$  from the land to the atmosphere. Although fossil fuel fluxes are spatially variable, they are constant in time, and therefore do not accurately portray diurnally, weekly, and seasonally variability [Gurney et al., 2005] or the correct atmospheric growth rates. A realistic fossil map with the correct temporal and spatial variability is not needed for this study.

### 2.2.3 *Oceanic Fluxes*

The ocean source is from the Community Climate System Model (CCSM-3) coupled ocean Biogeochemical Elemental Cycling model (BEC). This is a process based physical-

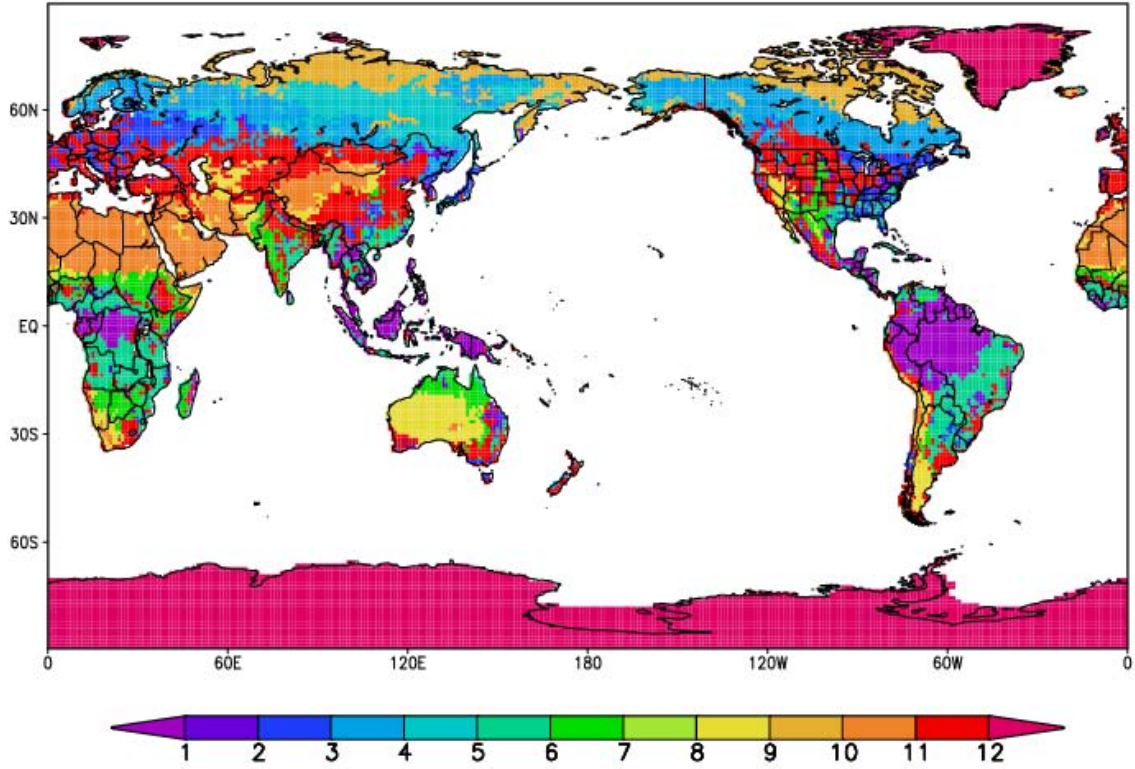


Figure 2.1: SiB3 Biome Map for GiMMSg NDVI. Vegetation Type by Number on Color Bar: 1) C3 Tall Broadleaf-Evergreen Trees, 2) C3 Tall Broadleaf-Deciduous Trees, 3) C3 Tall Broadleaf and Needleleaf Trees, 4) C3 Tall Needleleaf Trees, 5) C3 Tall Needleleaf-Deciduous Trees, 6) C4 Short Vegetation, Same as Types 6, 7, 8, 11 7) C4 Short Vegetation (Maize Optical Properties), 8) Same as 7, 9) Short Broadleaf Shrubs with Bare Soil, 10) C3 Short Ground Cover (Tundra), 11) C4 No Vegetation (Low Lat Desert), 12) Agriculture (Wheat) and C3 Grasslands, 13) Ice

biogeochemical-ecological model driven by reanalysis, capable of producing interannual variability in air-sea  $\text{CO}_2$  flux. The BEC model consists of upper ocean ecological [Moore et al., 2004] and full-depth biogeochemical [Doney et al., 2006] modules imbedded in a global 3-D Parallel Ocean Program (POP) ocean general circulation model [Smith and Gent, 2004; Collins et al., 2006]. The model is forced with atmospheric reanalysis from the CORE Inter-Annual Forcing (CIAF) version 2 dataset [Large and Yeager, 2004] and time-varying dust deposition [Mahowald et al., 2003] and saved at daily resolution. These models are described and analyzed in more detail in Doney et al [2009a; 2009b].

## 2.3 Transport Driver Data

Transport in PCTM is driven in offline mode; that is, transport fields are produced from a parent GCM, which are then saved for later transport in a chemistry transport model (CTM) such as PCTM. Parent GCMs used in this study are developed, tested, and run at NASA Goddard Global Modeling and Assimilation Office (GMAO). Weather fields in these GCMs are analyzed (or re-analyzed) through assimilation of data from conventional (radiosondes, wind profiles, radar, etc.) and satellite streams. Transport is driven by resolved processes, including horizontal winds and vertical divergence, which are used for advective processes, and unresolved processes solved through parameterization, including cumulus convection and PBL turbulence, which are used for vertical mixing. Transport in this study is determined from analyzed and re-analyzed meteorological fields from the GEOS Data Assimilation System (DAS). Goddard datasets are produced from two primary parent atmospheric general circulation models (AGCMs), GEOS version 4 (GEOS4) and GEOS version 5 (GEOS5), at three different spatial resolutions and two different vertical resolutions. Each parent AGCM has a unique set of vertical mixing schemes to generate moist convective mass fluxes and vertical mixing. The key differences between these models are specified below.

### 2.3.1 *GEOS4-1.25x1*

The first product is based on version 4 of GEOS-DAS (GEOS4-DAS). Kawa et al. [2004], Douglass et al. [2003], and Bloom et al. [2005] discuss this data in detail. Like PCTM, GEOS4-DAS uses finite volume dynamics based on the Lin-Rood dynamical core [Lin and Rood, 1996]. Observations used in the finite volume data assimilation system (FV-DAS) include both meteorological products and satellite data. Physical parameterizations are determined using the National Center for Atmospheric Research Community Climate Model, Version 3 (CCM3) package [Kiehl et al., 1998], which include deep convection [Zhang and McFarlane, 1995], shallow convection [Hack, 1994], and PBL turbulence [Holtslag and

Boville, 1993]. GEOS4-DAS transport fields are saved at 6-hourly resolution and  $1.25^\circ$  by  $1^\circ$  (representing the native resolution of GEOS4-DAS) grid spacing with 55 hybrid vertical levels up to 1 mb. This re-analysis driver data will be referred to as GEOS4-1.25x1.

Of the 55 vertical levels, only 14 are located within the troposphere. Since the mixing time scale between the troposphere and stratosphere is approximately 10 years, and this study is focused on processes in the troposphere at time scales much less than 10 years, the 55 vertical levels in GEOS4-1.25x1 are condensed to 25 levels while retaining all levels in the troposphere.  $1.25^\circ$  by  $1^\circ$  transport in PCTM is run with a 7.5 minute time step.

### *2.3.2 GEOS4-2.5x2*

GEOS4-1.25x1 is regridded horizontally to  $2.5^\circ$  by  $2^\circ$  to study transport at coarser horizontal resolution. Vertical mixing by moist convection and turbulent diffusion is identical to GEOS4-1.25x1. Wind vectors are also conserved, but regridding to coarser resolution has smoothed spatial gradients. All 25 vertical levels are retained, and transport fields are still saved at 6-hourly resolution. Time-stepping through PCTM is doubled to 15 minutes to be consistent with CFL and doubling of horizontal grid spacing. This regridded re-analysis driver data will be referred to as GEOS4-2.5x2.

### *2.3.3 GEOS5-0.67x0.5*

The third product is based on version 5 of GEOS-DAS [GEOS5-DAS, Rienecker et al., 2008]. GEOS5-DAS maintains the finite-volume dynamics used from GEOS4-DAS [Lin, 2004] and is integrated with physics packages under the Earth System Modeling framework (e.g., Collins et al., 2005), including the Relaxed Arakawa-Schubert (RAS) scheme for convection [Moorthi and Suarez, 1992] and separate PBL turbulent mixing schemes for stable [Louis et al., 1982] and unstable [Lock et al., 2000] conditions. DAS integrates GEOS5-DAS with three-dimensional variational analysis (3DVar) referred to as Gridpoint Statistical Interpolation (GSI) analysis, with a focus on wind fields for studies of tropospheric transport.



This GEOS5-DAS product is technically based on version 5.1.0 of GEOS5-DAS, which was used for tuning and development issues for the Modern Era Retrospective-analysis for Research Applications (MERRA, e.g., Bosilovich et al. [2006]). The native grid of GEOS5-DAS (and the grid at which the analysis is performed) is  $0.67^\circ$  by  $0.5^\circ$  horizontal resolution with 72 layers to 0.01 hPa, 31 of which are in the troposphere and the rest above the tropopause. The 72 vertical layers are condensed to 42 levels while retaining the 31 tropospheric levels. Transport fields are saved at 6-hourly resolution. In order to maintain CFL criteria,  $0.67^\circ$  by  $0.5^\circ$  transport is run with a 3.75 minute time step in PCTM. Transport fields are as instantaneous analysis (in contrast to the re-analysis version of GEOS5-DAS discussed below). Since the analysis is based on GEOS5-DAS, this driver data will be referred to as GEOS5-0.67x0.5.

#### *2.3.4 GEOS5-1.25x1*

The fourth (and final) product is similar to GEOS5-0.67x0.5 except is a newer version (5.2.0) of GEOS5-DAS used for MERRA which addresses a deficiency in the diurnal cycle in the high latitudes. Except for changes to address high latitude diurnal cycles, GEOS5-DAS version 5.2.0 is nearly the same as GEOS5-0.67x0.5. Transport fields are re-analyzed (saved during a corrector segment of the Incremental Analysis Update rather than the analysis segment) and saved at 3-hourly resolution at a reduced horizontal resolution of  $1.25^\circ$  by  $1^\circ$  to facilitate comparison with GEOS4-1.25x1. This re-analyzed driver data will be referred to as GEOS5-1.25x1.

## **2.4 Transport Simulations for Parts 1 and 2**

Transport simulations are run from January 1, 2000 through December 31, 2004 to properly spin up atmospheric gradients of  $\text{CO}_2$ . Transport in the spin up run is driven by GEOS4-1.25x1. All terrestrial, oceanic, and anthropogenic surface fluxes of carbon are used in the spin up. Flux and energy calculations in SiB are also driven by GEOS4-1.25x1 me-

teorology from 2000-2004 such that transport and surface fluxes of  $\text{CO}_2$  are synchronized. GEOS4-1.25x1 precipitation is scaled by monthly precipitation from the Global Precipitation Climatology Project (GPCP, see Huffman et al. [2001]) to force total monthly precipitation in GEOS4-1.25x1 to match that of GPCP. The time at which precipitation occurs remains unchanged so that covariance of anomalies in cloudiness, moisture, and vertical transport is conserved. Anthropogenic and oceanic surface flux maps are regridded to  $1.25^\circ$  by  $1^\circ$  using the conservative remapping scheme of the SCRIP software package [Jones, 1999].  $\text{CO}_2$ , which is given as a mole fraction in units of parts per million (ppm), is treated as a passive tracer such that chemical reactions and feedbacks between  $\text{CO}_2$  and weather are non-existent.

Three-dimensional  $\text{CO}_2$  fields from PCTM are saved at 0z on January 1, 2005 and regridded to the appropriate horizontal and vertical resolution depending on the driver data. PCTM is run from January 1, 2005 through December 31, 2005 to comprise the analysis period. Part 1 of this study includes one forward simulation in which PCTM is driven by GEOS4-1.25x1. Part 2 of this study consists of four forward simulations, one for each of the meteorological driver data sets. All driver data is run through the same version of PCTM. All transport simulations are prescribed with the same set of surface fluxes. SiB3 is run at  $1.25^\circ$  by  $1^\circ$  driven by two-dimensional surface meteorology from GEOS4-1.25x1. Ocean, fossil fuel, and terrestrial fluxes are regridded to the resolution of the meteorological driver data. For Part 2, this experimental design is important because all simulations use a common transport model and set of surface fluxes such that the only difference between experiments is atmospheric transport. Any differences in  $\text{CO}_2$  transport in Part 2 can therefore be attributed to differences in meteorological driver data.

With regard to Part 2, it is important to note that, while the dynamical core is consistent between GEOS models and with the dynamical core of PCTM, there are many other factors that may cause transport differences. Inconsistencies between horizontal and vertical grid spacing, data assimilation technique, mass balance errors associated with data

assimilation, number and type of observations used in data assimilation, physical parameterization schemes for vertical mixing are probably the most dominant factors. That being said, GEOS4-2.5x2 and GEOS4-1.25x1 are different horizontal grid spacing. The same can be said for GEOS5-1.25x1 and GEOS5-0.67x0.5, also some minor differences in the PBL scheme in the Arctic and the number and type of data assimilated may also be important (more discussion in Chapter 4). The biggest model differences in synoptic transport of CO<sub>2</sub> will most likely arise between GEOS4-DAS and GEOS5-DAS as a result of some combination of all of the factors listed above.

## 2.5 Methodology for Transport Analysis

PCTM solves the constituent continuity equation on  $\eta$ , but tracer transport can be diagnosed from the model output on any vertical coordinate (height, pressure, isentropes, etc). It is important to note that regridding between vertical coordinates will not change CO<sub>2</sub> distributions throughout the atmosphere, nor total meridional transport around a latitude circle. It will, however, change our interpretation of important physical processes in the atmosphere that are important for CO<sub>2</sub> variations and meridional transport. Because the underlying mechanisms governing CO<sub>2</sub> transport are open to interpretation, analysis of meridional transport in Part 1 will be performed on three vertical coordinate systems (discussed below) represent Eulerian and Lagrangians frames of reference. This section begins by expressing tracer transport as a function of a generalized vertical coordinate. The eddy decomposition technique for analyzing large scale tracer transport by symmetric and asymmetric components of the circulation is then described. Transport equations are then extended to three vertical coordinates.

### 2.5.1 Tendency Equation for Atmospheric $CO_2$

The equation expressing conservation of an arbitrary intensive scalar,  $\psi$ , on an arbitrary vertical coordinate,  $\zeta$ , can be written as

$$\begin{aligned} \frac{\partial}{\partial t} m_\zeta \psi_\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \lambda} m_\zeta u_\zeta \psi_\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} m_\zeta v_\zeta \psi_\zeta \cos \varphi \\ + \frac{\partial}{\partial \zeta} m_\zeta \dot{\zeta} \psi_\zeta = \frac{\partial}{\partial \zeta} S_\psi, \end{aligned} \quad (2.3)$$

where  $m_\zeta = -\frac{1}{g} \frac{\partial p}{\partial \zeta}$  is the pseudo-density (equivalent to the amount of mass (as measured by the pressure difference) between two  $\zeta$  surfaces),  $p$  is pressure,  $g$  is gravity,  $a$  is the radius of Earth,  $\varphi$  is longitude,  $\lambda$  is latitude,  $u$  is the zonal wind,  $v$  is the meridional wind, and  $S_\psi$  is the vertical flux of  $\psi$ .  $\dot{\zeta} \equiv \frac{D\zeta}{Dt}$  is defined as the rate of change of  $\zeta$  following a particle. The horizontal advection terms,  $\frac{1}{a \cos \varphi} \frac{\partial}{\partial \lambda} m_\zeta u_\zeta \psi_\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} m_\zeta v_\zeta \psi_\zeta \cos \varphi$ , can be rewritten as  $\nabla_\zeta \cdot m_\zeta \mathbf{V}_\zeta \psi_\zeta$ , where  $\mathbf{V}_\zeta$  is the horizontal wind vector pointing along a surface of  $\zeta$ . Equation 2.3 then reduces to

$$\frac{\partial}{\partial t} m_\zeta \psi_\zeta + \nabla_\zeta \cdot m_\zeta \mathbf{V}_\zeta \psi_\zeta + \frac{\partial}{\partial \zeta} m_\zeta \dot{\zeta} \psi_\zeta = \frac{\partial}{\partial \zeta} S_\psi. \quad (2.4)$$

To solve for conservation of  $\psi$  through an atmospheric column, Equation 2.4 is integrated vertically from the surface to the top of the atmosphere. This gives

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\zeta_S}^{\zeta_T} m_\zeta \psi_\zeta d\zeta + \nabla_\zeta \cdot \int_{\zeta_S}^{\zeta_T} m_\zeta \mathbf{V}_\zeta \psi_\zeta d\zeta \\ + (m\psi)_{\zeta_S} \left( \frac{\partial \zeta_S}{\partial t} + \mathbf{V}_{\zeta_S} \cdot \nabla \zeta_S - \dot{\zeta}_S \right) \\ - (m\psi)_{\zeta_T} \left( \frac{\partial \zeta_T}{\partial t} + \mathbf{V}_{\zeta_T} \cdot \nabla \zeta_T - \dot{\zeta}_T \right) \\ = S_{\psi_S} - S_{\psi_T}, \end{aligned} \quad (2.5)$$

where  $T$  refers to the top of the model atmosphere,  $S$  refers to the bottom of the atmosphere (or surface),  $\zeta_T$  is the value of  $\zeta$  at the top of the model atmosphere, and  $\zeta_S$  is the value of  $\zeta$  at the bottom of the model atmosphere. The lower boundary condition, i.e., that no mass crosses the Earth's surface, is expressed by requiring that a particle that is on the Earth's

surface remains there:

$$\frac{\partial \zeta_S}{\partial t} + \mathbf{V}_S \cdot \nabla \zeta_S - \dot{\zeta}_S = 0. \quad (2.6)$$

Similarly,  $\frac{\partial \zeta_T}{\partial t} + \mathbf{V}_T \cdot \nabla \zeta_T - \dot{\zeta}_T$  represents the mass flux across the top of the atmosphere, which we assume to be zero, i.e.,

$$\frac{\partial \zeta_T}{\partial t} + \mathbf{V}_T \cdot \nabla \zeta_T - \dot{\zeta}_T = 0. \quad (2.7)$$

Assuming also that there are no sources or sinks of  $\psi$  at the top of the atmosphere, and substituting Equations 2.6 and 2.7 into Equation 2.5, Equation 2.5 reduces to:

$$\frac{\partial}{\partial t} \int_{\zeta_S}^{\zeta_T} m_\zeta \psi_\zeta d\zeta + \nabla_\zeta \cdot \int_{\zeta_S}^{\zeta_T} m_\zeta \mathbf{V}_\zeta \psi_\zeta d\zeta = S_{\psi_S} \quad (2.8)$$

or, expanding the second term on the right hand side (RHS) of Equation 2.8,

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\zeta_S}^{\zeta_T} m_\zeta \psi_\zeta d\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \lambda} \int_{\zeta_S}^{\zeta_T} m_\zeta u_\zeta \psi_\zeta d\zeta \\ + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \int_{\zeta_S}^{\zeta_T} m_\zeta v_\zeta \psi_\zeta \cos \varphi d\zeta = S_{\psi_S}. \end{aligned} \quad (2.9)$$

Vertical integration through the atmospheric column and application of surface and upper boundary conditions has provided an equation for the column integrated time rate of change of constituent ( $\psi$ ) per unit mass (first term on the left hand side (LHS) of Equation 2.9) as a function of column integrated zonal advection (second term on LHS), meridional advection (third term on LHS), and sources and sinks of  $\psi$  at the Earth's surface (RHS of Equation 2.9). The zonal term is eliminated from Equation 2.9 by averaging around a latitude circle using the zonal mean operator  $[\ ]$ . Averaging in time using the time mean operator  $\overline{(\ )}$  reduces Equation 2.9 to

$$\frac{\partial}{\partial t} \left[ \overline{\int_{\zeta_S}^{\zeta_T} m_\zeta \psi_\zeta d\zeta} \right] + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \left[ \overline{\int_{\zeta_S}^{\zeta_T} m_\zeta v_\zeta \psi_\zeta \cos \varphi d\zeta} \right] = \overline{[S_{\psi_S}]}. \quad (2.10)$$

Finally, assuming  $\zeta_T$  and  $\zeta_S$  do not vary in time or space, the mean operators,  $\overline{(\ )}$  and  $[\ ]$ , may be moved inside the vertical integrals. Equation 2.10 is then rewritten as

$$\frac{\partial}{\partial t} \int_{\zeta_S}^{\zeta_T} \overline{[m_\zeta \psi_\zeta]} d\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \int_{\zeta_S}^{\zeta_T} \overline{[m_\zeta v_\zeta \psi_\zeta]} \cos \varphi d\zeta = \overline{[S_{\psi_S}]}. \quad (2.11)$$

Equation 2.11 reduces to the continuity equation for total mass (air) by setting  $\psi$  to unity and the source term on the RHS to zero; for total mass, the net meridional transport averages to zero over sufficiently long time periods. This represents total mass balance for a “ring” of air along a latitude circle.

By setting  $\psi$  to CO<sub>2</sub> mixing ratio ( $C$ ), the zonal and time averaged vertically integrated tendency equation for atmospheric CO<sub>2</sub> is expressed as

$$\frac{\partial}{\partial t} \int_{\zeta_T}^{\zeta_S} [\overline{m_\zeta C_\zeta}] d\zeta + \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \int_{\zeta_T}^{\zeta_S} [\overline{m_\zeta v_\zeta C_\zeta}] \cos \varphi d\zeta = \left[ \overline{(F_c)_S} \right], \quad (2.12)$$

where  $(F_c)_S$  is the surface flux of CO<sub>2</sub>, which has units of  $kgC \ m^{-2} \ s^{-1}$ .

### 2.5.2 Eddy Decomposition

The integrand of the second term on the LHS of Equation 2.12,  $[\overline{m_\zeta v_\zeta C_\zeta}]$ , represents net meridional transport of CO<sub>2</sub> by the atmosphere in the time and zonal mean at an arbitrary atmospheric layer. To assess influences of the mean circulation and baroclinic waves on meridional transport and polar seasonality, eddy decomposition is applied according to methods outlined by Peixoto and Oort [1992], in which quantities (lets call these  $x$ , representing mass flux  $(mv)_\zeta$  and tracer mixing ratio  $C_\zeta$ ) are “perturbed” at an arbitrary  $\zeta$  surface and rewritten as  $x = \bar{x} + x'$ , representing the time mean and its temporal deviation  $(\ )'$  or “transient component”, and  $x = [x] + x^*$ , representing the zonal mean and its zonal deviation  $(\ )^*$  or “eddy component”.

With these definitions, we consider various statistics derived from the three-dimensional fields of  $(mv)_\zeta$  and  $C_\zeta$ . To begin eddy decomposition, we first set  $a = (mv)_\zeta$  and  $b = C_\zeta$  and perturb quantities in time.  $\overline{ab}$  is then decomposed into its time mean and transient

components and rewritten as

$$\begin{aligned}
\overline{ab} &= \overline{(\bar{a} + a') (\bar{b} + b')} \\
&= \overline{\bar{a}\bar{b}} + \overline{\bar{a}b'} + \overline{a'\bar{b}} + \overline{a'b'} \\
&= \bar{a}\bar{b} + \overline{\bar{a}b'} + \overline{a'\bar{b}} + \overline{a'b'} \\
&= \bar{a}\bar{b} + \overline{a'b'}.
\end{aligned} \tag{2.13}$$

The RHS of Equation 2.13 simplifies from four terms (third line) to two terms (fourth line) because the time average of a prime, namely  $\bar{a}'$  and  $\bar{b}'$ , is approximately zero over sufficiently long time intervals. Next, zonal averaging is applied to Equation 2.13 and both terms on the RHS are decomposed into its zonal mean and eddy components. The first term is rewritten as

$$\begin{aligned}
[\bar{a}\bar{b}] &= \left[ \overline{([\bar{a}] + a^*) ([\bar{b}] + b^*)} \right] \\
&= [\bar{a}][\bar{b}] + [\bar{a}]b^* + a^*[\bar{b}] + a^*b^* \\
&= [\bar{a}][\bar{b}] + [\bar{a}] [\bar{b}^*] + [a^*] [\bar{b}] + [a^*b^*] \\
&= [\bar{a}][\bar{b}] + [a^*b^*],
\end{aligned} \tag{2.14}$$

where the RHS of Equation 2.14 simplifies from four terms (third line) to two terms (fourth line) because the zonal average of an eddy, namely  $[a^*]$  and  $[b^*]$ , is by definition zero around a latitude circle. Similarly, the second term in the RHS of Equation 2.13 is rewritten as

$$\begin{aligned}
[\overline{a'b'}] &= \left[ \overline{([a'] + a'^*) ([b'] + b'^*)} \right] \\
&= [\overline{a'}][\overline{b'}] + [\overline{a'}]b'^* + a'^*[\overline{b'}] + a'^*b'^* \\
&= [\overline{a'}][\overline{b'}] + [\overline{a'}][\overline{b'^*}] + [a'^*][\overline{b'}] + [a'^*b'^*] \\
&= [\overline{a'}][\overline{b'}] + [a'^*b'^*]
\end{aligned} \tag{2.15}$$

where the RHS of Equation 2.14 simplifies from four terms (third line) to two terms (fourth line) because the zonal average of the eddy of a prime, namely  $[a'^*]$  and  $[b'^*]$ , is by definition zero around a latitude circle.

Combining Equations 2.14 and 2.15 and substitution of  $(mv)_\zeta$  and  $C_\zeta$  for  $a$  and  $b$  yields

$$\overline{m_\zeta v_\zeta C_\zeta} = \overline{(mv)_\zeta} \overline{C_\zeta} + \overline{(mv)_\zeta^* C_\zeta^*} + \overline{(mv)_\zeta'} \overline{C_\zeta'} + \overline{(mv)_\zeta'^* C_\zeta'^*}. \quad (2.16)$$

Eddy decomposition is an extension of Reynolds averaging and applies to the large-scale, geostrophically and hydrostatically balanced flow. Time averaging of one month is applied, and assumed to be a sufficiently long period for time averaged prime quantities to disappear on the third line of Equation 2.13. The focus of this study is synoptic transport by baroclinic waves, which occurs on time scales of  $\sim 3$ -7 days and spatial scales of  $\sim 1000$  kms. By defining the time mean using a time window larger than the synoptic time scale, the net effect of eddies over multiple synoptic cycles is captured. Through these zonal and temporal averaging techniques, the statistics of  $\text{CO}_2$  transport is reduced to meridional and vertical statistics.

The first term on the RHS of Equation 2.16 refers to the mean meridional circulation (MMC), and accounts for transport of mean  $\text{CO}_2$  by the mean meridional mass flux. The second term refers to stationary eddies, and represents transport by stationary waves due to correlated zonal deviations of mass flux and  $\text{CO}_2$ . These are forced primarily by the geographic anchoring of planetary scale atmospheric waves by zonal inhomogeneities in boundary conditions such as continents, oceans, ice sheets and mountains. The third term accounts for temporal variations from zonal mean circulations, namely the transient symmetric circulation, and contributes to temporal fluctuations in circulations such as the Hadley Cell due to north-south seasonal migration. The fourth term accounts for temporal variations from the stationary planetary waves, that is transient eddies, representing transport by time varying dynamical processes such as baroclinic waves due to correlated temporal variations of mass flux and  $\text{CO}_2$ . In mid-latitudes, baroclinic instability is a major driver of transient eddies. The transient symmetric circulation is small in middle latitudes relative to the other terms and will be absorbed into transient wave term for the remainder of this study. Transport by transient and stationary waves together represents the eddy



component of the large-scale mid-latitude circulation.

### 2.5.3 *Vertical Coordinate*

Eddy decomposition allows for analysis of global transport of atmospheric constituents such as CO<sub>2</sub> by large scale mean and eddy circulations. These transport processes are analyzed in different vertical coordinate frameworks to simplify our understanding of eddy and mean transport of CO<sub>2</sub>. This study analyzes transport in both Eulerian and Lagrangian coordinate systems.

Eulerian coordinate systems represent a reference system that is fixed with respect to time and space. Pressure coordinates ( $p$ ) are a standard way of analyzing three-dimensional atmospheric fields. Pressure surfaces, however, intersect the ground at locations that change with time. Because of the spatio-temporal dependence of pressure levels, the column integral operation in Equation 2.10 cannot be moved outside of the averaging operators and eddy decomposition cannot be applied for single atmospheric layers. Terrain following coordinates such as  $\eta$ , on the other hand, work well for eddy decomposition because vertical layers are fixed in time and space.

Lagrangian vertical coordinate systems follow the motion of an air parcel as it moves vertically, and can be thought of as a reference system that is fixed with respect to the motion of a parcel of air. An example of a Lagrangian coordinate system is potential temperature, which is defined by  $\theta = T \left( \frac{p_0}{p} \right)^{\frac{R}{C_p}}$ , with  $p$  the pressure,  $R$  the ideal gas constant,  $C_p$  the specific heat,  $T$  the temperature, and  $p_0 = 1000$  mbar an arbitrary reference pressure. Surfaces of constant  $\theta$  are isentropic surfaces.  $\theta$  is conserved following a particle for reversible dry adiabatic transformations. This means that the vertical motion in  $\theta$ -coordinates is proportional to the heating rate, and that in the absence of diabatic heating (radiation, evaporation, sensible heating), there is “no vertical motion,” from the point of view of  $\theta$ -coordinates.  $\theta$  increases upwards in a statically stable atmosphere so that there is a monotonic relationship between  $\theta$  and height above ground level,  $z$ .

Eddy decomposition on isentropic coordinates is relatively straightforward, but some definitions must first be made to accommodate some complications. Namely, it is common for isentropic surfaces to intersect the lower boundary in the presence of strong horizontal temperature gradients, as occur for example along frontal zones. To avoid surfaces running through the ground, isentropic surfaces are assumed in this study to follow the surface boundary. This effect is imposed by defining imaginary “massless layers,” such that the pressure difference (or thickness) between two  $\theta$  surfaces is zero. Since no mass resides between the  $\theta$  surfaces in the portion of the domain where they “touch the Earth’s surface,” no mass conservation laws are violated when computing mass transport of tracer. This does, however, present a problem for tracer transport along massless surfaces because tracers such as  $C$  are undefined.

The following definitions are therefore made for quantities involving  $C$  so that massless layers can be applied to zonal and temporal averaging techniques (i.e., Equation 2.16) when isentropic coordinates intersect the surface:

$$\overline{C_\zeta(i, j, k)} = \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C_\zeta(i, j, k, t) \quad \text{for } m_\zeta(i, j, k, t) > 0 \quad (2.17)$$

$$C_\zeta(i, j, k, t) = \overline{C_\zeta(i, j, k)} \quad \text{for } m_\zeta(i, j, k, t) = 0 \quad (2.18)$$

$$C'_\zeta(i, j, k, t) = C_\zeta(i, j, k, t) - \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C(i, j, k, t), \quad (2.19)$$

where  $i, j, k, t$  are indices for longitude, latitude, height, time and  $T_{mass}$  is equal to the total number of layers in the time averaging window (1 month) that have non-zero mass ( $m_\zeta > 0$ ). These definitions guarantee that  $C' = 0$  for massless layers.

As a proof of concept, these definitions are applied to Equation 2.13

$$\begin{aligned} \overline{(mv)C} &= \overline{(\overline{mv} + (mv)') (\overline{C} + C')} \\ &= \overline{(\overline{mv} + (mv)') \left( \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C + \left( C - \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C \right) \right)}. \end{aligned}$$

Rearranging terms gives

$$= \overline{(\overline{mv} + (mv)')} \left( \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C - \frac{1}{T_{mass}} \sum_{t=1}^{T_{mass}} C + C \right),$$

Finally, terms involving  $C$  cancel out, leaving

$$= \overline{(mv)} \overline{(C)}.$$

This guarantees that the LHS and RHS of the first line in Equation 2.13 are equal for massless layers as long as definitions 2.17-2.19 are applied consistently throughout eddy decomposition.

Next I show how definitions 2.17-2.19 are applied to Equation 2.13 so that the first line of Equation 2.13 simplifies to the last line of Equation 2.13. Substitution of  $mv$  for  $a$  and  $C$  for  $b$  in Equation 2.13 gives

$$\begin{aligned} \overline{(mv)C} &= \overline{(\overline{mv} + (mv)')} \overline{(\overline{C} + C')} \\ &= \overline{(\overline{mv})\overline{C}} + \overline{(\overline{mv})C'} + \overline{(mv)'\overline{C}} + \overline{(mv)'C'} \\ &= \overline{(\overline{mv})\overline{C}} + \overline{(\overline{mv})C'} + \overline{(mv)'\overline{C}} + \overline{(mv)'C'}. \end{aligned} \quad (2.20)$$

Averages of terms involving  $C$  are taken for layers with mass ( $\overline{C}$  and  $\overline{C'}$ ) using Equation 2.17 and 2.19, averages of terms involving  $mv$  are taken over all layers ( $\overline{mv}$  and  $\overline{(mv)'}$ ), and averages of terms involving both  $mv$  and  $C$  are taken over all layers ( $\overline{(\overline{mv})C}$  and  $\overline{(\overline{mv})'C'}$ ) using Equation 2.18. Using these definitions, the time average of primes in Equation 2.20 are zero and Equation 2.20 reduces to

$$\overline{(mv)C} = \overline{(\overline{mv})\overline{C}} + \overline{(mv)'\overline{C'}}. \quad (2.21)$$

Note also that, with these definitions, averages of terms involving  $C$  can be taken for all layers with the same result. This is more algebraically correct, but either definition can be used. In the results that follow,  $\overline{C}$  is defined for layers with non-zero mass.

The LHS of Equation 2.21 is independent of our definition of  $C$  for massless layers because by definition  $m$  is zero and thus  $C$  is always multiplied by zero. The balance

of terms on the RHS is however strongly influenced. For example, by setting  $C$  to zero for massless layers and using all layers in the calculation of  $[\overline{C}]$  in the first term on the RHS,  $[\overline{C}]$  attains a physically unrealistic value for a particular  $\theta$  surface in the presence of massless layers due to including layers with  $C = 0$  in the average, and is therefore biased towards small values. This small bias also reduces mean meridional transport of  $\text{CO}_2$ , forcing the eddy term to compensate in order to keep balance. It is more desirable for mean meridional transport of  $\text{CO}_2$  to involve transport of a physically meaningful value of tracer that represents an average value for layers with mass. As long as  $C = C' + \overline{C}$  and  $C = C^* + [C]$  is consistent throughout eddy decomposition, the LHS of Equation 2.13 is always be equal to the RHS when massless layers are encountered in the statistics. The definitions provided here, however, assure that partitioning of eddy and mean transport is done in a physically realistic manner. Furthermore, since these definitions result in non-zero mixing ratios for massless layers, no additional definitions are necessary for zonal quantities, and Equations 2.14 and 2.15 can be solved using all layers.

When describing eddy and mean transport processes in an averaged sense, “dry isentropes” (referred to here as surfaces of constant  $\theta$ ) and “moist isentropes” (defined as  $\theta_e \sim \theta + \frac{L_v}{c_p} q$ , with  $L_v$  the latent heat of vaporization and  $q$  the water vapor concentration) are more indicative of parcel trajectories through mid-latitude stormtracks than Eulerian averages. Horizontal flow along isentropic surfaces contains (approximately) the adiabatic component of vertical motion (assuming radiative cooling and sensible heating are relatively unimportant for synoptic motions) that is often neglected in a Eulerian reference system. In addition,  $\theta_e$  is approximately conserved (internal energy neither lost nor gained) in the presence of latent heating [Pauluis et al., 2008]. Mass transport on  $\theta_e$  surfaces includes a large contribution from moist air rising within stormtracks, while other forms of averaging (Eulerian or dry isentropes) misses entirely this contribution (the corresponding transport ends up in the mean term). The vertical component of motion gained from latent heat release in condensing moist air is critical for describing tracer transport by moist ascent

within baroclinic waves. In Chapter 3, eddy and mean transport is examined by setting  $\zeta$  to  $\eta$ ,  $\theta$ , and  $\theta_e$  in Equation 2.16. In Chapter 4, transport is examined only with  $\zeta$  set to  $\theta_e$ .

## 2.6 Methodology for Inversion used in Part 3

End-to-end inversion experiments are carried out in Chapter 5 to determine whether transport bias discussed in Chapter 4 is aliased into errors in surface flux estimation. The inversion system uses the ensemble-based MLEF optimization technique, which was discussed briefly toward the end of Chapter 1.1.3. The following section includes discussion of the inversion technique, followed by discussion of the strategy for sampling forward transport simulations for synthetic satellite data, then a detailed description of the major steps in the MLEF data assimilation cycle, and ending with details regarding covariance smoothing, prior flux estimates, basic terminology, and net flux and uncertainty calculations. MLEF discussion follows techniques and discussion provided by Lokupitiya et al. [2008].

### 2.6.1 Assimilation Scheme

The optimization problem is represented as solving for unknown persistent multiplicative biases in photosynthesis, respiration, and air-sea gas exchange. Net exchange of carbon within a single grid point is represented as

$$F(x, y, t) = (1 + \beta_{RESP}(x, y))RESP(x, y, t) - (1 + \beta_{GPP}(x, y))GPP(x, y, t) + (1 + \beta_{OCEAN}(x, y))OCEAN(x, y, t) + FF(x, y, t), \quad (2.22)$$

where  $x$  is latitude,  $y$  is longitude, and  $t$  is time.  $\beta$ 's represent persistent multiplicative biases in time-varying grid-scale component land and ocean fluxes. Land fluxes are represented by heterotrophic respiration, or  $RESP(x, y, t)$ , and gross primary production, or  $GPP(x, y, t)$ , which are derived from SiB3 at hourly resolution. Ocean fluxes are represented by a net daily flux, or  $OCEAN(x, y, t)$ , and are derived from Doney et al. [2009a; 2009b]. Fossil fuel flux,  $FF(x, y, t)$ , is derived from the Vulcan Project [Gurney et al., 2009] and is assumed to be perfect (i.e., biases aren't included in the fossil fuel term).

Unlike ocean flux, which is prescribed as a net flux, fluxes in SiB3 have components into and out of the land ( $GPP$  and  $RESP$ , respectively). The strategy over land is therefore not to solve for biases in net flux, but assume that biases exist in component fluxes. SiB3 is balanced annually such that the sum of  $GPP$  at some grid point  $(x, y)$  from  $t = 1$  to  $t = n$ , where  $n$  is the total number of hours in a year, is equal to the sum of  $RESP$  from  $t = 1$  to  $t = n$ . The real world has slowly varying terrestrial sinks that add up to something like  $2.0\text{-}3.0\text{ GtC yr}^{-1}$  globally. To represent net annual sinks in this framework net uptake must exceed net respiration; i.e.,  $GPP$  must be larger than  $RESP$  in the annual mean. This is accomplished with  $\beta$ 's. Let's assume, for example, that SiB3 predicts respiration and GPP fluxes of  $120\text{ GtC yr}^{-1}$  for the entire globe. There are an infinite number of combinations of  $\beta_{GPP}$  and  $\beta_{RESP}$  that could be used to obtain a  $2.0\text{ GtC yr}^{-1}$  net terrestrial sink, as long as  $\beta_{GPP} > \beta_{RESP}$ .  $\beta_{RESP} = -0.02$  and  $\beta_{GPP} = 0.02$  are two possible values. Put into context, these values mean that modeled respiration is overestimated in SiB3 by about 2% and GPP is underestimated by about 2%.  $\beta_{GPP}$  can also be less than zero, suggesting SiB3 overestimates GPP, as long as  $\beta_{RESP}$  is more negative.

The idea is that high frequency time variations in respiration and photosynthesis are driven by relatively well-understood and easily modeled processes that are determined to first order by variations in solar radiation [Zupanski et al., 2007a]. Mechanisms governing the physiological response of plants to weather and quick turnover times in the soil carbon pool are fairly well represented in SiB3 by process based equations, reanalyzed weather fields, vegetation cover and soil texture maps, and satellite derived phenology. Processes not represented in SiB3 such as underestimation of available nitrogen, forest management, agricultural land-use may lead to persistent biases in photosynthesis, while poor prescription of soil carbon pools or coarse woody debris may bias respiration calculations. This technique allows for  $RESP$  and  $GPP$  to vary on hourly, synoptic, and seasonal timescales and for air-sea exchange on daily time scales, but assumes that biases in these fluxes persist for longer periods.

Lokupitiya et al. [2008] solved for 8 week biases using synthetic data based on the in-situ network. The 8 week window was chosen because it takes several weeks to months for the atmosphere to mix sufficiently such that in-situ stations can feel CO<sub>2</sub> pulses from distant source regions. Here biases are allowed to persist for 2 weeks because satellite data provides approximately global coverage in just a few days (GOSAT has global coverage in 3 days) such that surface flux signals are sufficiently constrained.

### 2.6.2 *Synthetic Satellite Data*

Synthetic satellite data refers to column averaged CO<sub>2</sub> mixing ratios obtained from a tracer transport model. This data is generated by sampling pressure weighted CO<sub>2</sub> output from GEOS5-0.67x0.5 along the GOSAT orbit using an orbit simulator provided by David Baker [personal communication, November 15, 2010]. GOSAT (discussed in Chapter 1.1.1) uses a sun-synchronous orbit with early afternoon sun-lit equator crossing time ( $\sim 1:30$  pm local time) and orbital inclination near  $98^\circ$ . Since GOSAT measures CO<sub>2</sub> absorption using reflected solar radiation, the transport model is sampled only during daytime during the descending mode of the orbit. Subsequent orbits are separated by  $\sim 25^\circ$  in longitude and  $\sim 99$  minutes apart. GOSAT points near-nadir as well as at the sun glint spot, which greatly increases the signal over the ocean. GOSAT measures with 13.4 cross-scans per minute (5 total cross scans) with a field of view of  $\sim 100$  km<sup>2</sup>.

The number of points collected along a GOSAT orbit in one year (assuming cloudy pixels are not thrown away) according to the orbit simulator is shown in Figure 2.2A. A maximum of 281 measurements are sampled by GOSAT in one hour. This corresponds to 6744 points in a day and 94,416 points collected over a two week period. The number of samples over land is much larger than over ocean because of cross-scans. Also note this map represents the number of samples collected in a  $2.5^\circ \times 2^\circ$  grid box. Since multiple samples may occur over these large grid boxes during a single GOSAT pass, it is possible for the total number of data at a grid cell to exceed the actual number of samples at a single point.

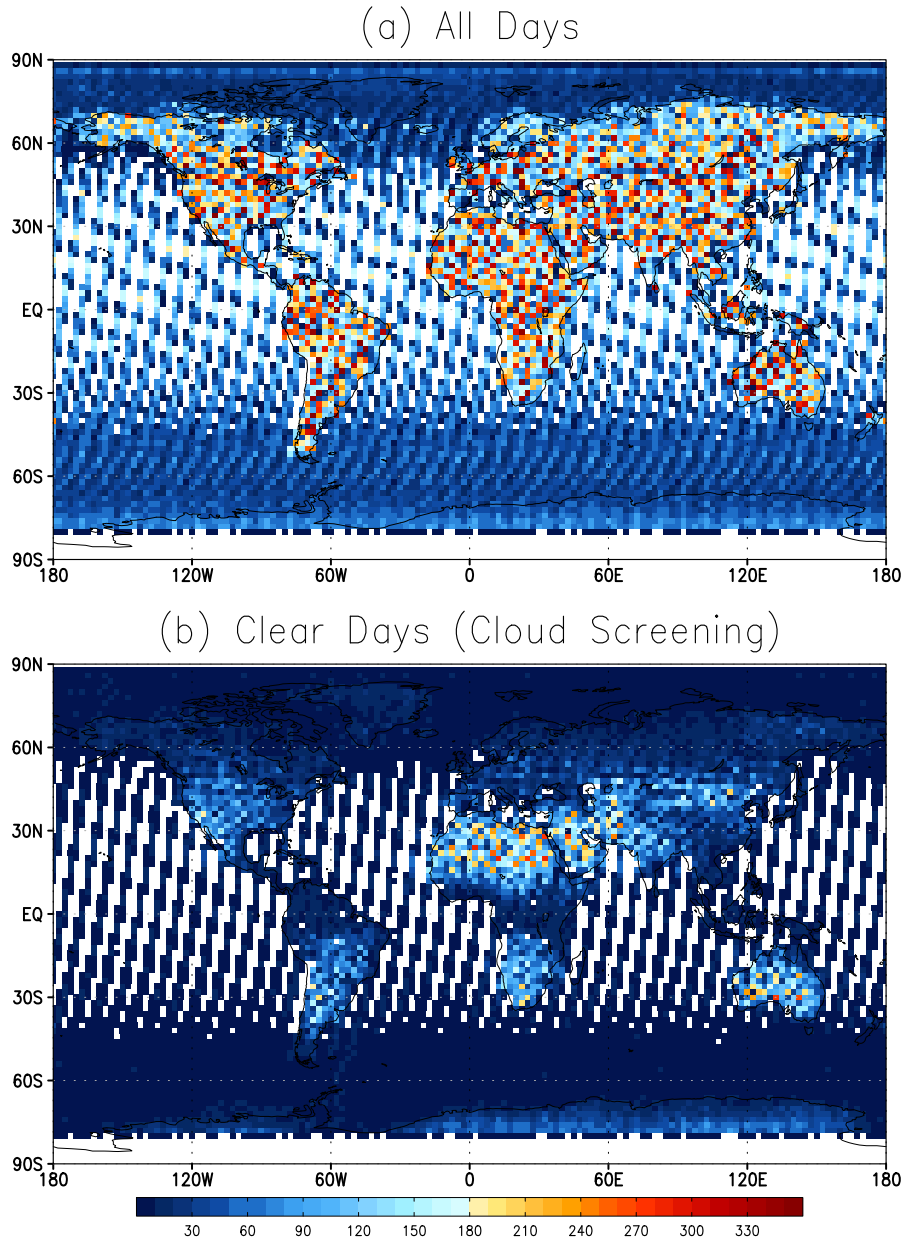


Figure 2.2: Number of points sampled by pixels in GEOS4-2.5x2 along the GOSAT orbit in one year for all days (a) and clear days (b).

The total number of measurements per year is independent of the size of the grid box used for sampling. For example, the number of points per grid box is reduced significantly when  $0.67^\circ \times 0.5^\circ$  grid spacing is used at the expense of more grid boxes. Nevertheless, these numbers represent a significant increase over the number of in-situ points.

A significant portion of satellite data must be thrown away due to cloud contami-



nation. Cloud contamination is especially bad for GOSAT due to its large field of view relative to an OCO like instrument. This means GOSAT will only detect CO<sub>2</sub> mixing ratios during fair weather days which could lead to temporal sampling biases as discussed in Chapter 1.1.1, especially if simulated transport differs greatly from reality (or other models) underneath cloudy pixels. Assuming satellites could see through gaps in clouds, and were therefore able to sample the atmosphere during inclement weather, it's possible temporal sampling errors would be much smaller (as might be the case for a laser sounder). As discussed in Chapter 5, synthetic data is generated, and then inverted, for all conditions (clear + cloudy sky) and clear conditions only to test for the effect of fair weather bias on flux optimization. Clear conditions are defined as grid cells whose cloud optical depth ( $\tau$ ) is less than 0.3. Cloud optical depth is prescribed from GEOS5-1.25x1 to be consistent with the transport model, and is defined as the grid scale value (total in-cloud optical depth from ice and liquid water times the three dimensional total cloud fraction in a grid box). Data reduction due to discarding pixels with  $\tau > 0.3$  is shown in Figure 2.2B. Regions where clouds persist, such as tropical and boreal latitudes, become very poorly sampled as a result. The effect of data reduction is obvious when is plotted as a percentage (clear days divided by all days) in Figure 2.3.

Observations, whether in-situ or remote, must be assigned some measure of uncertainty. This so-called observation error corresponds to the diagonal of the observation covariance matrix,  $\mathbf{R}$  (see next section). Observation error is determined by factors including (1) measuring instrument error, (2) transport error, and (3) the error due to scale mismatch between observations and the transport model, also referred to as representativeness error. Corbin et al. [2006], for example, investigated representativeness error for satellite data due to sampling only in clear conditions (clear-sky sampling error, discussed in more detail in Chapter 1) and found bias up to -0.4 ppm. Measurement errors, especially for high precision in-situ measurements, are typically very small ( $< 0.5$  ppm) compared to model transport error (up to 20 ppm depending on location, time of day, and type of transport).

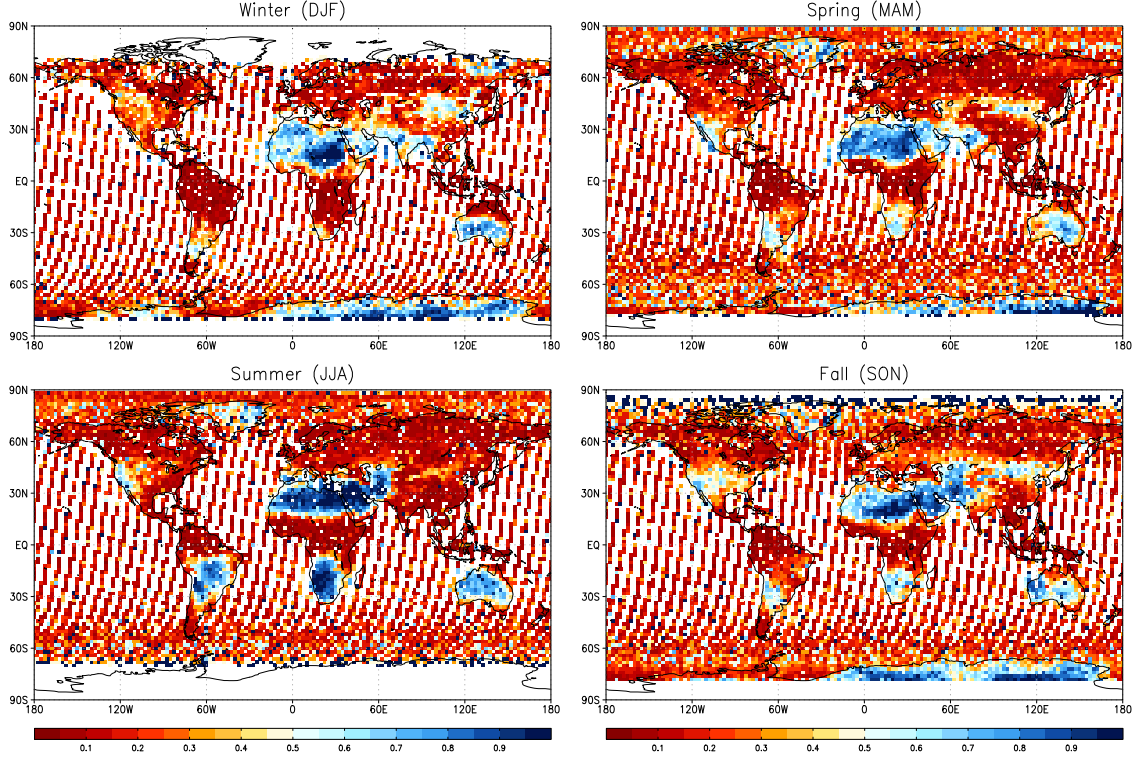


Figure 2.3: Percent of points discarded due to cloud screening for winter (top left), spring (top right), summer (bottom left), and fall (bottom right).

For satellites, measurement errors are typically much larger, more variable (depending on the satellite, active versus passive measurements, albedo effects over the surface, latitude, viewing angle, etc.), and generally larger than 1 ppm. Measurement errors may also be purely random or systematic and biased.

All OSSEs in Chapter 5 assume synthetic satellite data from GOSAT has a 3 ppm observation error ( $\sigma = 3$  ppm) due to random measurement error. This error is added to synthetic measurements as a measure of instrument precision by creating a vector of white gaussian noise, with mean of zero and  $\sigma = 3$ , that is equal to the size of the observation vector for one year. 3 ppm measurement noise is chosen to represent the approximate uncertainty of GOSAT. This noise is added to synthetic measurements after sampling from the forward runs.

### 2.6.3 MLEF

There are two major steps involved in a data assimilation cycle: (1) the analysis step and (2) the forecast step. In the analysis step the optimal state is found by minimizing the cost function provided the observations inside the analysis window. In the forecast step, a prior for the next cycle is found by applying a forecast model, in this case PCTM, to the optimal state found in the analysis step.

MLEF attempts to find the maximum likelihood solution for the bias parameters  $\beta_{GPP}$ ,  $\beta_{RESP}$ , and  $\beta_{OCEAN}$  using ensemble-based data assimilation by minimizing the following cost function

$$J(\beta) = \frac{1}{2}[\mathbf{y} - H(\beta)]^T \mathbf{R}^{-1}[\mathbf{y} - H(\beta)] + \frac{1}{2}[\beta - \beta_b] \mathbf{P}_f^{-1}[\beta - \beta_b], \quad (2.23)$$

where  $\beta$  is the vector of unknown bias correction terms (the state vector being solved),  $\mathbf{y}$  is a vector of observations,  $H(\beta)$  is an observation operator,  $\beta_b$  is the prescribed prior estimate for bias correction terms,  $\mathbf{R}$  is the observation error covariance matrix, and  $\mathbf{P}_f$  is the forecast (prior) error covariance matrix. The vector of observations is represented by the synthetic satellite data. GEOS4-2.5x2 serves as the observation operator in Equation 2.23 to communicate surface fluxes to observation locations. The matrix  $\mathbf{P}_f$  defines the prior uncertainty of  $\beta$ , also known as the forecast error covariance.

Equation 2.23 is minimized via an iterative conjugate-gradient algorithm with Hessian preconditioning (described below), and is robust for non-linear processes (not considered in this study). Iterative minimization converges in a single iteration to the Kalman filter solution,

$$\beta = \beta_b + P_f H^T (H P_f^T H^T + R)^{-1} [y - H(\beta_b)], \quad (2.24)$$

when the observation operator is linear and the ensemble size is equal to the size of the control vector [Zupanski 2005, Appendix A]. Here the control variable is the vector of unknowns. However, in the following experiments, the ensemble size is much smaller than the size of the control variable, which may degrade the MLEF solution. As demonstrated by

Zupanski et al [2007], however, the MLEF solution smoothly converges to the KF solution as the ensemble size approaches the size of the control vector, which provides justification for using smaller ensemble sizes.

A major advantage of MLEF is that it produces an estimate of the analysis uncertainty, or analysis error covariance. Unlike variational techniques which require full rank of the solution, the prior and posterior uncertainties of Equation 2.24 are defined in ensemble subspace as square roots of the forecast error covariance  $P_f^{\frac{1}{2}}$  and the analysis error covariance  $P_a^{\frac{1}{2}}$ :

$$P_a^{\frac{1}{2}} = P_f^{\frac{1}{2}}(I + A)^{-\frac{1}{2}} \quad (2.25)$$

where  $A$  is the information matrix in ensemble subspace [Zupanski et al., 2007]. The analysis error covariance changes in time (decreasing, ideally) due to the impact of assimilated observations involved in  $A$ . The information matrix can also be used as a guide when determining an adequate ensemble size. According to Lokupitiya et al. [2008] the ensemble size can also be determined from the information measure referred to as Degrees of Freedom for Sigma, or DFS, with an ensemble size considered appropriate if further increase in ensemble size, given a certain number of observations, does not change DFS.

The major steps of MLEF, as described in Zupanski [2005], are summarized below.

#### *Step 1: Ensemble Forecast*

This step generates ensemble members using a parallel computing environment, using initial ensemble perturbations in the first assimilation cycle (“cold start”) or the analysis covariance matrix in subsequent assimilation cycles (“warm start”). Use of the analysis covariance matrix to generate ensemble members is an important way of propagating information between successive assimilation cycles. The purpose of this step is to generate a bunch of random guesses used in minimization of the cost function; the more ensembles used, the more certain the posterior probability distribution the solution generates. Ensemble members are also used in this step to calculate the forecast error covariance, which is used

in a later step for calculation of the updated analysis error covariance.

*Step 2: Innovation Vector*

This is the first step in the analysis, in which the ensemble forecast is interpolated to observation locations. These calculations can be done on separate processors for each ensemble member.

*Step 3: Hessian Preconditioning*

The first step in minimizing the cost function is to compute second order partial derivatives of the cost function so that a maximum likelihood analysis solution is obtained in a single step. This is known as Hessian preconditioning and requires calculation of matrices from innovation vectors calculated in Step 2, which can also be done through parallel computing. Eigenvalue decomposition is used to speed up matrix inversion used in Hessian preconditioning. The motivation for this step is that minimizing the cost function directly from ensemble-spanned subspace can sometimes be difficult because multimodal distributions may be present and gradient search algorithms may not converge on the correct solution. Hessian preconditioning transforms the initial distribution to something that can be minimized in just a couple steps.

*Step 4: Gradient Calculation*

The next step is to calculate the gradient in ensemble-spanned subspace. Here the cost function is redefined through the change of variable introduced through Hessian preconditioning. The first derivative of the new cost function is defined and minimization is calculated using the preconditioned steepest descent.

*Step 5: Analysis Error Covariance*

MLEF calculates the analysis error covariance using the optimized analysis determined from Step 4 (the inverse Hessian calculated at the minimum). The analysis error

covariance is then used as initial perturbations for the next assimilation cycle.

#### 2.6.4 *Covariance Smoothing and Localization*

Strong covariance smoothing in the first assimilation cycle and covariance localization in all cycles are required in experiments where the observation vector is small compared to the number of unknowns [Lokupitiya et al., 2008]. This was certainly the case in Lokupitiya et al. [2008] because of the use of relatively sparse in-situ measurements. The observation vector is significantly larger in this experiment, but smoothing and localization are carried out nevertheless. This may especially helpful when the number of observations is reduced by cloud screening. Covariance smoothing is applied during the “cold start” by adding spatial correlations with an e-folding length of 800 km over land and 1600 km over ocean. After the first assimilation cycle, covariance smoothing is not necessary. Spatial covariance improves with time because of propagation of the error covariance through the analysis system and the opportunity to learn about spatial structure from observations later in the year. Localization is important in ensemble data assimilation because it prevents sampling errors at large distances and thereby reduces the ensemble size required for the analysis.

#### 2.6.5 *Defining Priors*

As a starting point, simulated carbon fluxes are assumed to be unbiased such that  $\beta_{GPP} = \beta_{RESP} = \beta_{OCEAN} = 0$  at every grid point (in contrast to  $\beta$ ’s centered around a value of 1 as in previous MLEF experiments). In all subsequent cycles, the estimated biases from the previous cycle are used as priors. If satellite data contains information that is different from the priors,  $\beta$ ’s will evolve with time to match the data. Otherwise, there will be no need for  $\beta$ ’s to change. Several OSSEs are run in Chapter 5 to determine if the inversion can distinguish between biased and unbiased satellite data.  $\beta$ ’s may also change in time if the inversion “thinks” the satellite data contains information different from the priors (as might occur, for example, in the presence of biased transport).

Prior uncertainties for biases are selected that allow  $\beta$ 's to learn from the satellite data. Choosing too tight or too loose uncertainties may prevent reaching a reasonable solution. Since the biases aren't known, constant prior standard deviations are assumed for the biases at each grid point of  $\sigma_{GPP} = 0.2$ ,  $\sigma_{RESP} = 0.2$ ,  $\sigma_{OCEAN} = 0.1$ . These uncertainties are somewhat generous considering they act on large component fluxes of  $GPP$  and  $RESP$ . In contrast, if operating on NEE, which is much smaller, more variable, and therefore more uncertain, these uncertainties would be considered too small.

### 2.6.6 Terminology

Five types of surface fluxes are discussed in Chapter 5: (1) prior fluxes or “priors”, (2) analyzed or optimized fluxes, (3) true fluxes or “truth”, (4) flux estimates, and (5) flux errors. Prior fluxes refer to modeled flux components including  $GPP$ ,  $RESP$ , and  $OCEAN$  that are used as a-priori fluxes for the inversion. These fluxes actually take the form  $(1 - \beta_{GPP}) * GPP$ ,  $(1 - \beta_{RESP}) * RESP$ ,  $(1 - \beta_{OCEAN}) * OCEAN$ , but are assumed to be unbiased at the start of the inversion with  $\beta_{GPP} = \beta_{RESP} = \beta_{OCEAN} = 0$ . Analyzed, or optimized, fluxes refer to a-posteriori fluxes that take the same form as the priors but with possibly non-zero  $\beta$ 's in the case that the inversion finds a signal in the satellite data different from prior constraints. Flux components are assumed to be modeled perfectly and it is the job of the inversion to search the satellite data for possible biases. True fluxes are fluxes that are prescribed in forward simulations to generate synthetic satellite data. These can be “hypothetical” fluxes if used to mimic real world sinks by setting  $\beta$  correction terms to non-zero values, or can reduce to priors when  $\beta$ 's are set to zero. Flux estimates refer to the difference between analyzed and prior fluxes. Flux errors refer to the difference between analyzed and true fluxes.

### 2.6.7 Flux and Uncertainty Calculation

Fluxes and uncertainties are calculated according to Lokupitiya et al. [2008]. To estimate monthly average terrestrial NEE, the following calculation is used:

$$F_{CO_2} = (1 - \beta_{RESP})\overline{RESP} - (1 - \beta_{GPP})\overline{GPP}, \quad (2.26)$$

where  $\overline{(\dots)}$  represents the monthly average. The corresponding uncertainty estimate is given by

$$\begin{aligned} \sigma_{CO_2}^2 &= \overline{RESP}^2 \sigma_{\beta_{RESP}}^2 + \overline{GPP}^2 \sigma_{\beta_{GPP}}^2 \\ &\quad - 2 \times \overline{RESP} \times \overline{GPP} \times Cov(\beta_{RESP}, \beta_{GPP}), \end{aligned} \quad (2.27)$$

where  $\sigma_{\beta_{RESP}}^2 = Var(\beta_{RESP})$  and  $\sigma_{\beta_{GPP}}^2 = Var(\beta_{GPP})$ , where  $Var(\dots)$  represents the variance, which are represented by diagonal values within the covariance matrix. Similarly, the following calculation is used for air-sea exchange:

$$F_{CO_2} = (1 - \beta_{OCEAN})\overline{OCEAN}, \quad (2.28)$$

with the corresponding uncertainty estimate given by

$$\sigma_{CO_2}^2 = \overline{OCEAN}^2 \sigma_{\beta_{OCEAN}}^2, \quad (2.29)$$

where  $\sigma_{\beta_{OCEAN}}^2 = Var(\beta_{OCEAN})$ .

Mean terrestrial NEE over a region of land comprised of multiple grid boxes can be estimated as

$$\bar{F}_{Region} = \frac{1}{n} \sum_i F_i = \frac{1}{n} \sum_i (1 - \beta_{RESP})\overline{RESP}_i - \frac{1}{n} \sum_i (1 - \beta_{GPP})\overline{GPP}_i. \quad (2.30)$$

The corresponding uncertainty in regional monthly NEE is given by

$$\begin{aligned} \sigma_{F_{Region}}^2 &= \frac{1}{n^2} \sum_i \sum_j \overline{RESP}_i \times \overline{RESP}_j \times Cov(\beta_{RESP_i}, \beta_{RESP_j}) \\ &\quad + \frac{1}{n^2} \sum_i \sum_j \overline{GPP}_i \times \overline{GPP}_j \times Cov(\beta_{GPP_i}, \beta_{GPP_j}) \\ &\quad - \frac{2}{n^2} \sum_i \sum_j \overline{RESP}_i \times \overline{GPP}_j \times Cov(\beta_{RESP_i}, \beta_{GPP_j}), \end{aligned} \quad (2.31)$$



where  $i$  and  $j$  indicate the grid boxes within the region and  $n$  indicates the total number of grid boxes in the region. Similarly for air-sea exchange,

$$\bar{F}_{Region} = \frac{1}{n} \sum_i F_i = \frac{1}{n} \sum_i (1 - \beta_{OCEAN}) \overline{OCEAN}_i, \quad (2.32)$$

with the corresponding uncertainty in regional monthly air-sea exchange given by

$$\sigma_{F_{Region}}^2 = \frac{1}{n^2} \sum_i \sum_j \overline{OCEAN}_i \times \overline{OCEAN}_j \times Cov(\beta_{OCEAN_i}, \beta_{OCEAN_j}). \quad (2.33)$$

As noted in Lokupitiya et al. [2008], cross-correlations are added to smooth analyzed fluxes because the number of unknowns is large. It is also noted that the filter develops cross-correlations from the data during assimilation, an important advantage of MLEF, and that these cross-correlations need to be considered during interpretation of uncertainty estimates.

## Chapter 3

### MOIST SYNOPTIC TRANSPORT OF CO<sub>2</sub> ALONG THE MIDLATITUDE STORM TRACK

This first results section analyzes forward simulations of CO<sub>2</sub> from PCTM to understand and quantify how seasonal interaction of synoptic processes with ecosystem metabolism modulates seasonality in northern middle and polar latitudes. The concept of the large scale atmospheric circulation is first discussed from the perspective of streamfunctions on Eulerian and Isentropic coordinate systems in order to provide insight into how the mid-latitude circulation operates. Meridional CO<sub>2</sub> transport is then diagnosed in a moist isentropic framework and parsed to mean and eddy components to demonstrate the role of moist synoptic storms for meridional exchange of CO<sub>2</sub>. The significance of these storms for carbon cycle studies is demonstrated first through a case study that shows the covariance of transport by these storms with cloudiness and then through budget calculations of atmospheric CO<sub>2</sub> tendencies for a column of air around a latitude circle to show how eddy transport compares to fossil fuel emissions and terrestrial net ecosystem exchange. Finally, a conceptual diagram is provided to illustrate key interactions between baroclinic waves and atmospheric CO<sub>2</sub>.

#### 3.1 Introduction

It has long been recognized that seasonal and diurnal covariance between terrestrial ecosystem metabolism and fine-scale vertical transport in the atmosphere is a strong determinant of vertical structure in CO<sub>2</sub> mixing ratio (the CO<sub>2</sub> rectifier, Denning et al. [1995],

among others). Like the CO<sub>2</sub> rectifier, synoptic transport by baroclinic waves along the mid-latitude storm track involves strong vertical motion and is correlated with ecosystem metabolism because large-scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. Baroclinic wave activity is enhanced in winter when respiration and decomposition dominate ecosystem metabolism, and is suppressed in summer when photosynthesis dominates.

Unlike the CO<sub>2</sub> rectifier, tracer transport by slantwise convection in baroclinic waves involves a strong meridional component and is intimately tied to condensation and precipitation processes. A recent analysis by Pauluis et al. [2008] found that mass transport along  $\theta_e$  on baroclinic waves represents an important component of the atmospheric heat engine that operates between the equator and poles. This is also an important vehicle for tracer transport. Poleward transport by rising warm moist air follows a warm conveyor belt (WCB) above warm fronts, and is compensated by equatorward transport by sinking cold dry air following a dry air intrusion (DI) behind cold fronts [Cooper et al., 2001; Stohl, 2001].

Pollution transport by WCBs has received much attention in the scientific literature. Carlson [1981] presented empirical evidence indicating aerosol transport into the Arctic along  $\theta_e$ . More recently, models and aircraft measurements from intensive field campaigns (e.g., North American Regional Experiment) have helped create conceptual models that explain how WCBs transport gases such as O<sub>3</sub> and CO upward, poleward, and eastward away from continents and across ocean basins [e.g., Bethan et al., 1998; Cooper et al., 2001]. Strong pollution transport by WCBs is not surprising considering WCBs originate near the surface in polluted boundary layers and are responsible for most of a mid-latitude cyclones meridional energy transport [Eckhardt et al., 2004].

CO<sub>2</sub> transport by baroclinic waves was described in detail by Fung et al. [1983] and was included in the study by Tans et al. [1990] that established the concept of a terrestrial carbon sink but has not received much attention since [However, see Miyazaki et

al., 2008; Keppel-Aleks et al., 2010]. In Chapter 3, I attempt to refine our understanding of synoptic transport of CO<sub>2</sub> by including the contribution from moist processes embedded within stormtracks. Analysis of global transport simulations on the Lagrangian vertical coordinate  $\theta_e$  (described in Section 2) provides a means for describing how synchrony of baroclinic disturbances with seasonal ecosystem metabolism modulates seasonal variations of CO<sub>2</sub> in northern mid-latitudes. Analysis of meridional CO<sub>2</sub> transport is also provided on other vertical coordinates to facilitate interpretation of transport on  $\theta_e$ , which is a bit confusing and counter-intuitive at first.

Meridional transport of seasonally-varying CO<sub>2</sub> is analyzed in PCTM [Kawa et al., 2004; Parazoo et al., 2008], discussed in Section 2.1.1. PCTM produces realistic seasonal and synoptic variability, as well as interhemispheric mixing. “Eddy decomposition,” discussed in Chapter 2.5.2, is used to isolate transport signals associated with synoptic processes. There is extensive discussion of the influence of mean and eddy components of the atmospheric circulation on seasonal variations of CO<sub>2</sub>, including vertical and horizontal spatial structure, zonal and time average column integrated transport, and sensitivity to vertical coordinate. Analysis of covariance of baroclinic waves with ecosystem scale metabolism is carried out to determine the influence of meridional sloshing of CO<sub>2</sub> between middle latitudes, where the bulk of the terrestrial biosphere resides, and high latitudes, where vegetation is sparse. A conceptual diagram that summarizes key interactions between baroclinic waves and CO<sub>2</sub> mixing ratio is provided at the end of Chapter 3.

### 3.2 The Zonal Mean Circulation - Streamfunctions

Before analyzing atmospheric transport of CO<sub>2</sub> by mean and eddy circulations, the large scale atmospheric circulation operating between equator and pole is described. The global atmospheric circulation transports high energy air from equatorial regions to polar regions. This includes poleward flow of high energy parcels of air and equatorward flow of low energy parcels. Due to the turbulent nature of the atmosphere, parcel trajectories

vary widely, and it becomes necessary to describe the circulation in an averaged sense by computing zonal and temporal means over a sufficiently long period. Pauluis et al. [2008] give an excellent overview of global mean circulations calculated from the NCEP-NCAR Reanalysis on  $p$ ,  $\theta$ , and  $\theta_e$ . In this section, calculations of streamfunctions made by Pauluis et al. [2008] are repeated using GEOS4-1.25x1 to illustrate key points and determine whether circulations in GEOS4-1.25x1 are consistent with NCEP-NCAR.

### 3.2.1 *Streamfunction on Terrain Following Surfaces ( $\eta$ )*

The Eulerian mean circulation is a common technique for describing the atmospheric circulation. It is produced by averaging the flow on constant pressure, geopotential height, or terrain following coordinates. The circulation is conceptually and quantitatively similar in all cases. Since transport in PCTM is calculated on the terrain following coordinate  $\eta$ , calculate of the Eulerian mean streamfunction  $\Psi$  is performed on  $\eta$ , where  $\Psi_\eta$  is defined as:

$$\Psi_\eta(\eta_0, \phi) = \frac{1}{\tau} \int_0^\tau \int_0^{2\pi} \int_0^{p_s} H(\eta_0 - \eta) v a \cos \phi \frac{dp}{g} d\lambda dt. \quad (3.1)$$

Here,  $\eta_0$  is some reference value of  $\eta$ ,  $p$  is pressure,  $\phi$  is latitude,  $\tau$  is the time period over which the average is computed,  $\lambda$  is longitude,  $a$  is Earth's radius,  $v$  is the meridional velocity,  $g$  is the gravitational acceleration, and  $H(x)$  is the Heavyside function, with  $H(x) = 1$  for  $x \geq 0$  and  $H(x) = 0$  for  $x < 0$ .

Plots of  $\Psi$  effectively show zonally averaged vertical and meridional velocity, together in one diagram. Lines of constant  $\Psi$  cannot intersect the Earth's surface; if they did, that would imply a flow of air across the Earth's surface. Similarly, lines of constant  $\Psi$  cannot extend upward into space. Figure 3.1 shows the annual mean  $\Psi_\eta$  as a function of latitude for GEOS4-1.25x1 for 2005.

The annual mean circulation consists of a three cell structure in both hemispheres: the Hadley cell in the tropics, the Ferrel cell in mid-latitudes, and the polar cell at high latitudes. The Hadley and polar cells, with air parcels moving poleward at high altitude and equatorward at low altitude, are direct circulations that transport energy toward the

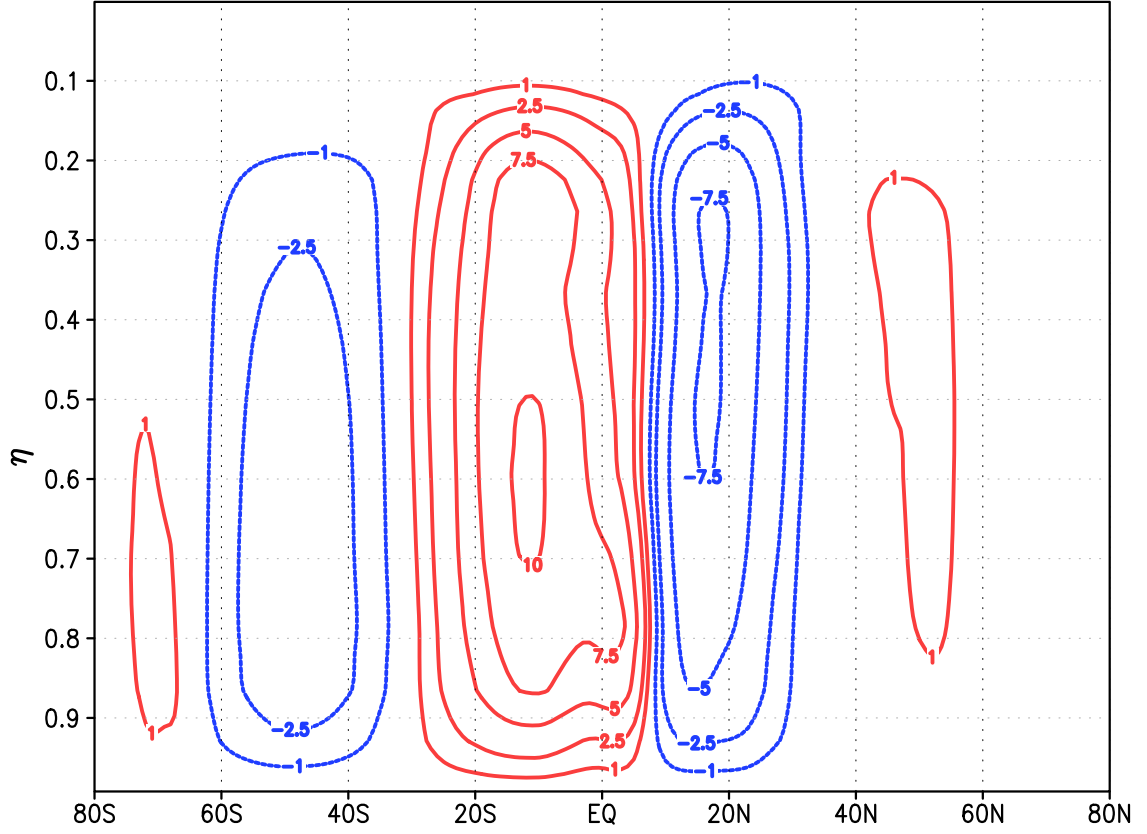


Figure 3.1: The global mean circulation computed on Eta surfaces from GEOS4-1.25x1 for 2005. Contour interval is  $2.5 \times 10^{10} \text{ kg s}^{-1}$ . Solid red contours are positive values of  $\Psi$  and correspond counterclockwise circulations, with northward flow at low levels and southward flow at high levels. Dashed blue contours are negative values of  $\Psi$  and correspond to clockwise circulations, with southward flow at low levels and northward flow at high levels.

poles. Deep rising motion occurs near the equator in the annual mean, with sinking motion on either side in the sub-tropics. The zonal band of deep rising motion shifts north and south seasonally with the Inter-Tropical Convergence Zone (ITCZ) and the position of the sun. The strongest vertical motion occurs near  $\eta = 0.5$ , which roughly corresponds to the middle troposphere, or 500 mb. The polar cells, in comparison, are very weak in the annual mean. The Ferrel Cell, an indirect circulation (energy transport is from low to high energy) in mid-latitudes, is strongest in the Southern Hemisphere, and corresponds to poleward flow near the surface and equatorward flow at high altitude that together transport energy towards the equator. Comparison to streamfunctions computed from NCEP-NCAR

Reanalysis [Pauluis et al., 2008] indicates that the general circulation is slightly stronger in GEOS4-1.25x1.

Synoptic scale eddies, which result from baroclinic instability in the unstable symmetric circulation, are not apparent in the Eulerian mean circulation. Eddies transport more energy toward the poles than the Ferrel cell transports toward the equator such that total energy transport in the mid-latitudes is poleward. The Ferrel cell is a response of the MMC to energy and momentum transport by synoptic eddies. Energy transport acts to cool the subtropics, inducing the sinking branch of the Ferrel cell, and warm higher latitudes, inducing the rising branch of the Ferrel cell. Eddy transport of angular momentum causes the Ferrel cell to adjust the meridional branches of the circulation; i.e., angular momentum transport out of the subtropics to middle latitudes is balanced in the MMC by Coriolis acceleration associated with an equatorward component of the Ferrel Cell.

### 3.2.2 *Streamfunction on $\theta$*

The mean circulation can also be described through averaging on isentropic surfaces. The streamfunction on  $\theta$ ,  $\Psi_\theta$ , is defined by

$$\Psi_\theta(\theta_0, \phi) = \frac{1}{\tau} \int_0^\tau \int_0^{2\pi} \int_0^{p_s} H(\theta_0 - \theta) v a \cos \phi \frac{dp}{g} d\lambda dt. \quad (3.2)$$

This is identical to Equation 3.1 except  $\eta_0$  has been replaced with  $\theta_0$ . Figure 3.2 shows the annual mean streamfunction on  $\theta$ , which looks very different from  $\eta$ . Direct circulations extend from the Equator all the way to the poles. Since the atmosphere is statically stable and  $\theta$  is monotonic with height, it can be said that the isentropic circulation is characterized by poleward flow at high altitudes and equatorward return flow near the Earth's surface. Radiative cooling causes  $\theta$  to decrease during poleward flow while sensible heating causes  $\theta$  to increase during the equatorward return flow. Strong diabatic heating by solar radiation, latent heat release, and sensible heating cause cross-isentropic flow and deep upward vertical motion in the tropics, while strong radiative cooling causing cross-isentropic flow and subsidence at high latitudes. Ferrel cells do not appear in the isentropic circulation.

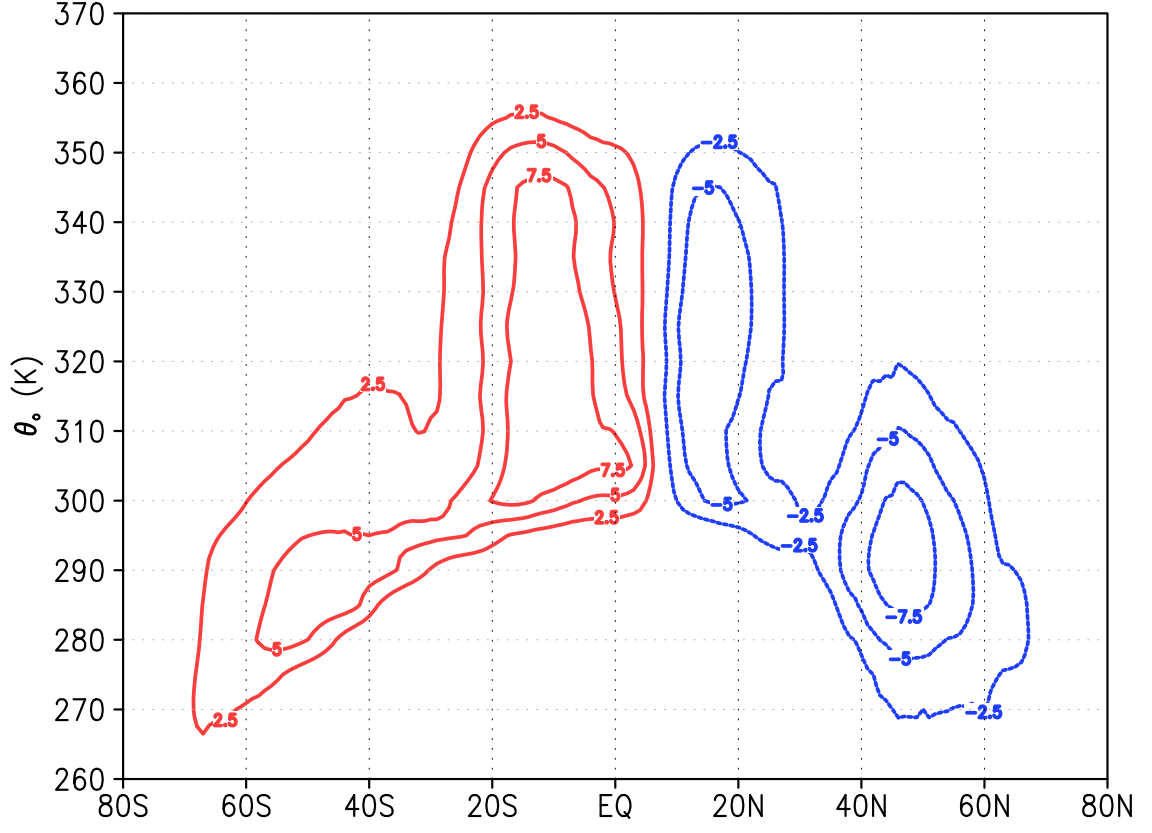


Figure 3.2: Same as Figure 3.1, except computed on potential temperature surfaces.

To understand why isentropic coordinates exhibit poleward transport at all latitudes while Eulerian coordinates exhibit poleward transport everywhere except mid-latitudes, we write the following equation for mass transport using eddy decomposition technique described in Chapter 2:

$$[mv] = [m][v] + [m^*v^*], \quad (3.3)$$

where  $m$  is the pseudo-density and  $v$  is the meridional wind. The first term on the RHS of Equation 3.3 is the product of the zonally averaged pseudo-density with the zonally averaged meridional wind. The second term arises from correlated fluctuations of the pseudo-density and meridional wind. The first term represent mass transport by the mean flow, which is supplemented by eddy mass flux when variations of meridional wind are correlated strongly enough with variations of pseudo-density.



Transport by the zonally averaged flow,  $[m][v]$ , dominates the circulation in the tropics where variations of mass are small because of weak temperature gradients. This leads to small values of  $m^*$  on both  $\eta$  and  $\theta$  and the eddy mass flux is trivial in the tropics. In the mid-latitudes, however, the zonally averaged meridional wind is small because of geostrophic flow. Eddy mass flux dominates mass transport in midlatitudes because stationary and baroclinic waves force large variations of mass and hence  $m^*$  becomes large. Variations of mass on  $\eta$  are still too small, however, to stand out in Eulerian averaged circulations, and consequently the zonally averaged flow dominates the Eulerian circulation as seen in Figure 3.1. Variations of mass are large on isentropic surface such that  $[m^*v^*]$  takes over from the zonally averaged flow. Variations of mass are much larger on  $\theta$  than  $\eta$  because, as discussed in Chapter 2, isentropic circulations are more indicative of parcel trajectories, where potential temperature of an air parcel is approximately conserved in the absence of condensation.

### 3.2.3 *Streamfunction on $\theta_e$*

It is clear from Figure 3.2 that, while energy transport is poleward in both hemispheres, the isentropic flow is only approximately conserved. In particular, the Northern Hemisphere mid-latitudes experience fairly severe cross isentropic flow near 40°N due to latent heat release by condensation and precipitation along stormtracks. This problem is alleviated somewhat by accounting for latent heat release associated with moist air parcels.  $\theta_e$ , defined in Chapter 2.5.3, is conserved for reversible adiabatic transformations. Because  $\theta_e$  includes a contribution from the latent heat content of water vapor, phase transitions from vapor to liquid are also conserved.

The streamfunction on  $\theta_e$ ,  $\Psi_{\theta_e}$ , is defined as:

$$\Psi_{\theta_e}(\theta_{e0}, \phi) = \frac{1}{\tau} \int_0^\tau \int_0^{2\pi} \int_0^{p_s} H(\theta_{e0} - \theta_e) v a \cos \phi \frac{dp}{g} d\lambda dt. \quad (3.4)$$

This is identical to Equation 3.2 except  $\theta_0$  has been replaced with  $\theta_{e0}$ . Figure 3.3 shows the annual mean streamfunction on  $\theta_e$ , which looks similar to  $\theta$  except for a few key differences.

First, and perhaps most important, calculations of Pauluis et al [2008] show that total mass

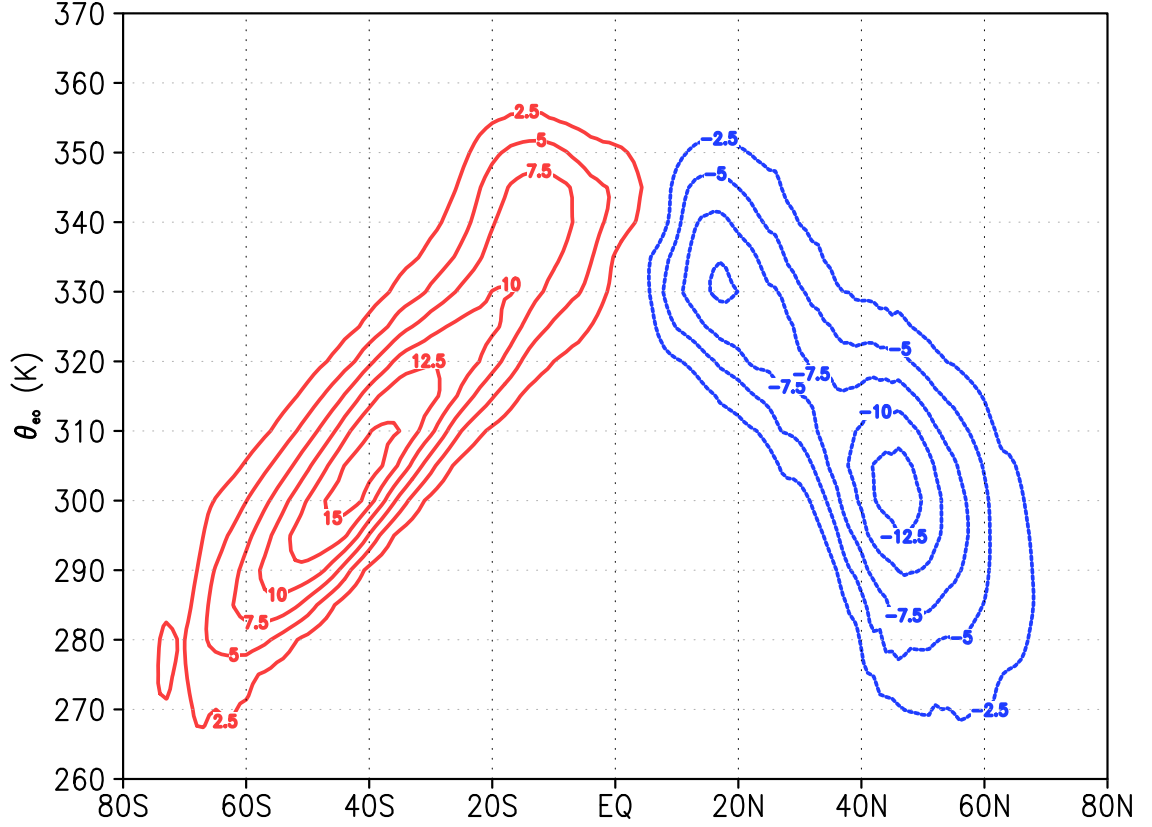


Figure 3.3: Same as Figure 3.2, except computed on equivalent potential temperature surfaces.

transport on  $\theta_e$  is approximately twice that on  $\theta$ . The authors use the joint distribution of mass transport on  $\theta$  and  $\theta_e$  to show that the additional mass transport corresponds to a low-level poleward flow of warm moist air with high  $\theta_e$  indicative of air parcels that are nearly convectively unstable. The circulation is strongest in mid-latitudes along the stormtrack where synoptic scale eddies advect low-level warm moist air upward and poleward before subsiding over the poles. Accounting for latent heating in tropical and midlatitude regions of strong precipitation has eliminated much of the cross-isentropic flow. It is also worth noting that the potential temperature of the surface equatorward return flow increases at a faster rate on  $\theta_e$  because of surface evaporation.

### 3.3 Energy Transport

The atmospheric heat engine operating between the equator and poles is an important vehicle for tracer transport. If streamfunctions provide insight into what the general circulation looks like, eddy decomposition describes how zonally symmetric and asymmetric components of the general circulation move energy and tracer around. Interpretation of the general circulation is also sensitive to the vertical coordinate; as will be shown later, eddy decomposition and hence tracer transport by difference components of the circulation is also sensitive to vertical coordinate. Such sensitivity will ultimately aid perception of important mechanisms that govern meridional CO<sub>2</sub> transport.

The streamfunction calculated on  $\theta_e$  suggests that the mid-latitude circulation is driven strongly by asymmetries from the zonally symmetric circulation (i.e., eddies). To illustrate the relative importance of heat transport by eddies compared to the zonally symmetric circulation, eddy decomposition is used to parse moist static energy ( $h = c_p T + gz + Lq$ , where  $T$  is temperature,  $g$  is gravity,  $z$  is altitude, and  $L$  is latent heat) transport to mean, transient, and stationary terms as discussed in Chapter 2.5.2 according to the expression:

$$\begin{aligned}
 2\pi a \cos \phi \int_{\zeta_{sfc}}^{\zeta_{top}} [\overline{m_\zeta v_\zeta h_\zeta}] \frac{d\zeta}{g} &= 2\pi a \cos \phi \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta} \right] \left[ \overline{h_\zeta} \right] \frac{d\zeta}{g} \\
 &+ 2\pi a \cos \phi \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta^* h_\zeta^*} \right] \frac{d\zeta}{g} \\
 &+ 2\pi a \cos \phi \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta'} \right] \left[ \overline{h'_\zeta} \right] \frac{d\zeta}{g} \\
 &+ 2\pi a \cos \phi \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta'^* h_\zeta'^*} \right] \frac{d\zeta}{g}. \tag{3.5}
 \end{aligned}$$

Equation 3.5 gives units of energy transport in PW. As mentioned at the end of Section 2.5.2, the third and fourth terms on the RHS of Equation 3.5 will collectively be referred to as transport by transient eddies.

Moist static energy is essentially a measure of the energy within an air parcel due to potential energy, sensible heat or internal energy, and latent heating. Moist static energy

transport is calculated on  $\eta$ ,  $\theta$ , and  $\theta_e$ , vertically integrated according to Equation 3.5, and plotted in Figure 3.4. Positive values indicate northward transport and negative values

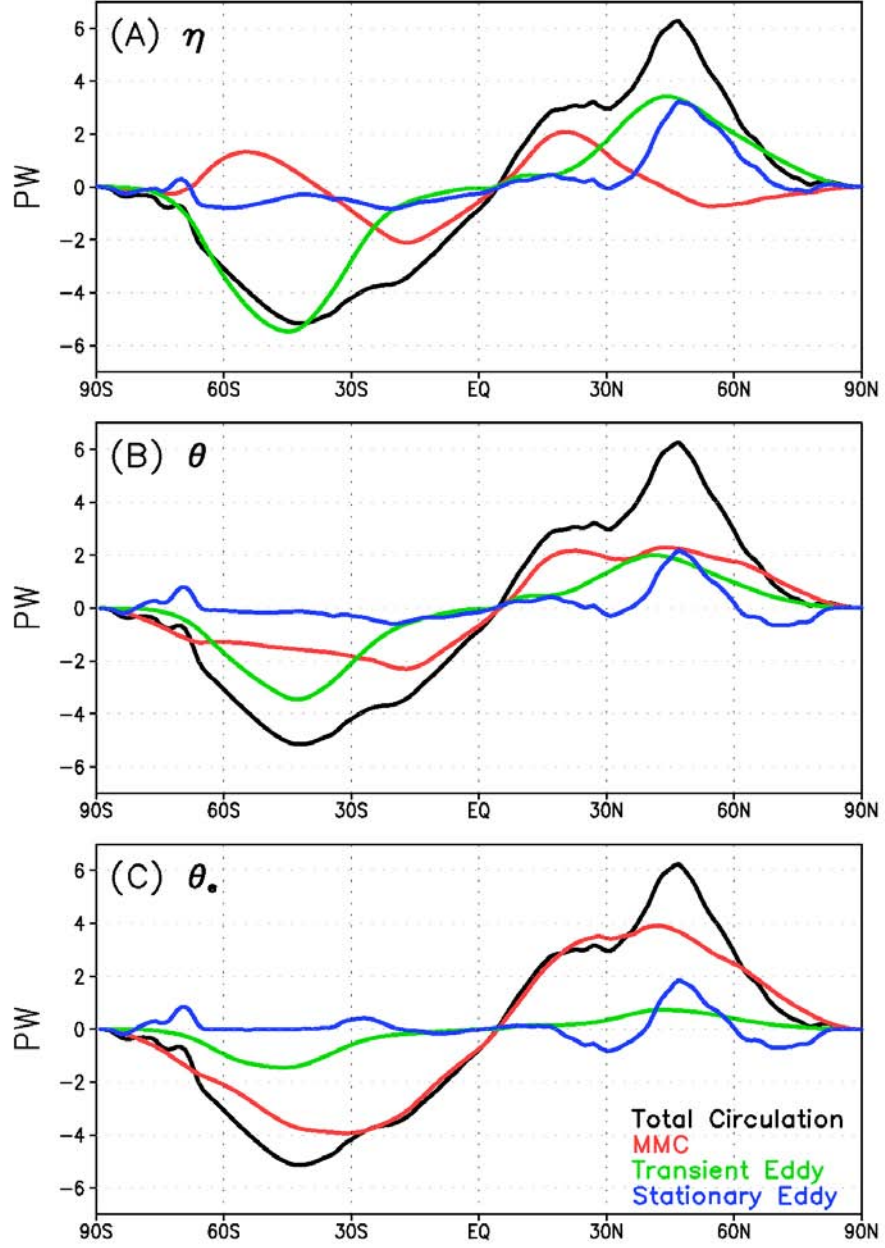


Figure 3.4: Annual mean transport of moist static energy by various components of the mid-latitude circulation. Energy has units of PW month<sup>-1</sup>

indicate southward transport. Energy transport peaks in middle latitudes in both hemispheres near 45°N and crosses zero near the equator, indicating the atmospheric heat engine

is moving energy away from low latitudes and into high latitudes.

Interpretation of energy transport by mean and eddy terms is sensitive to vertical coordinate and somewhat misleading. When calculated on  $\eta$  (Figure 3.4A), the MMC (mean circulation) dominates transport out of the subtropics through the Hadley Cell. Transport by the MMC at lower latitudes is actually stronger during individual seasons, but this transport largely cancels in the annual mean. Eddies dominate transport in the annual mean in mid-latitudes, with transient eddies important in both hemispheres. While transient and stationary waves contribute equally in the NH, eddies are dominated by transient eddies in the SH due to the large percentage of ocean coverage.

Eddy transport is much weaker on  $\theta$  and almost non-existent on  $\theta_e$ . In general, transport by the MMC is dominant in the isentropic framework. These results are misleading, however, because  $h$  and  $\theta_e$  are closely related by  $c_p T d \ln \theta_e \approx dh$ , where  $T$  is the temperature of the condensation level, and are thus very strongly correlated in the lower troposphere. Because  $\theta_e$  and  $h$  within air parcels are both conserved during adiabatic vertical motion, and synoptic eddies tend to follow the path of isentropes during moist ascent along stormtracks, zonal and temporal perturbations of  $h$  are approximately zero when calculated on  $\theta_e$  such that “eddy” transport of  $h$  becomes small in the isentropic framework. It is therefore more convenient to use Eulerian coordinates such as  $\eta$  when describing energy or heat transport by baroclinic waves.

### 3.4 Seasonal Covariance of Baroclinicity and Ecosystem Metabolism

The motivation behind including eddy decomposition of energy transport in this study is to show how seasonal variations of meridional transport of energy and heat correspond with plant growth and decay. To this end, the seasonal distribution of eddy transport of  $h$ , as calculated on  $\eta$ , is plotted in black in Figure 3.5. Transport of  $h$  is much more vigorous during winter (DJF) than summer (JJA) in the NH due to the strong meridional temperature gradient. By comparison, there is relatively little seasonal change in eddy transport

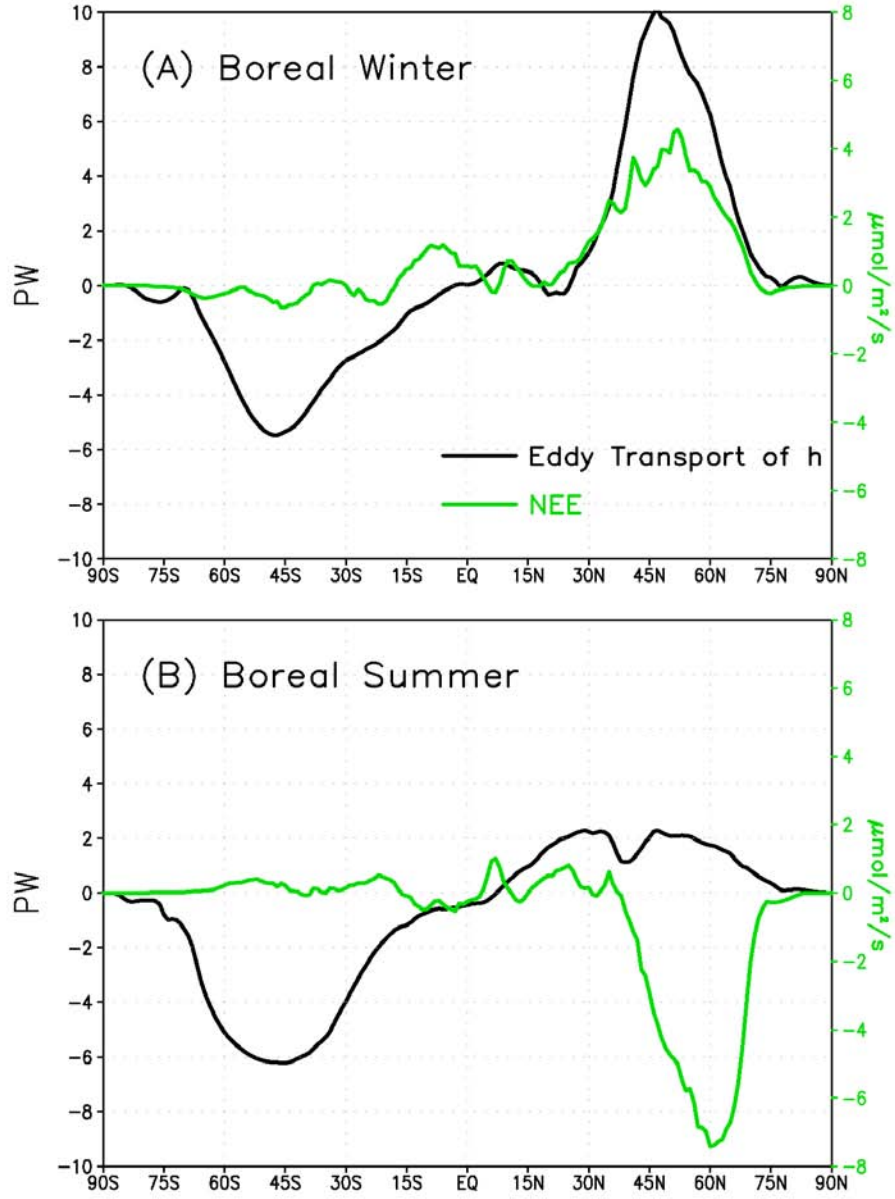


Figure 3.5: Seasonal covariance of moist static energy transport ( $h$ , black line) and zonally-monthly averaged terrestrial net ecosystem exchange (NEE, green line) for boreal winter (A, Dec-Jan-Feb) and boreal summer (B, Jun-Jul-Aug). NEE is plotted as a zonal mean in units of  $\mu\text{mol m}^{-2} \text{s}^{-1}$  on the right y-axis.

in the SH. The green line in Figure 3.5 represents zonally averaged NEE as simulated by SiB. NEE has the largest seasonality in northern mid-latitudes where the bulk of global land area resides and ecosystem metabolism is limited by seasonal variations in solar radiation. During DJF there is net respiration of  $\text{CO}_2$  to the atmosphere due to microbial

decomposition of plant litter. During JJA there is net uptake of  $\text{CO}_2$  from the atmosphere due to photosynthetic assimilation. NEE is much weaker in the SH where total land area and biomass are reduced. While the tropics are the most biologically active part of the planet, seasonality of zonally averaged NEE is small due to cancellation of respiration and GPP. Synoptic transport of  $h$  is correlated with ecosystem metabolism in the NH because large-scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. Baroclinic wave activity, as depicted by eddy energy transport in Figure 3.5, is enhanced in winter when respiration and decomposition dominates ecosystem metabolism, and is suppressed in summer when photosynthesis dominates.

### 3.5 Meridional Transport of $\text{CO}_2$ by Mean and Eddy Circulations

This section provides analysis of the influence of covariance of heat transport and surface fluxes of  $\text{CO}_2$  on meridional transport of  $\text{CO}_2$ . Eddy decomposition is used to assess how transport by eddy and mean circulations contribute separately to  $\text{CO}_2$  seasonality in northern middle and polar latitudes. Total average meridional transport at an arbitrary level  $\zeta$  is calculated according to the expression

$$2\pi a \cos \phi \frac{M_c}{M_{air}} \overline{[m_\zeta v_\zeta C_\zeta]} \frac{d\zeta}{g}, \quad (3.6)$$

giving units of  $\text{PgC K}^{-1} \text{ month}^{-1}$ , where  $M_c$  is the molecular weight of carbon and  $M_{air}$  is the molecular weight of air. Mean and eddy components of transport are calculated the same way with total transport decomposed according to Equation 2.16 in Chapter 2. Equation 3.6 is used to show vertical distributions of  $\text{CO}_2$  transport. The zonal mean is removed to discuss spatial distributions of transport (this analysis assumes zonal wind is zero). The vertically integrated distributions of eddy and mean transport is also analyzed. Column integrated meridional transport is calculated in units of  $\text{PgC month}^{-1}$ . The equation for

CO<sub>2</sub> transport is similar to that discussed above for  $h$  transport:

$$\begin{aligned}
2\pi a \cos \phi \frac{M_c}{M_{air}} \int_{\zeta_{sfc}}^{\zeta_{top}} \overline{[m_\zeta v_\zeta C_\zeta]} \frac{d\zeta}{g} &= 2\pi a \cos \phi \frac{M_c}{M_{air}} \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta} \right] \left[ \overline{C_\zeta} \right] \frac{d\zeta}{g} \\
&+ 2\pi a \cos \phi \frac{M_c}{M_{air}} \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta}^* \overline{C_\zeta}^* \right] \frac{d\zeta}{g} \\
&+ 2\pi a \cos \phi \frac{M_c}{M_{air}} \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta}' \right] \left[ \overline{C_\zeta}' \right] \frac{d\zeta}{g} \\
&+ 2\pi a \cos \phi \frac{M_c}{M_{air}} \int_{\zeta_{sfc}}^{\zeta_{top}} \left[ \overline{(mv)_\zeta}'^* \overline{C_\zeta}'^* \right] \frac{d\zeta}{g}. \quad (3.7)
\end{aligned}$$

Again, the third and fourth terms on the RHS of Equation 3.7 are referred to as transient eddies.

### 3.5.1 CO<sub>2</sub> Transport by the MMC

Seasonal variations in the vertical distribution of CO<sub>2</sub> transport by the MMC, as calculated on  $\theta_e$ , are summarized in Figure 3.6. Zonal mean poleward and equatorward transport are in opposite directions within a vertical column throughout the year in northern latitudes. Transport is strongest during winter (Figure 3.6B) and tends to dominate the annual mean signal (Figure 3.6A). Transport during summer (Figure 3.6C) resembles the winter pattern except that transport throughout the column is weaker and occurs at higher  $\theta_e$  due to warmer temperatures.

The pattern of seasonal transport by the MMC is explained to some degree by the mean distribution of atmospheric CO<sub>2</sub> and meridional mass fluxes, which are plotted together in Figure 3.7. During boreal winter (Figure 3.7B), for example, the northern mid-latitudes are a source region for atmospheric CO<sub>2</sub> through ecosystem respiration (see Figure 3.5A) and fossil fuel emission such that CO<sub>2</sub> accumulates near the surface (dark red shading) and decreases vertically into the upper atmosphere (shading tapers from dark red near 50°N and 260K to yellow and blue upward and outward; note that red values indicate CO<sub>2</sub> that is high relative to the column mean, i.e. positive CO<sub>2</sub> anomaly, and blue shading indicates low CO<sub>2</sub> relative to the column mean). It should be noted that  $\theta_e$  is not continuous from equator to pole because of surface evaporation and sensible heating of air parcels



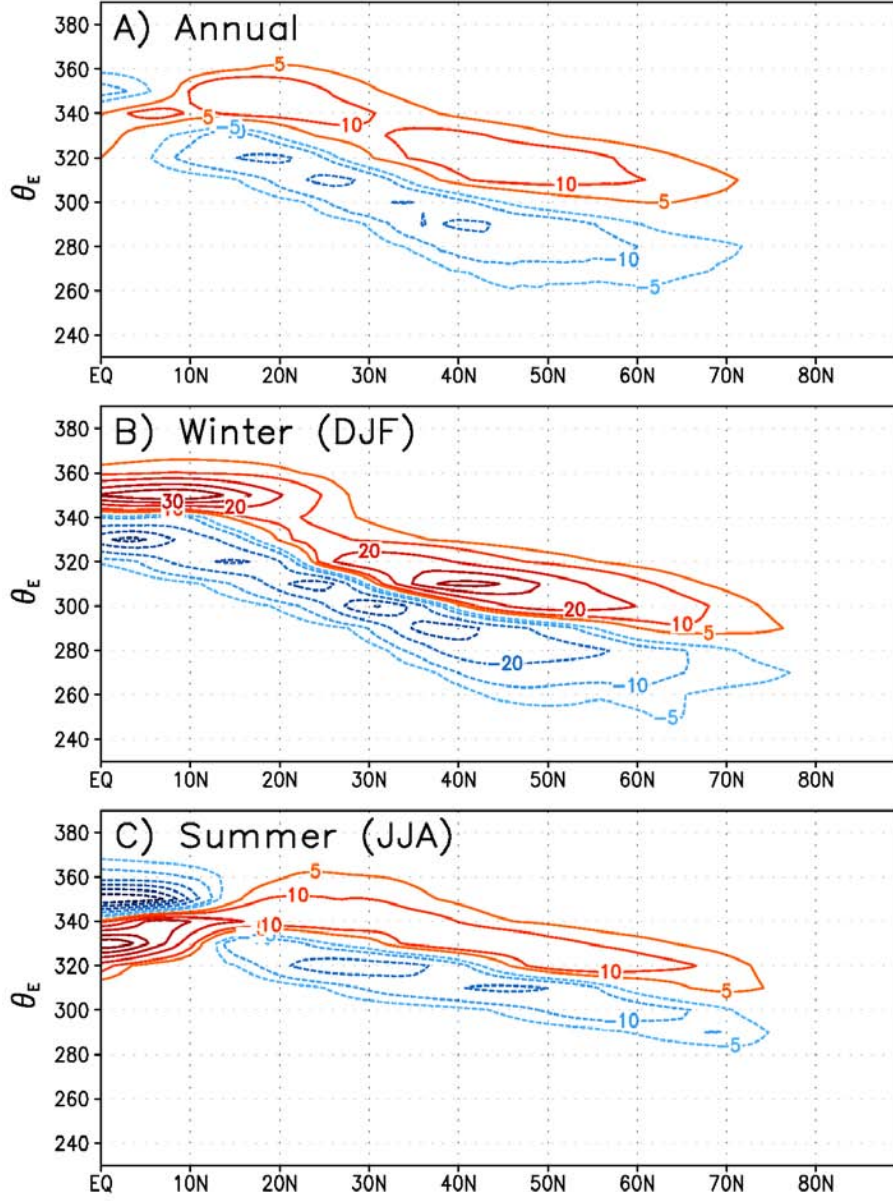
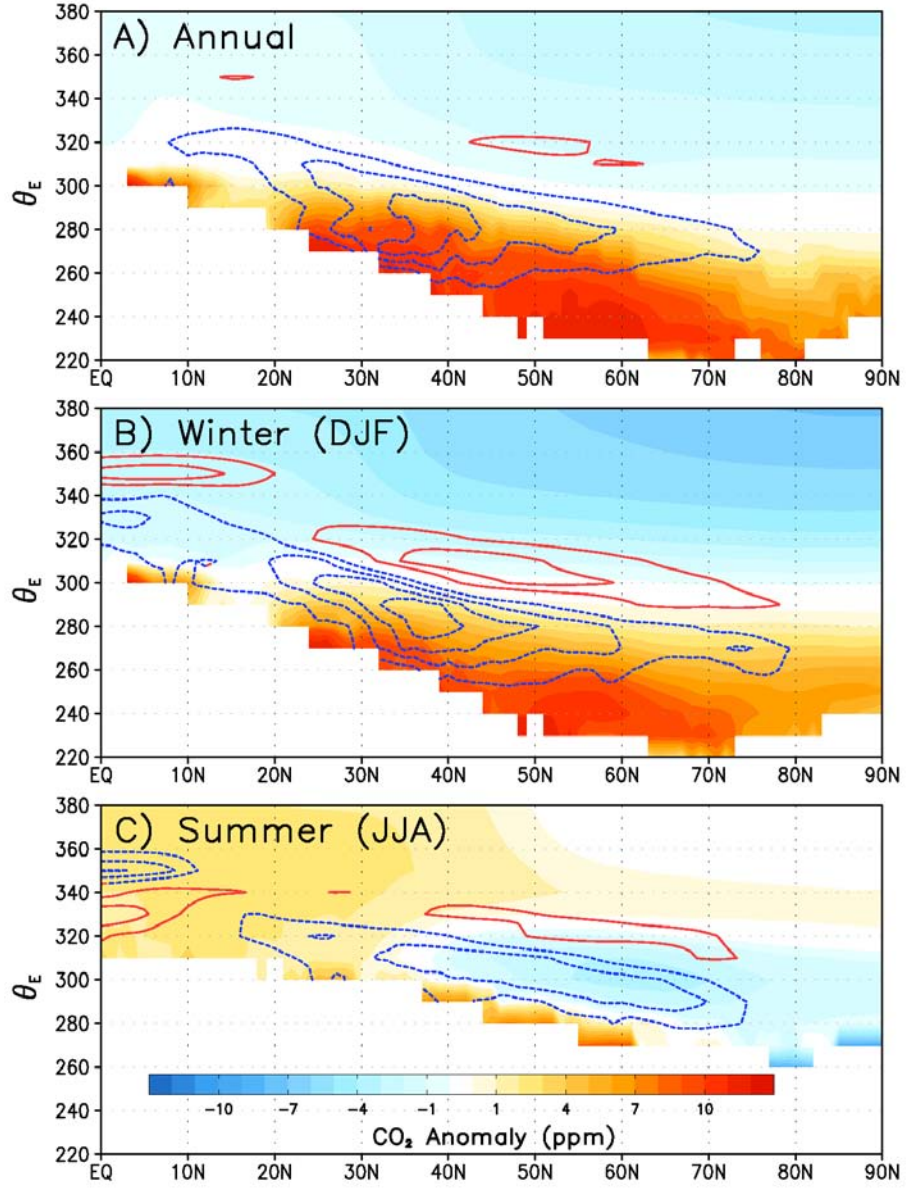


Figure 3.6: Diagnosis of vertical distribution of atmospheric CO<sub>2</sub> transport by the MMC on  $\theta_e$  in the annual mean (A) and during boreal winter (B) and summer (C). Red (blue) contours indicate poleward (equatorward) transport. Units are PgC K<sup>-1</sup> month<sup>-1</sup>

moving equatorward along the surface [Pauluis et al., 2008]. Also, because  $\theta_e$  slopes upward and poleward toward cold air, there is net gain in altitude following  $\theta_e$  poleward (with loss in altitude due to radiational cooling of  $1^\circ\text{C day}^{-1}$ ).

Red and blue contours in Figure 3.7 indicate the mean poleward and equatorward meridional mass flux, respectively. According to Figure 3.7B, transport by the MMC during



by the MMC during boreal winter is indicated by red contours in Figure 3.6B. Equatorward transport of high  $\text{CO}_2$  is indicated by blue contours.

Mean mass circulation during boreal summer occurs in the same sense as winter transport (see red and blue contours in Figure 3.7C); during summer, however, the equator-to-pole temperature gradient is strongly reduced and the strength of the atmospheric circulation weakens. At the same time, warmer temperatures and increased precipitation create a sink region for atmospheric  $\text{CO}_2$  in the mid-latitudes through enhanced photosynthetic uptake (see Figure 3.5B), causing reversal in the large-scale vertical  $\text{CO}_2$  gradient north of  $\sim 30^\circ\text{N}$  (except near the surface because of deep convective uplifting of low  $\text{CO}_2$  [see Miyazaki et al., 2008]).

### 3.5.2 $\text{CO}_2$ Transport by Synoptic Eddies

Unlike the mean circulation, which essentially is a “statistical circulation” determined by zonal and temporal averaging, eddy circulations are more physical and have vertical and horizontal structure that can be witnessed and experienced in real life. This section is therefore organized into subsections based on the spatial structure being characterized. Other subsections, including “Cyclonic Case Study” and “Significance of Eddy Transport” are aimed at clarifying mechanisms associated with eddy transport.

#### *Vertical distribution of Eddy Transport*

Transport by eddy circulations is weaker than mean transport by several orders of magnitude, but poleward through most of the troposphere in northern latitudes (north of  $30^\circ\text{N}$ ) in the annual mean (Figure 3.8A). Poleward transport by synoptic eddies is strongest during boreal winter (Figure 3.8B) with peak transport near  $45^\circ\text{N}$  near  $\theta_e = 300\text{ K}$ . During boreal summer (Figure 3.8C) there is convergence of poleward and equatorward eddy transport near  $60^\circ\text{N}$  due to net biological uptake of  $\text{CO}_2$ . Opposing directions of transport in the column south of  $30^\circ\text{N}$ , most noticeable during boreal winter, are a result of transport

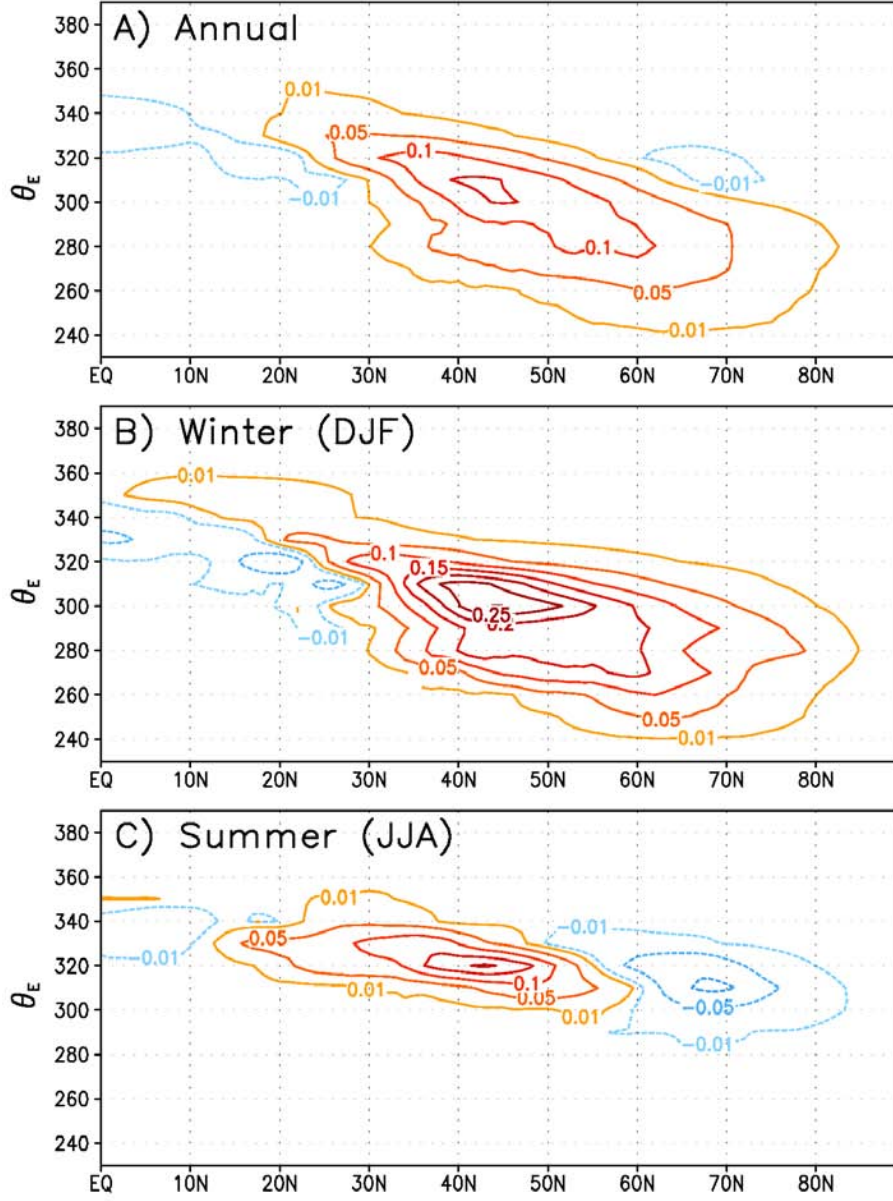


Figure 3.8: Diagnosis of vertical distribution of atmospheric CO<sub>2</sub> transport by synoptic eddies (stationary and transient waves combined) on  $\theta_e$  in the annual mean (A) and during boreal winter (B) and summer (C). Red (blue) contours indicate poleward (equatorward) transport. Units are  $\text{PgC K}^{-1} \text{ month}^{-1}$

by monsoonal circulations in tropical latitudes, which are manifested in stationary waves due to deviations from the zonal mean circulation.

Eddy transport is forced through covariance between perturbations (temporal and zonal) to the mean circulation and to mean CO<sub>2</sub>. Positive (negative) perturbations to the

mean circulation represent anomalous poleward (equatorward) mass flux, with poleward flow of warm moist air ( $\theta_e > \sim 280\text{K}$ ) representing WCBs and equatorward flow of cold dry air ( $\theta_e < \sim 280\text{K}$ ) representing DIs. Anomalous mass flux along  $\theta_e$  causes corresponding perturbations to mean  $\text{CO}_2$  (with respect to latitude). Because  $\text{CO}_2$  decreases northward on  $\theta_e$  throughout much of the troposphere in the annual mean (Figure 3.7A), anomalous poleward mass flux causes positive perturbations of  $\text{CO}_2$ . Poleward transport in the annual mean and during boreal winter is a result of positive covariance of anomalous mass flux and  $\text{CO}_2$ . Equatorward transport north of  $60^\circ\text{N}$  during summer, on the other hand, is a result of negative covariance. Negative covariance and thus equatorward transport result primarily from changes in the sign of meridional  $\text{CO}_2$  gradients.

### *Cyclonic Case Study*

Examination of a particular case illustrates the eddy transport mechanism. A composite surface map (Figure 3.9A) shows an extratropical cyclone that passed over the continental United States on 13 January 2005. A low-pressure system is centered over Lake Michigan, with corresponding warm and stationary fronts (solid red line and red-blue line, respectively) passing into New England and a cold front (solid blue line) over the Great Plains extending to Mexico. In the warm sector of the cyclone (southeast United States) there is poleward flow (depicted by wind vectors at surface stations pointing north) of warm moist air (temperatures from  $10\text{-}20^\circ\text{C}$ , high dewpoints). Following the classic mid-latitude cyclone [Cooper et al., 2001; Stohl et al., 2001], this air follows a WCB, rises above the surface warm front, and spreads into Canada and the Atlantic. In the cold sector is equatorward flow (wind vectors point south) of cold dry air out of Canada (temperatures from  $-10\text{-}0^\circ\text{C}$ , low dewpoints), which descends behind the cold front.

Figure 3.9B shows the cyclone influence on meridional  $\text{CO}_2$  transport at  $\theta_e = 300\text{K}$ . Since poleward moving air originates at subtropical latitudes where air is warmer and moister than equatorward moving air,  $\theta_e$  intersects the surface at higher latitude in the



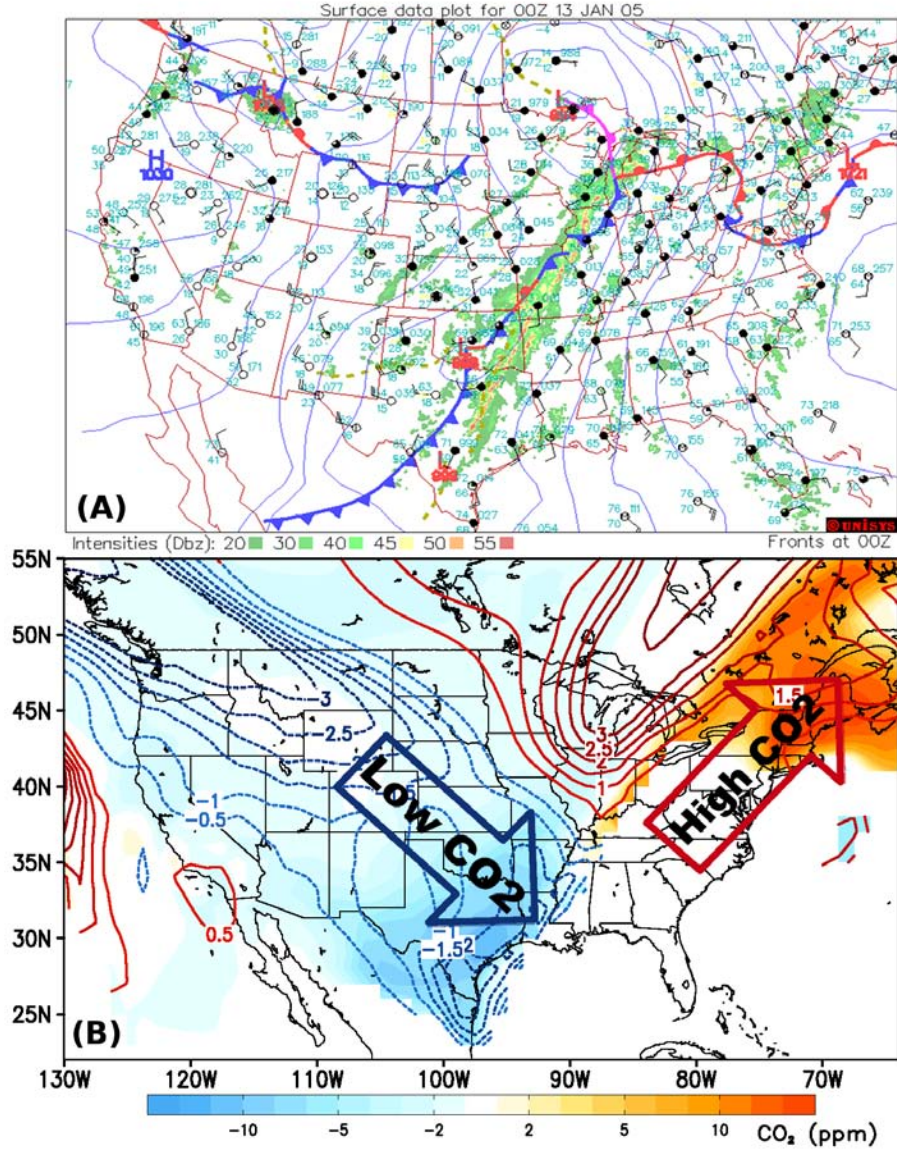


Figure 3.9: Case study from 13 Jan 2005 showing an example of eddy  $\text{CO}_2$  transport associated with a typical mid-latitude cyclone. The surface composite map in (A) is reproduced with permission of Unisys Corporation (c) 2011. Model output of anomalous  $\text{CO}_2$  (ppm, shaded) and mass flux ( $1 \times 10^{-4} \text{ kg m}^{-1} \text{ K}^{-1}$ , contour) along  $\theta_e = 300 \text{ K}$  is plotted in (B).

warm sector ( $\sim 42^\circ \text{N}$ ) than the cold sector ( $\sim 27^\circ \text{N}$ ). The primary pattern is for equatorward mass flux of low  $\text{CO}_2$  air in the cold sector and poleward flux of high  $\text{CO}_2$  air in the warm sector. As equatorward moving air sinks it crosses many pressure levels and creates large anomalous mass fluxes. Cold sector transport is dry;  $q$  is zero on  $\theta_e$  and equatorward transport is equal on  $\theta$  (not shown) and  $\theta_e$ . Similarly, poleward moving air

crosses many pressure levels during moist ascent, with equally large anomalous mass fluxes. *Warm sector transport is moist and  $q$  is nonzero; since  $\theta_e$  conserves energy in the presence of latent heating by condensation while  $\theta$  gains energy,  $\theta_e$  is more indicative of parcel trajectories along moist conveyors.* This portion of eddy transport is parsed to the mean term when diagnosed on  $\theta$ , reducing eddy  $\text{CO}_2$  transport by nearly half (see discussion below). Because covariance of  $\text{CO}_2$  and mass flux is positive in the warm and cold sectors in the time and zonal mean at  $\theta_e = 300\text{K}$ , net eddy transport over North America is poleward. Eddy transport over Eurasia, as well as along other  $\theta_e$  surfaces, has different latitudinal origins, with corresponding transport dependent on the latitudinal  $\text{CO}_2$  gradient and the amount of convective instability of  $\theta_e$ , but follows the same principle.

#### *Cloudiness Associated with Eddy Transport*

Migrating cyclones responsible for  $\text{CO}_2$  advection also cause a great deal of precipitation and cloudiness. In the case illustrated above, radar (color filled stippling in Figure 3.9B) indicates observed precipitation and filled surface stations cloudiness. Radar shows that precipitation during this day occurs primarily near cold and warm fronts. Surface station analysis indicates that moist conveyors in the warm sector are also very cloudy. *Covariance among condensing air (i.e. clouds), precipitation,  $\text{CO}_2$ , and heat transport means that  $\text{CO}_2$  transport along the east side of baroclinic waves will be hidden from satellite observing systems.* Transport along the DI, however, is dry and sunny (open surface stations) and detectable by satellites.

#### *Significance of Eddy Transport*

While covariance of zonal and temporal anomalies of mass flux and  $\text{CO}_2$  are positive and thus poleward on average for stationary and transient waves, there are many instances of net equatorward transport of  $\text{CO}_2$  by eddies, as seen for example over the Great Lakes in

Figure 3.9B. It is a useful exercise to calculate the correlation coefficient of the data series used to calculate the eddy terms to get a sense at the noisiness of correlations of mass flux and  $\text{CO}_2$ . A plot of the annual mean correlation coefficient for transient and stationary waves is shown in Figure 3.10.

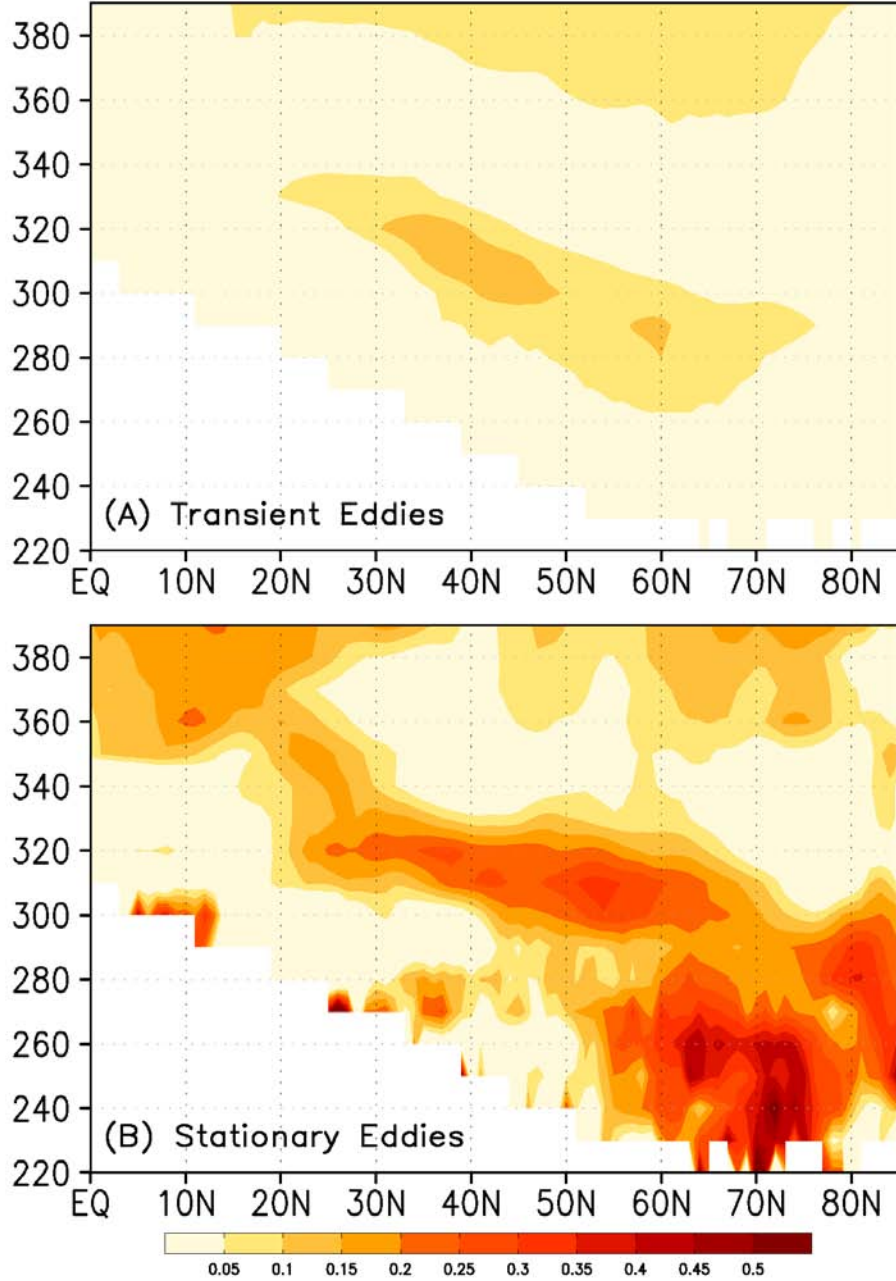


Figure 3.10: Annual mean correlation coefficient of mass flux and  $\text{CO}_2$  deviations for (A) transient eddies and (B) stationary eddies (bottom plot), plotted as a function of  $\theta_e$ . Colorbar is the correlation coefficient.



The average correlation is positive throughout the northern atmosphere, consistent with predominantly poleward transport in Figure 3.8A. Stationary waves exhibit correlations between mass flux and CO<sub>2</sub> deviations of about 0.25 in northern middle latitudes and values close to 0.4-0.5 at higher northern latitudes. Correlations within transient waves peak near 30-50°N and are closer to 0.1. Higher correlations in stationary waves suggest a stronger and more systematic correspondence between mass flux and CO<sub>2</sub> deviations, with net flow dominated by poleward transport of CO<sub>2</sub>. Low correlations in transient waves in the annual mean suggest that transport by traveling cyclones is noisy and mixed between poleward and equatorward transport. Eddy correlations have some seasonality with maximum values attained during boreal winter, but are not significantly different in shape or magnitude from the annual mean (not shown).

#### *Horizontal Distribution of Eddy Transport*

The previous discussion, including case studies, have facilitated understanding of some of the key underlying mechanisms responsible for patterns of eddy CO<sub>2</sub> transport in the NH. The next step is discussion of the longitudinal distribution of time-averaged column integrated eddy transport. This analysis will provide an idea of the horizontal extent of zonally averaged eddy terms discussed in Figure 3.8. The spatial distribution of column integrated stationary and transient waves is plotted in Figure 3.11.

Patterns of eddy CO<sub>2</sub> transport emerge over spatial scales of 1000's of km's. Comparison to a climatology of WCBs presented in Figure 3 of Eckhard et al. [2004] suggests that dominant patterns of combined poleward CO<sub>2</sub> transport by stationary and transient eddies (Figure 3.11A) are coincident with climatologically favorable regions for WCBs. Like WCBs, CO<sub>2</sub> transport tends to follow a northward and eastward track originating in subtropical latitudes ( $\sim 20\text{-}30^\circ\text{N}$ ), peaking in middle latitudes ( $\sim 40\text{-}50^\circ\text{N}$ ), and terminating at the southern base of the polar vortex ( $\sim 70^\circ\text{N}$ ). The most favorable regions tend to originate over land (e.g., southern portion of the United States and Tropical Asia) and terminate over

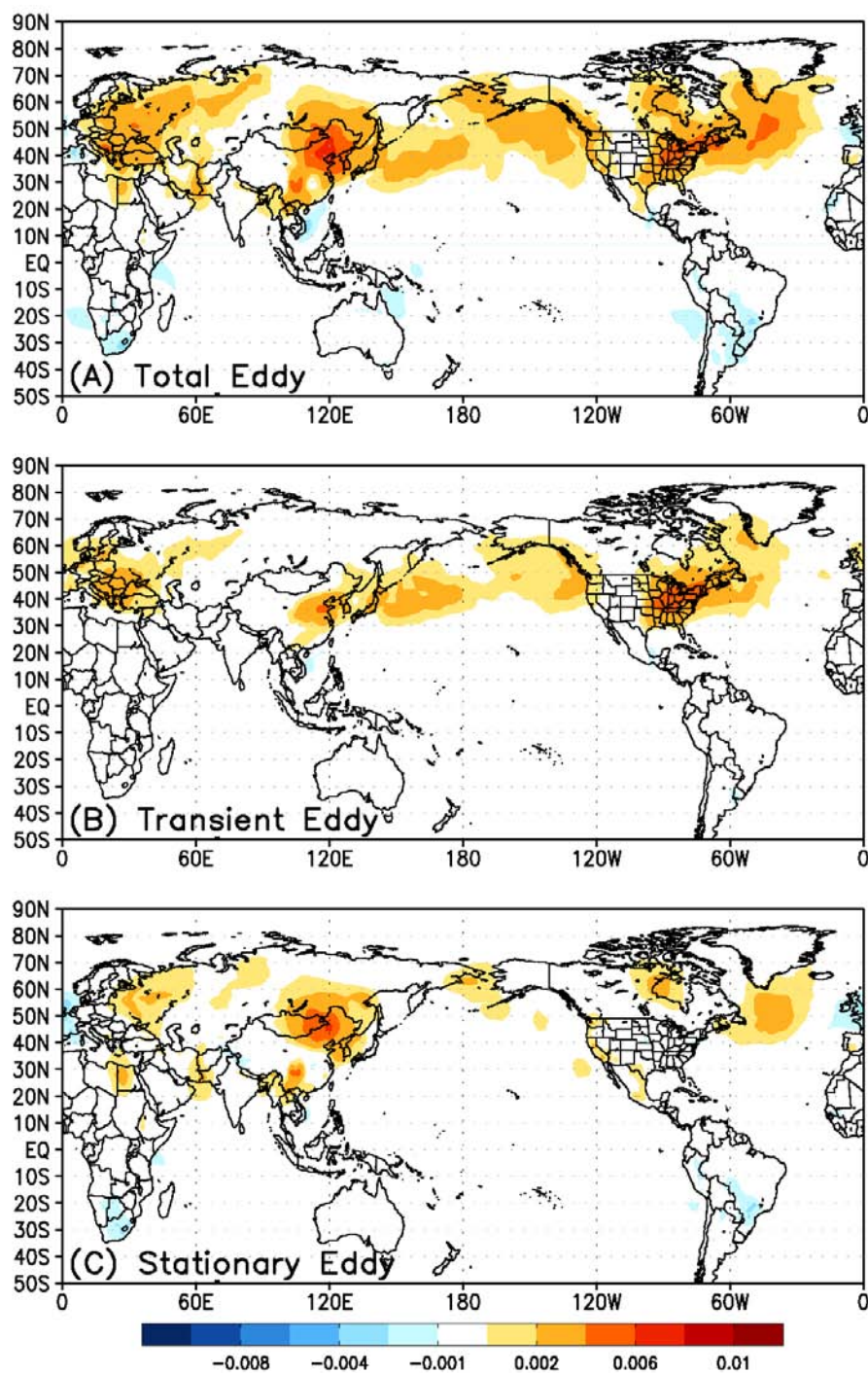


Figure 3.11: The annual mean spatial distribution of column integrated transport by (A) transient and stationary eddies, (B) transient eddies and (C) stationary eddies. The colorbar for each plot is shown at the bottom in units of  $\text{PgC month}^{-1}$

the eastern portion of the adjacent ocean (e.g., Pacific Ocean and Atlantic Ocean, respectively). Transport tends to be dominated by transient eddies in the early stage of an eddy transport event (Figure 3.11B) after which stationary eddies take over (Figure 3.11C) suggesting that instabilities in baroclinic waves initiate poleward transport by synoptic eddies.

### 3.5.3 *Vertically Integrated Transport*

Column integrated transport in the annual mean and during boreal winter and summer are shown in Figure 3.12. In the annual mean, total transport in the column is equatorward south of  $45^\circ\text{N}$  and poleward north of  $45^\circ\text{N}$  (Figure 3.12A). Since net terrestrial biological surface flux is zero in the annual mean due to forced balance between GPP and respiration, net transport in the annual mean primarily reflects transport of fossil fuels out of northern mid-latitudes. Seasonal transport is, however, much larger and more strongly influenced by seasonal biology. Transport during boreal winter (Figure 3.12B), for example, is reminiscent of annual mean transport in pattern but much stronger in amplitude due to strong decomposition and respiration in the same zone as fossil fuel emissions. Transport during boreal summer (Figure 3.12C) leads to convergence of  $\text{CO}_2$  near  $60^\circ$  where strong terrestrial uptake occurs.

Vertically integrated transport by the MMC gives insight into the interplay of the strong vertical  $\text{CO}_2$  gradient and the single overturning cell exhibited in the mean circulation field on  $\theta_e$  shown in Figures 3.6 and 3.7. While poleward and equatorward transport by the MMC are large, they largely cancel one another with smaller residual equatorward transport in the annual mean with maximum equatorward transport of  $-0.6 \text{ PgC month}^{-1}$ . During summer, residual mean transport is poleward in the column integral, reflecting the change in sign of the vertical  $\text{CO}_2$  gradient. Net eddy transport, however, is poleward throughout the column (Figure 3.8) and in the column integral (Figure 3.12) in the annual mean. Although eddy transport is orders of magnitude smaller than mean transport within a single moist isentropic layer, it is poleward through the entire column. *Annual mean transport by*

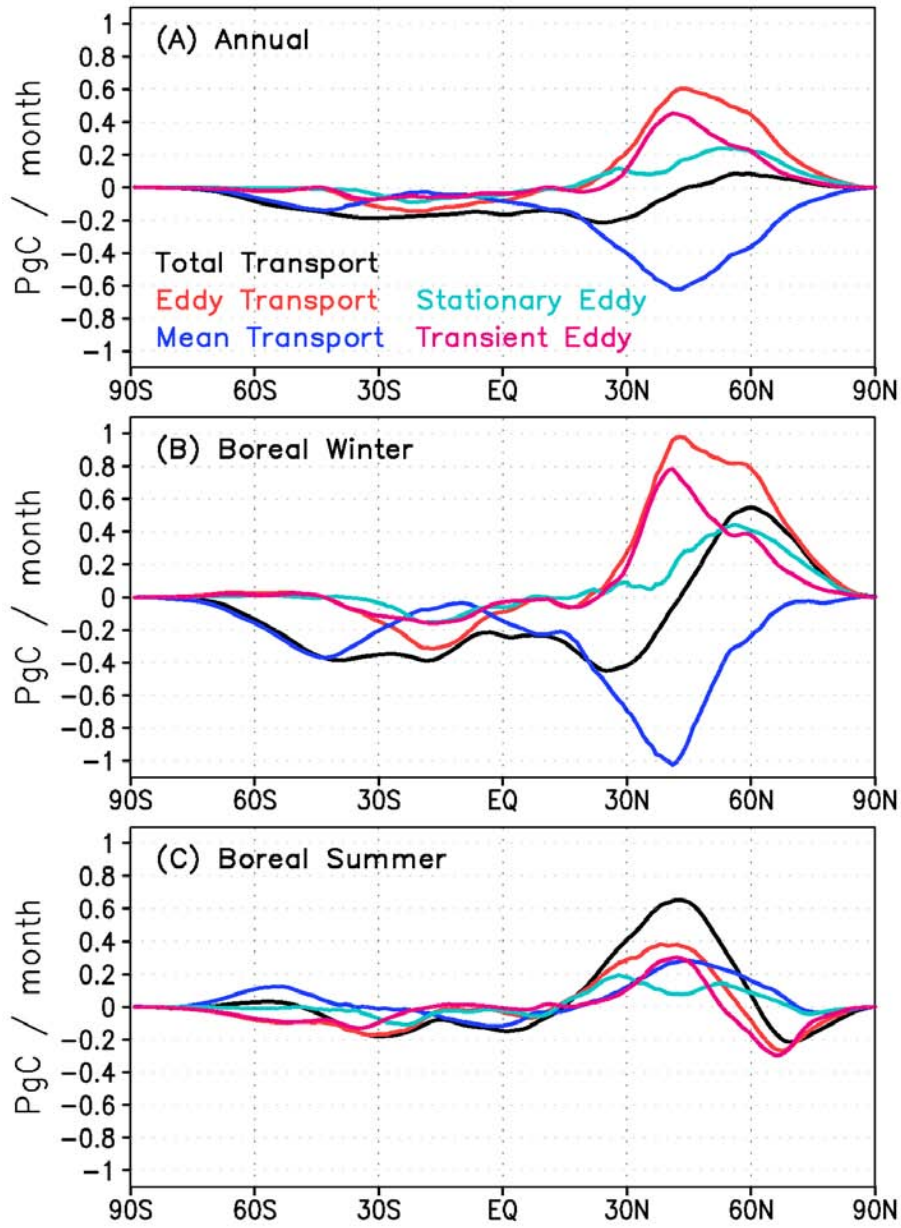


Figure 3.12: Column integrated transport in the annual mean (A) and during boreal winter (B) and summer (C). Total transport (black line) is parsed into transport by total eddy (red, transient + stationary), transient eddy (magenta), stationary eddy (cyan) and mean (blue) components.

*synoptic eddies is therefore systematically opposed in the column by transport by the MMC.* Poleward eddy transport peaks at  $0.6 \text{ PgC month}^{-1}$  in the annual mean and  $1.0 \text{ PgC month}^{-1}$  during boreal winter. Partitioning of eddy transport to stationary and transient components reveals similar patterns to those discussed in the spatial plots of

Figure 3.11, namely that eddy transport is dominated by transient eddies at lower latitudes and determined equally by transient and stationary eddies at higher latitudes.

#### 3.5.4 Dependence on Vertical Coordinate

Although PCTM is run on a terrain following coordinate system, partitioning of transport to eddy and mean components is sensitive to vertical coordinate. To demonstrate this, eddy and mean  $\text{CO}_2$  transport are diagnosed on  $\eta$ ,  $\theta_e$  and  $\theta$  coordinate systems as discussed in the beginning of this section and in Chapter 2.5.3, column integrated as in Figure 3.12, and plotted as annual means in Figure 3.13. Note that column integrated transport along coordinates which are monotonic with height ( $\eta$  and  $\theta$  at all latitudes and

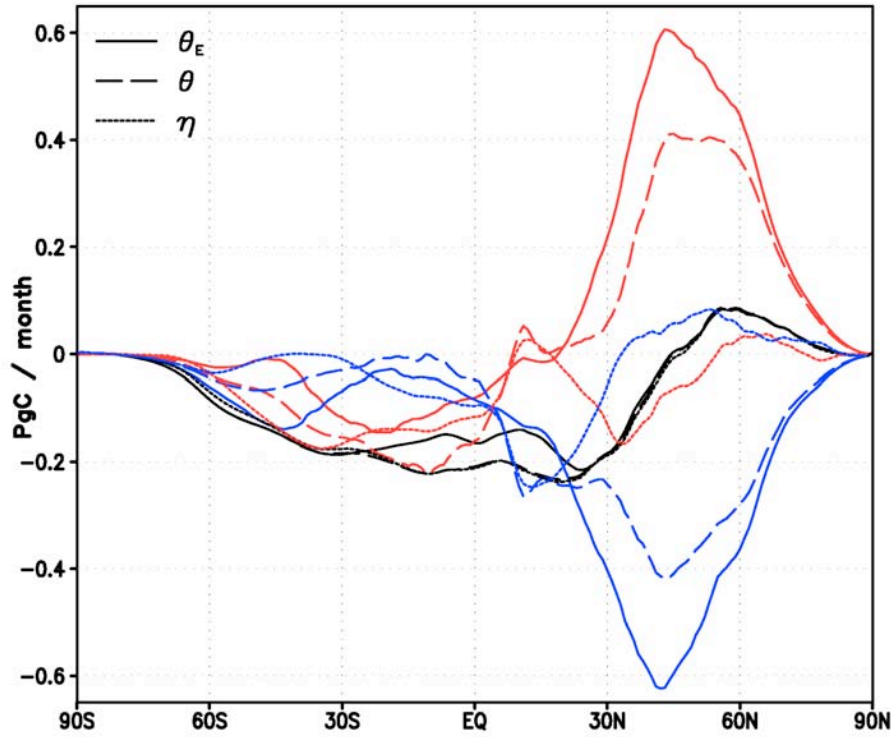


Figure 3.13: Column integrated transport in the annual mean as a function of vertical coordinate. Transport is calculated on moist isentropic surfaces ( $\theta_e$ , solid line), dry isentropic surfaces ( $\theta$ , dashed line), and terrain following surfaces ( $\eta$ , dotted line). Total transport (black) is partitioned to transport by eddy (red) and mean (blue) components.

$\theta_e$  outside the tropics) is equal and therefore conserved. There is, however, a small error

in  $\theta_e$  in the tropics where a mid-tropospheric minimum in  $\theta_e$  occurs. If total mass and  $\text{CO}_2$  are conserved in a column at any given latitude, total transport will (and should be) independent of coordinate system. This error has been investigated and is attributed to small errors in mass conservation when regridding in the tropics from monotonic  $\eta$  to non-monotonic  $\theta_e$ .

Outside of this small error in the tropics, total transport is conserved and therefore independent of coordinate system. *The partitioning of total transport to eddy and mean components, however, is strongly sensitive to vertical coordinate.* Focusing on the northern mid-latitudes with transport on  $\theta_e$  the reference (solid lines), eddy transport decreases by nearly half when calculated on  $\theta$  (red dashed line). Because total transport is conserved, decreased poleward transport by eddies in a  $\theta$  framework is compensated by decreased equatorward transport by the MMC (dashed blue line).

Mass transport on  $\theta_e$  includes a large contribution from moist air rising within storm-tracks [Pauluis et al., 2008], while averaging on dry coordinate systems such as  $\theta$  misses entirely this contribution. As demonstrated in Figure 3.13, the corresponding transport ends up in the mean term and in the annual mean case of  $\text{CO}_2$  transport strongly reduces eddy  $\text{CO}_2$  transport. The vertical component of motion gained from latent heat release in condensing moist air is therefore critical for describing tracer transport by moist ascent within synoptic waves. Increased transport due to latent heat release is most important in regions of condensing air such as moist conveyors in the warm sector of extratropical cyclones. For example, in the cold dry sector of the case study shown in Figure 3.9 (west side of a typical baroclinic wave), water vapor ( $q$ ) is zero,  $\theta_e$  is reduced to  $\theta$ , and equatorward transport is equal on  $\theta$  and  $\theta_e$ . Warm sector transport, however, is moist and  $q$  is nonzero. While  $\theta_e$  conserves energy in the presence of latent heating by condensation,  $\theta$  gains energy. As seen in Figure 3.13, this portion of eddy transport is parsed to the mean term when transport is diagnosed on  $\theta$ . As a result, column integrated transport by eddies is strongly diluted on  $\theta$ .

While eddy and mean transport patterns behave very similarly due to the Lagrangian nature of parcel following coordinates such as  $\theta$  and  $\theta_e$ , transport in an Eulerian framework such as the terrain following coordinate  $\eta$  provides an entirely different interpretation of transport within an atmospheric fluid. Column integrated eddy transport on  $\eta$  (red dotted line) actually opposes eddy transport on isentropes. The same phenomenon occurs in the case of mean transport. Both terms are nearly an order of magnitude smaller than their isentropic counterparts.

To demonstrate the effect of an Eulerian framework on interpretation of CO<sub>2</sub> transport, vertical distributions of CO<sub>2</sub> and mass flux are calculated on  $\theta_e$  and  $\eta$ . Annual mean vertical profiles are shown in Figure 3.14. The annual mean CO<sub>2</sub> distribution is clearly sensitive to coordinate system. In the Lagrangian framework, CO<sub>2</sub> gradients are very strong with respect to latitude and height and carry potential for large eddy CO<sub>2</sub> fluxes in response to baroclinic instabilities in the flow (Figure 3.14A). In the Eulerian framework, CO<sub>2</sub> gradients are extremely weak and carry little capacity for strong transport by eddies (Figure 3.14B). While  $\eta$  is tied to the surface and has little variability in space,  $\theta_e$  is strongly sensitive to moisture and heating and is highly variable in time and space. Since  $\theta_e$  surfaces tilt upward towards cold air, their climatologically favorable distribution is upward slant from equator to pole. This means parcels of air traveling along isentropes start in polluted boundary layers where CO<sub>2</sub> mixing ratios are enhanced for much of the year due to fossil fuel emission and respiration and rise upward and poleward along stormtracks out of polluted boundary layers into the free troposphere where CO<sub>2</sub> mixing ratios are depleted relative to the surface. CO<sub>2</sub> therefore decreases from equator to pole following  $\theta_e$  but increases following  $\eta$ .

The diagram in Figure 3.15 further illustrates this paradox. A vertical cross section of the atmosphere is shown for the midlatitudes from 20-70°N (north and poleward is to the left) from the surface up to the tropopause. For the purpose of this argument, attention is focused on the polar front and the red and blue air streams in the center of the diagram



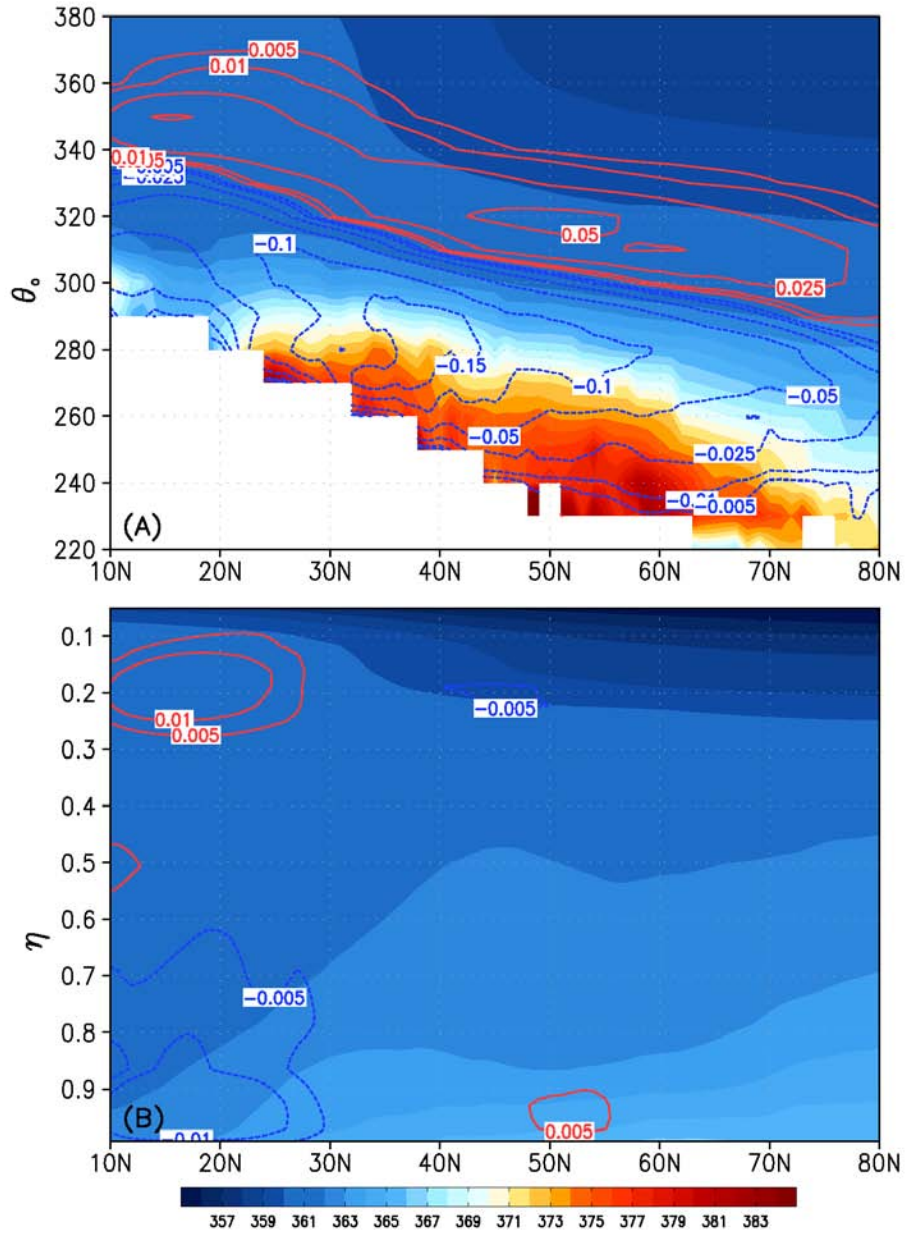


Figure 3.14: Zonal and month mean distribution of CO<sub>2</sub> (shaded) and meridional mass flux (contours) in the annual mean on  $\theta_e$  (A) and  $\eta$  (B). The colorbar for CO<sub>2</sub> (in ppm) is shown on the bottom. CO<sub>2</sub> is plotted on the same scale.

where extratropical cyclones and anticyclones occur. Dashed lines are not exact but can be thought of as isentropes, which are tilted upward and poleward. From our discussion of CO<sub>2</sub> distributions in the Eulerian framework it was shown that CO<sub>2</sub> tends to build up near the surface with peak values north of 50°N. From an Eulerian framework along for example



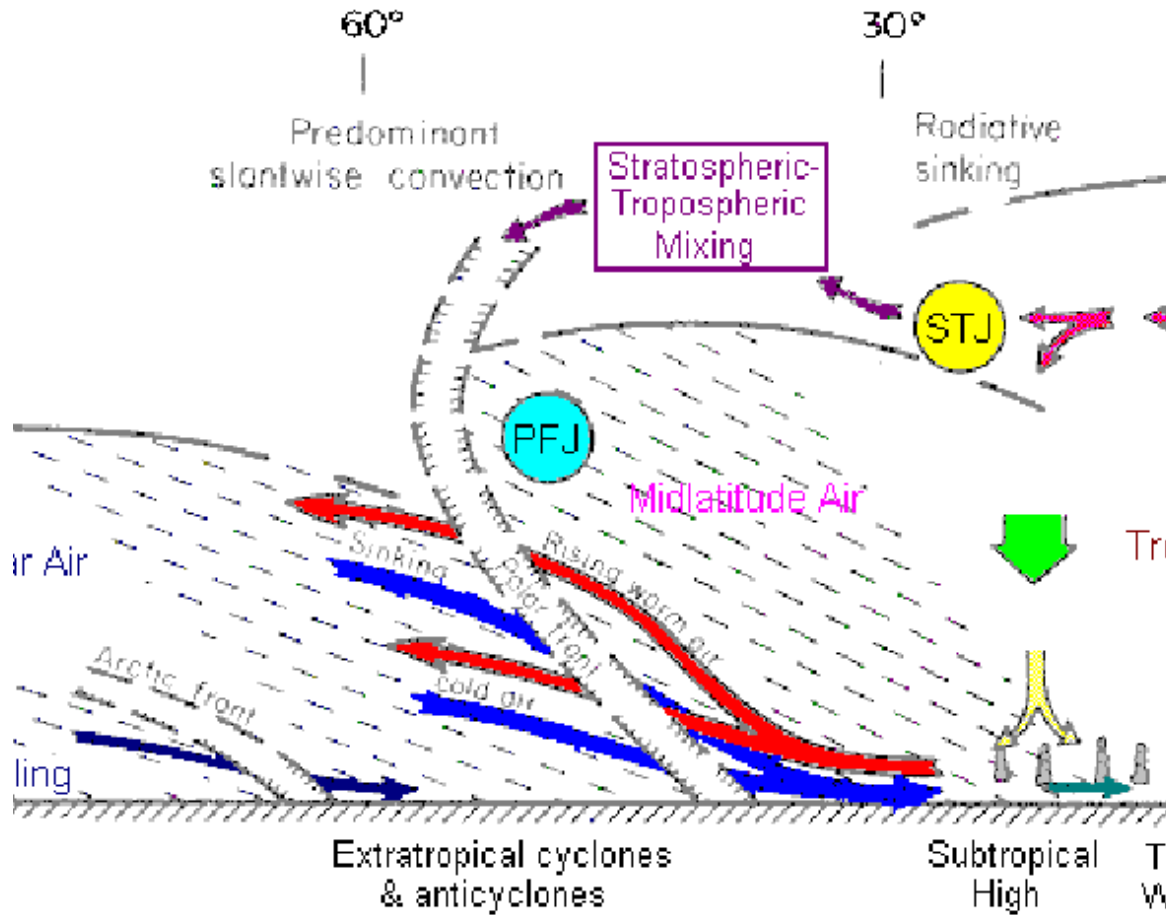


Figure 3.15: Diagram of isentropic surfaces and air streams along polar front (courtesy <http://www-das.uwyo.edu/geerts/cwx/notes/chap01/tropo.html>).

the surface, poleward moving air flows, on average, directly into  $\text{CO}_2$  enriched air from the south such that  $\text{CO}_2$  depleted air is carried poleward. This results in negative covariance of  $\text{CO}_2$  and mass flux and net eddy  $\text{CO}_2$  transport is equatorward, as shown in the dashed red line near  $30^\circ\text{N}$  in Figure 3.13. From a Lagrangian perspective of isentropes, poleward moving air follows the red line of rising warm air (the slope of which depends on whether  $\theta$  or  $\theta_e$  is used) and flows, on average, directly out of  $\text{CO}_2$  enriched air near the surface into the free troposphere such that  $\text{CO}_2$  enriched air is carried poleward. This parcel following flows results in positive covariance of  $\text{CO}_2$  and mass flux and net eddy  $\text{CO}_2$  transport is poleward, as shown in the solid red line in Figure 3.13.

### 3.6 Seasonal Variations

Meridional transport strongly mediates the seasonal cycle of CO<sub>2</sub> in mid-latitudes and amplifies it in high latitudes, accounting for most of the observed seasonality at sites like Barrow, Alaska and Alert, Canada [Fung et al., 1983]. Comparison of satellite observed vegetation with modeled transport implies that atmospheric mixing produces signals at high latitudes where vegetation is sparse [Fung et al., 1987]. To convey the influence of eddy and mean transport on seasonality in the Arctic from a  $\theta_e$  perspective, the seasonal cycle of the atmospheric CO<sub>2</sub> tendency (tendencies in ppm month<sup>-1</sup> are approximated by dividing Equation 1 through by zonal and time averaged surface pressure) binned by latitude into the Arctic and mid-latitudes is plotted as a function of eddy transport, mean transport and surface flux tendencies (Figure 3.16).

CO<sub>2</sub> seasonality is stronger in the Arctic despite weak surface fluxes (Figure 3.16A). This is driven strongly by synoptic eddies and therefore moist synoptic storms. Polar seasonality is strongly opposed by the mean meridional circulation, with net equatorward transport (at all northern latitudes) from September-June and poleward transport during summer. The same transport processes responsible for seasonality in the Arctic also strongly damp the seasonal cycle in mid-latitudes to about 50% of the seasonality implied by NEE, with eddy transport of similar magnitude to fossil fuel emissions during summer and fall (Figure 3.16B). Eddy transport tends to lead mean transport by 2-4 weeks during summer in the Arctic.

### 3.7 Conceptual Diagram

Results of Part 1 of this study are summarized in Figure 3.17. Synchrony of baroclinic waves with terrestrial ecosystem metabolism governs “meridional sloshing” of CO<sub>2</sub> between middle and high latitudes, which modulates seasonality at mid-latitudes and amplifies seasonality at high latitudes. Poleward flow of warm moist air, high in CO<sub>2</sub> relative to high latitudes, is swept along moist isentropic surfaces along the east side of baroclinic waves.

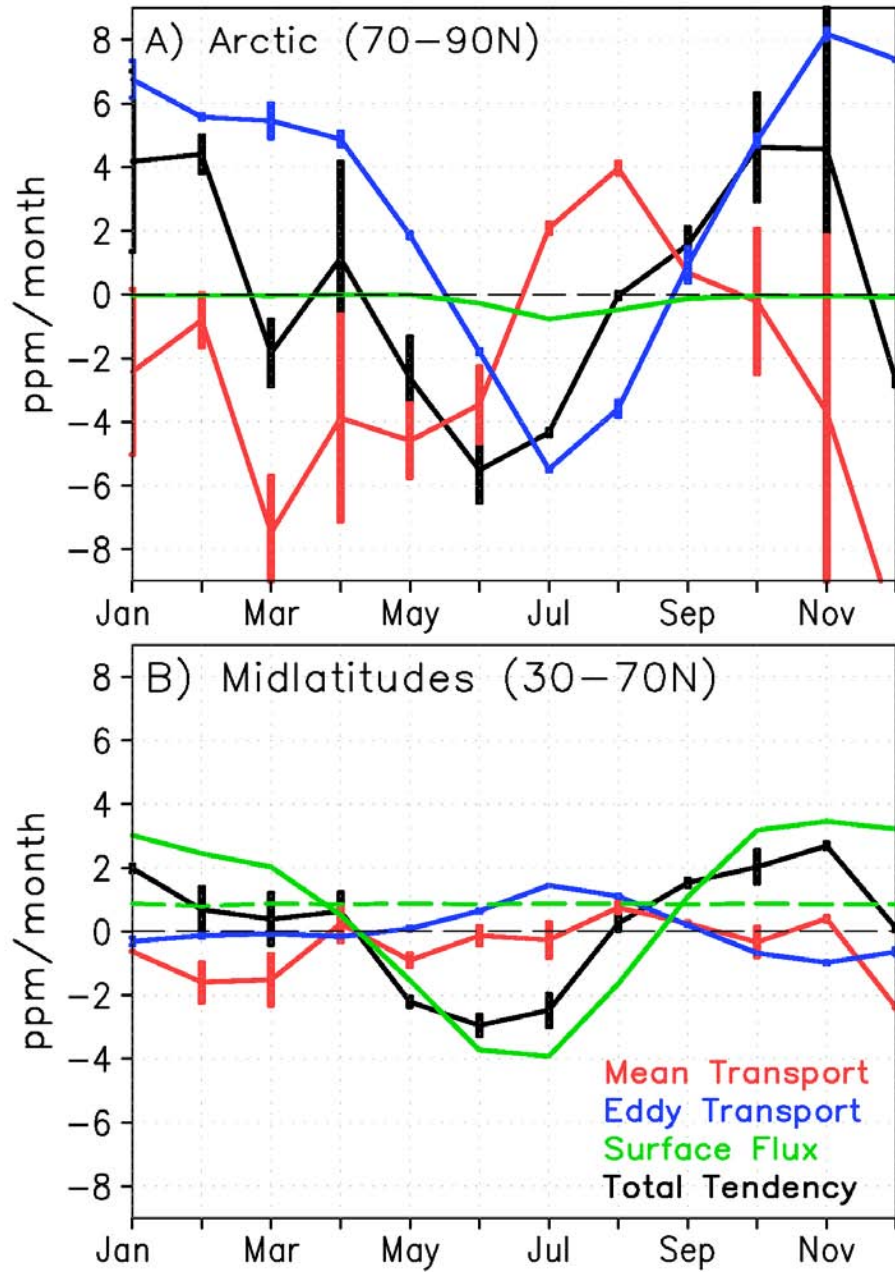


Figure 3.16: Column integrated seasonal CO<sub>2</sub> budget tendencies (ppm month<sup>-1</sup>) on  $\theta_e$  binned into (A) high latitudes and (B) mid-latitudes. Transport by transient and stationary eddies is plotted in blue, mean meridional transport in red, the total CO<sub>2</sub> tendency in black, and the total surface flux of carbon from land, ocean, and fossil fuels in green. The fossil fuel tendency is plotted as a dashed green line. The sum of individual tendencies (red, blue and green lines) is equal to the total tendency (black line). Error bars represent the root mean squared error for the two years of averaging used.



Figure 3.17: Illustration of important  $\text{CO}_2$  transport mechanisms associated with a baroclinic wave.

As this air is forced up over the polar front along the mid-latitude stormtrack, high  $\text{CO}_2$  air is dumped into the polar vortex and a huge comma-shaped cloud forms. Poleward flow is compensated by equatorward flow of cold dry air, low in  $\text{CO}_2$  relative to low latitudes. This air descends from high polar altitudes into lower latitudes from the vortex along the surface in the western trough portion of the baroclinic wave. Mean polar flow, not depicted graphically in this study, opposes transport by synoptic eddies with net equatorward transport of  $\text{CO}_2$  during winter. The directions of eddy and mean transport are reversed during boreal summer as baroclinic waves weaken and the terrestrial biosphere draws  $\text{CO}_2$  out of the atmosphere.

### 3.8 Discussion and Conclusions from Part 1

Synoptic weather systems transport a large amount of CO<sub>2</sub> north and south in mid-latitudes, up to 1 PgC month<sup>-1</sup> during winter when baroclinic waves ramp up. As a result, seasonality is strongly damped in mid-latitudes and amplified in the Arctic. To put this in perspective, synoptic eddies transport more CO<sub>2</sub> out of mid-latitudes than is emitted by fossil fuels. The total amount of CO<sub>2</sub> transported by synoptic eddies is most likely sensitive to model choice and should increase as models become more skilled at representing important eddy processes such as frontal circulation, moist convection and precipitation. Because eddies mediate seasonality in northern latitudes, it is critical that inversion scientists (modelers and observationalists) consider seasonal meridional transport, and the underlying mechanisms, carefully.

Sensitivity of eddy transport to factors such as storm track position, seasonal tendencies in CO<sub>2</sub> mixing ratio and the pattern of seasonal change in surface CO<sub>2</sub> flux over the globe poses a challenging task for inversion modelers. Additionally, interpretation of synoptic variations in continental records requires accurate simulation of frontal weather systems [Geels et al., 2004; Wang et al., 2007; Parazoo et al., 2008] and moist conveyors. As discussed in Chapter 3 and by Corbin et al. [2006], moist processes associated with synoptic weather systems hide much of the dynamics from satellites, and likely from other observing systems [e.g., aircraft flask samples, Stephens et al., 2007]. Additionally, warm conveyors transport CO<sub>2</sub> and other trace gases into the polar vortex where they are hidden during polar winter. Continuous in-situ records can supplement airborne and remotely observed measurements during inclement weather, but only at a few locations. This fair-weather bias in measurements puts stringent requirements on models of moist transport. We therefore recommend that inversion modelers pay special attention to modeling of wet synoptic storms in addition to in-situ and satellite observations, with particular attention given to factors such as grid spacing (e.g., frontal circulations better resolved with finer grid spacing), representation of moist convection along fronts and within the warm sector

of cyclones, and assimilation of observations of moisture.

An additional complication arises in dry parts of the mid-latitude circulation where high-pressure systems transport  $\text{CO}_2$  in different directions within an atmospheric column due to convergence aloft and divergence near the surface. Transport aloft in mid-latitudes, for example, is driven in large part by deep convection in tropical latitudes, which is typically parameterized and therefore poorly represented in global models. That being said, observations are ideally suited for constraining the dry portion of the atmosphere. For example, satellite measurements help constrain flow in the column, in-situ measurements constrain flow at the surface, and periodic aircraft measurements from field campaigns such as the HIPER Pole-to-Pole Observations Experiment (HIPPO) constrain flow throughout the lower and free troposphere. Without observations such as these, there is otherwise over-reliance on models to distinguish between large and opposing directions of transport.

Regional inversion modelers will be one of the primary victims of the vast amounts of seasonal transport by synoptic storms. It is important to correctly represent large-scale meridional  $\text{CO}_2$  gradients flowing over regional domains because interaction with synoptic waves generates a source of variability that may confound flux attribution [Keppel-Aleks et al., 2010]. It is not surprising, given the unobserved and unresolved nature of moist synoptic transport, that Schuh et al. [2010] find that the annual sink over N. A. varies by  $\sim 30\%$  when lateral boundary conditions are prescribed from two different global transport models. Because of large meridional transport by synoptic eddies, regional inversions should carefully account for meridional advection at north and south borders when prescribing lateral boundary conditions.

Furthermore, it is possible that interannual variability in eddy (or mean) transport, in response for example to teleconnection patterns such as the North Atlantic Oscillation [e.g., Hurrell, 1995], has contributed to trends in  $\text{CO}_2$  seasonality in the Arctic [e.g., Keeling et al., 1996; Randerson et al., 1997]. Keeling et al. [1996] reported that the seasonal cycle of  $\text{CO}_2$  increased in amplitude by 40% in the Arctic from the early 1960's to the mid-

90's. While the study also provided evidence that increased seasonality is linked to global warming and increased uptake by plants, a subsequent investigation by Randerson et al. [1997] showed that such large trends in uptake couldn't be explained by CO<sub>2</sub> fertilization alone, and that other processes must also contribute. Given the large year-to-year changes in the amount of mass that sloshes back and forth between middle and polar latitudes, and the possible link between global warming, NAO index, and winter storm strength, it is reasonable to expect that trends in seasonal exchange of CO<sub>2</sub> between middle and polar latitudes may contribute to trends in seasonal CO<sub>2</sub> in the Arctic. While such analysis is beyond the scope of this study, it is an excellent topic for future studies.

Finally, it is worth exploring in more detail to what degree synoptic transport into the polar vortex influences equatorward moving air at a later point in time. Previous studies show that point emissions of tracer can scatter widely north, south, and across ocean basins through interaction with baroclinic waves [e.g., Cooper et al., 2001, 2004]. It is common, for example, for baroclinic waves to advect polluted Asian air into the Arctic. A portion of this air is advected back south into N. A. with the same system, but the rest mixes in the polar vortex. Counter to intuition CO<sub>2</sub> anomalies in air arriving at mid-latitude sites from the north may actually reflect fluxes that occurred earlier to the south, and mixing with the polar vortex scrambles any coherent link between the longitudinal origin of poleward and equatorward moving air.

## Chapter 4

### UNCERTAINTY IN MOIST SYNOPTIC TRANSPORT OF CO<sub>2</sub>

Part 1 of this study, presented in Chapter 3, showed that CO<sub>2</sub> transport by moist synoptic storms in northern mid-latitudes is significant, intimately tied to frontal and moist processes and therefore poorly resolved and poorly represented in most global transport models, and systematically unobserved by satellites. Because a significant portion of CO<sub>2</sub> transport is unobserved, there is a lot of pressure on models of moist transport to represent synoptic transport processes with high fidelity. Part 2 of this study seeks a better understanding of how well different transport products represent moist synoptic transport. Of particular interest to me is sensitivity of net meridional transport by mid-latitude storms to grid spacing and fine-scale vertical mixing. It is entirely possible that transport differences between models exist but are mostly random and largely cancel out over long enough time scales. It is just as likely, however, that systematic transport differences exist, in which case biased transport errors are more prone to be aliased into inverse estimates of CO<sub>2</sub> flux.

In order to quantify uncertainty associated with resolved and unresolved synoptic processes, Part 2 includes analysis of meridional transport of seasonally varying CO<sub>2</sub> using a global transport model, identical surface fluxes and initial conditions, and transport driven by four reanalysis products from GEOS-DAS (described in Chapter 2.3). Using a common transport model (PCTM), transport differences arise strictly from differences in meteorological forcing, unlike in some previous studies such as TransCom which also differ in transport algorithms. With confirmation in Chapter 3 of the importance of moist synoptic processes for carbon transport, transport is diagnosed strictly along  $\theta_e$ .



## 4.1 Introduction

Model projections of climate change depend critically on accurate predictions of carbon sources and sinks [Friedlingstein et al., 2006; IPCC, 2007]. Inverse methods combine information from atmospheric models and CO<sub>2</sub> mixing ratio measurements to provide process-based and observed constraint to source and sink estimates. This “top-down” approach is useful at global scales but mostly unsuccessful to date at regional scales where information about underlying processes is contained [Baker et al., 2006]. While in-situ and spaced-based measurement constraints have improved in quantity and quality in recent years, strong sensitivity to transport error continues to be a key limitation at regional scales [Gurney et al., 2004; Houweling et al., 2010]. For example, seasonal and diurnal covariance between ecosystem metabolism and fine-scale vertical transport in the atmosphere is a strong determinant of vertical structure in CO<sub>2</sub> (the CO<sub>2</sub> “rectifier effect,” Denning et al. [1995, 1996, 1999]). Numerical treatment of subgrid-scale vertical mixing is a leading source of uncertainty in CO<sub>2</sub> inverse models [Denning et al., 1999; Yi et al., 2004; Helliker et al., 2004; Baker et al., 2005; Yang et al., 2007; Stephens et al., 2007]. As discussed in Chapter 3, transport by baroclinic waves is another process that controls the distribution of CO<sub>2</sub> on seasonal time scales.

The global CO<sub>2</sub> observing system has grown to include continuous records at continental locations close to terrestrial sources. These records allow quantitative flux estimation at finer spatial scales than previously possible [e.g., Law et al., 2003; Peylin et al., 2005; Zupanski et al., 2007; Lauvaux et al., 2008; Schuh et al., 2009], but feature much greater variance at synoptic scales. Interpretations of these variations require accurate simulation of structure, timing, and vertical motions associated with synoptic waves and frontal weather systems [Geels et al., 2004; Hurwitz et al., 2004; Wang et al., 2007; Parazoo et al., 2008].

Tracer transport by baroclinic waves is associated with large- and small-scale features. Representation of large-scale features ( $\sim 1000$  km’s) in numerical models, such as the position and timing of a baroclinic wave, is strongly dependent on grid spacing, with

finer grid spacing improving the chances of representing features within a wave. Model inter-comparison studies show that simulations at higher spatial resolution produce better match with observed synoptic variations at northern mid-latitudes sites [Geels et al., 2007; Patra et al., 2008; Ahmadov et al., 2009].

Unfortunately, certain features of synoptic waves occur at scales smaller than most global inverse models ( $1^\circ$  globally or coarser,  $\sim 100$  km) can resolve. Frontal ascent of tracer, for example, is typically confined to a 60 km wide region [Donnel et al., 2001], and is therefore unresolved down to grid spacing of 10 km or less. Additionally, transport by moist convection and turbulent mixing typically occur at scales smaller ( $< 10$  km) than most atmospheric models can resolve. These processes are solved instead through parameterization, which is highly variable between models, not as easily constrained in meteorological reanalysis as resolvable flows, and therefore highly uncertain. It is unclear what role slantwise vertical transport by synoptic storms plays in top-down  $\text{CO}_2$  flux estimates, but because vertical mixing, surface  $\text{CO}_2$  flux intensity, and  $\text{CO}_2$  spatial gradients are tightly linked in inverse calculations, errors in vertical mixing translate into errors in vertical  $\text{CO}_2$  distribution and thus  $\text{CO}_2$  flux estimates [Stephens et al., 2007]. For example, models in TransCom3 that trap more  $\text{CO}_2$  near the surface due to weak vertical mixing tend to overestimate the strength of the NH terrestrial sink.

In the era of remotely-sensed observations of trace gas from satellites, an additional complication arises because the strongest horizontal gradients in  $\text{CO}_2$  occur along frontal boundaries that are typically hidden from orbital sensors by clouds [Corbin et al., 2006], and much of the seasonality at northern latitudes is determined by moist conveyors which will hide  $\text{CO}_2$  transport from satellites [Parazoo et al., 2011]. Corbin et al. [2008] analyzed simulated column-mean  $\text{CO}_2$  using a clear-sky sampling regime along hypothetical satellite orbits and found that systematic sampling errors of up to 1.5 ppm in the seasonal mean were caused by cloud masking of frontal  $\text{CO}_2$  gradients. Covariance of synoptic transport with surface  $\text{CO}_2$  flux is likely to cause errors in top-down flux estimates if not represented

correctly in transport models. Since a portion of the synoptic signal is unobserved, it is important to understand uncertainty associated with synoptic processes.

In order to quantify uncertainty associated with resolved and unresolved synoptic processes, we analyze meridional transport of seasonally varying CO<sub>2</sub> using a global transport model, identical surface fluxes, and transport driven by four reanalysis products from GEOS-DAS. This experiment is aimed at isolating transport as the sole cause of differences between simulations. Synoptic scale eddy CO<sub>2</sub> transport is diagnosed along  $\theta_e$  following Chapters 2 and 3. Reanalysis products were described in detail in Chapter 2.3. Discussion of transport uncertainty will be the primary focus of the first half of Section 3.2. Modeled synoptic scale variations of CO<sub>2</sub> are compared to observed variations at continental and remote sites at the beginning of Section 3.2 so that model fidelity may be evaluated. This research also seeks to identify key differences between models that may be responsible for uncertainties in eddy transport. Possible explanations for transport uncertainty are discussed in the latter half of Section 3.2. Discussion and concluding remarks are provided at the end of this chapter.

While the major focus of Part 2 is quantification of uncertainties in moist eddy transport and identification of possible causes of uncertainty, comparison to observations are made in order to (1) instill some sense of realism in the simulations and (2) provide indication that model uncertainty in eddy transport is a result of improved representation and resolution of physical processes as models move towards higher fidelity. Model comparisons are made to well-calibrated in-situ observations of CO<sub>2</sub> from northern mid-latitude surface and tall tower sites at hourly resolution. The 12 sites chosen for this study are described in more detail in Parazoo et al. [2008]. Lag-correlations are used as a measure of correlation and timing (or phase) of synoptic variations, with seasonal and diurnal variations removed from time series of modeled and observed CO<sub>2</sub> using a butterworth filter.

## 4.2 Comparison to Observations

Modeled synoptic variations are fairly well correlated with observations at zero time lag with an average correlation among all sites and models of  $r=0.496$  (see Figure 4.1). The strongest correlations occur at remote sites (ZEP, PAL;  $r=0.66$ ) and coastal sites (MHD, SBL;  $r=0.64$ ). Weaker correlations occur over continental terrestrial sites (e.g., LEF, CDL, and WPL), which tend to be more sensitive to surface fluxes. Overall correlations vary only slightly between models, with  $r = 0.476$  for GEOS4-2.5x2,  $r = 0.482$  for GEOS4-1.25x1,  $r = 0.509$  for GEOS5-1.25x1, and  $r = 0.517$  for GEOS5-0.67x0.5, although model dependence of correlation is probably better interpreted on a site-to-site basis. Relative to a select set of continuous observations, day-to-day variability is generally more realistic in GEOS5-0.67x0.5.

These plots provide little indication of the amplitude of day-to-day variability at these locations. To demonstrate amplitude between models relative to observations, modeled and observed surface  $\text{CO}_2$  mixing ratios are plotted for a 30-day period in January 2005 at LEF in Figure 4.2. While the timing and structure of synoptic features is consistent between models and observations, the amplitudes of GEOS5-0.67x0.5 and GEOS5-1.25x1 are significantly overestimated relative to observations and GEOS4-2.5x2 and GEOS4-1.25x1. Although more extensive analysis of model spread in amplitude and correlation relative to observations would be useful, to avoid a lengthy discussion I will summarize by mentioning that these strong differences in amplitude occur at most of the other mid-latitude and boreal sites during winter, especially at higher latitudes, but that model spread in amplitude is significantly weaker during summer. Whether these model differences in correlation and amplitude at synoptic timescales can be attributed to differences in eddy and/or mean transport is the subject of the remainder of this paper.

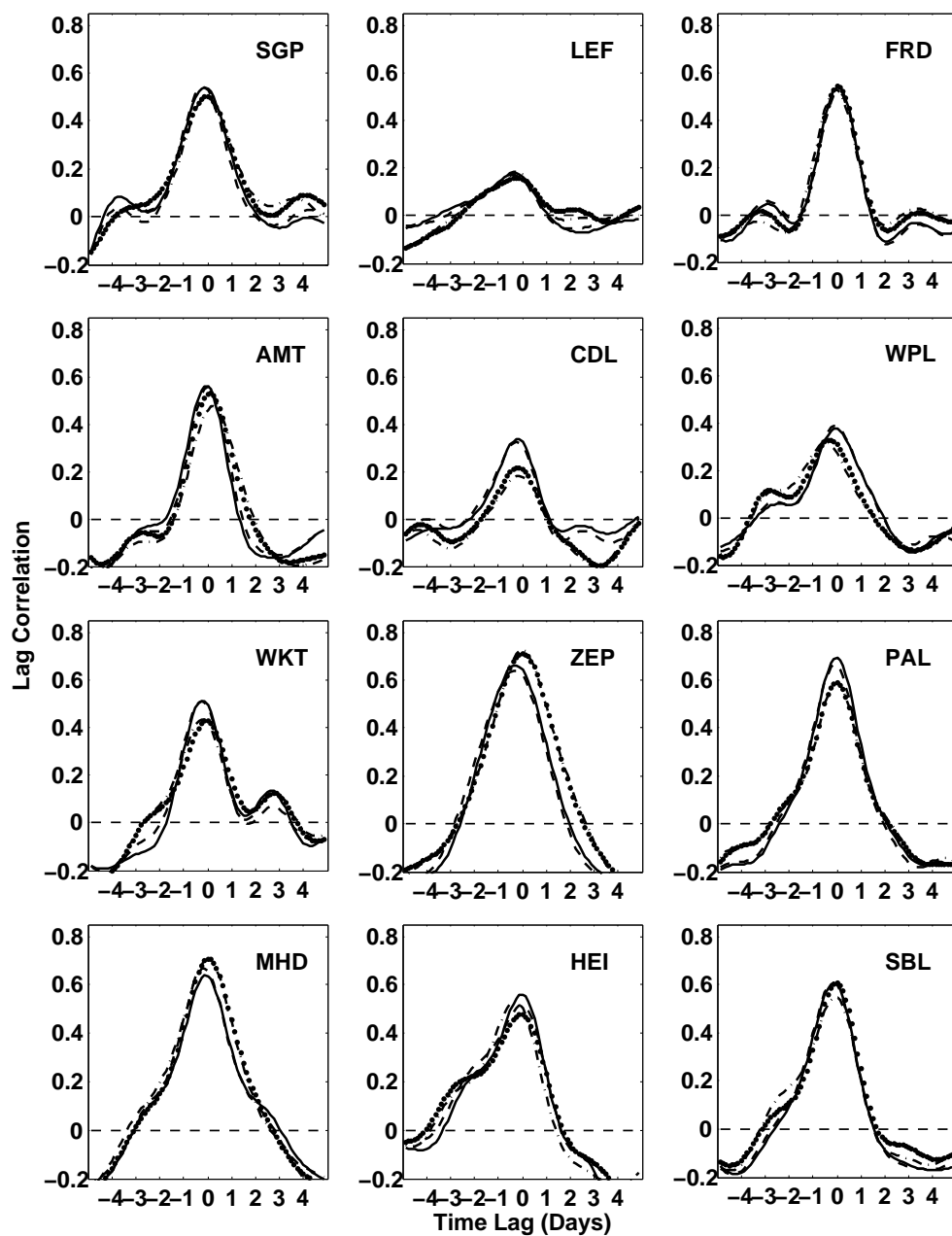


Figure 4.1: Lag Correlations between model and observations at mid-latitude sites in North America and Europe. The four line styles refers to the model used in the correlation, where GEOS5-0.67x0.5 is solid, GEOS5-1.25x1 as dashed, GEOS4-1.25x1 is dotted, and GEOS4-2.5x2 is dash-dotted. These same line styles are used in the remaining figures in the Chapter.

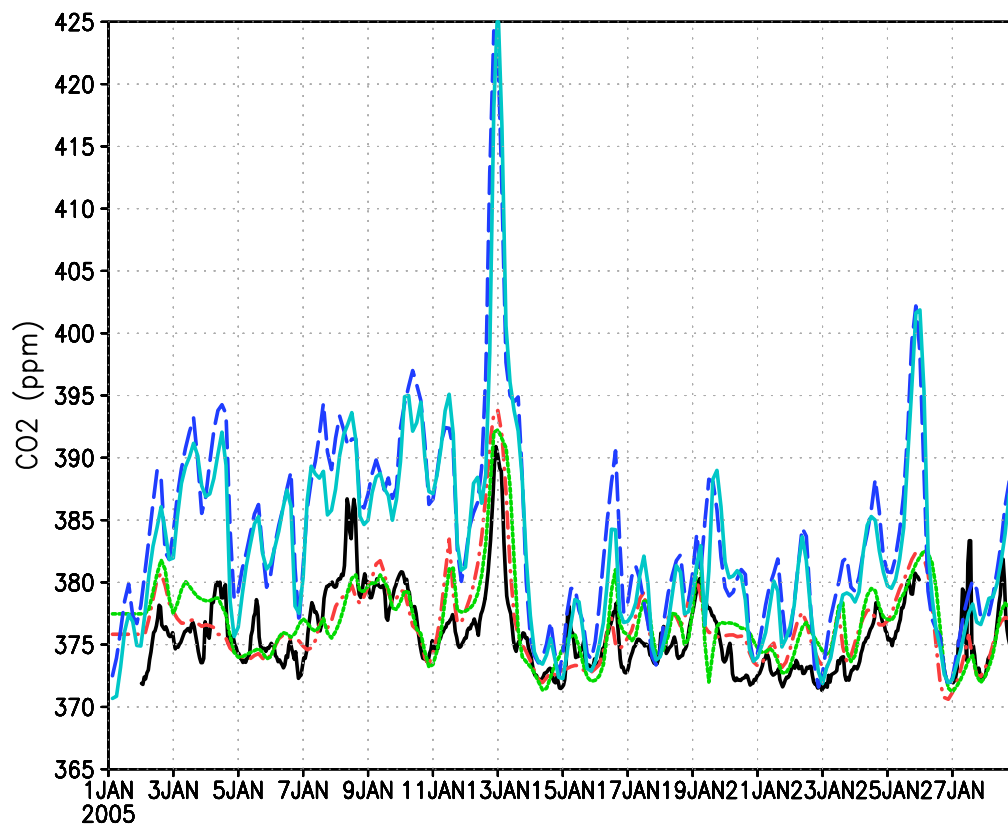


Figure 4.2: Modeled and observed surface  $\text{CO}_2$  mixing ratio from January 1-30, 2005 at a site in northern Wisconsin (LEF). Observations are plotted at 76 m above the surface in black. Line styles for models are the same as in Figure 4.1

### 4.3 Estimates of Transport Uncertainty

Total column transport in the annual mean is approximately conserved between models (a small exception exists at low latitudes), with net northward transport north of  $50^\circ\text{N}$  and net southward transport south of  $50^\circ\text{N}$  (Figure 4.3A). Eddy transport is poleward in both hemispheres and opposed by mean transport, which is strongly southward in northern latitudes in the annual mean. Eddy and mean transport peak in northern mid-latitudes, consistent with strong terrestrial breathing of the planet and industrial emissions, and during boreal winter when baroclinic waves are most active. Mean transport in the column is a small residual of cancellation between strong poleward transport aloft and equatorward transport near the surface; eddy transport is weak but poleward through most of the

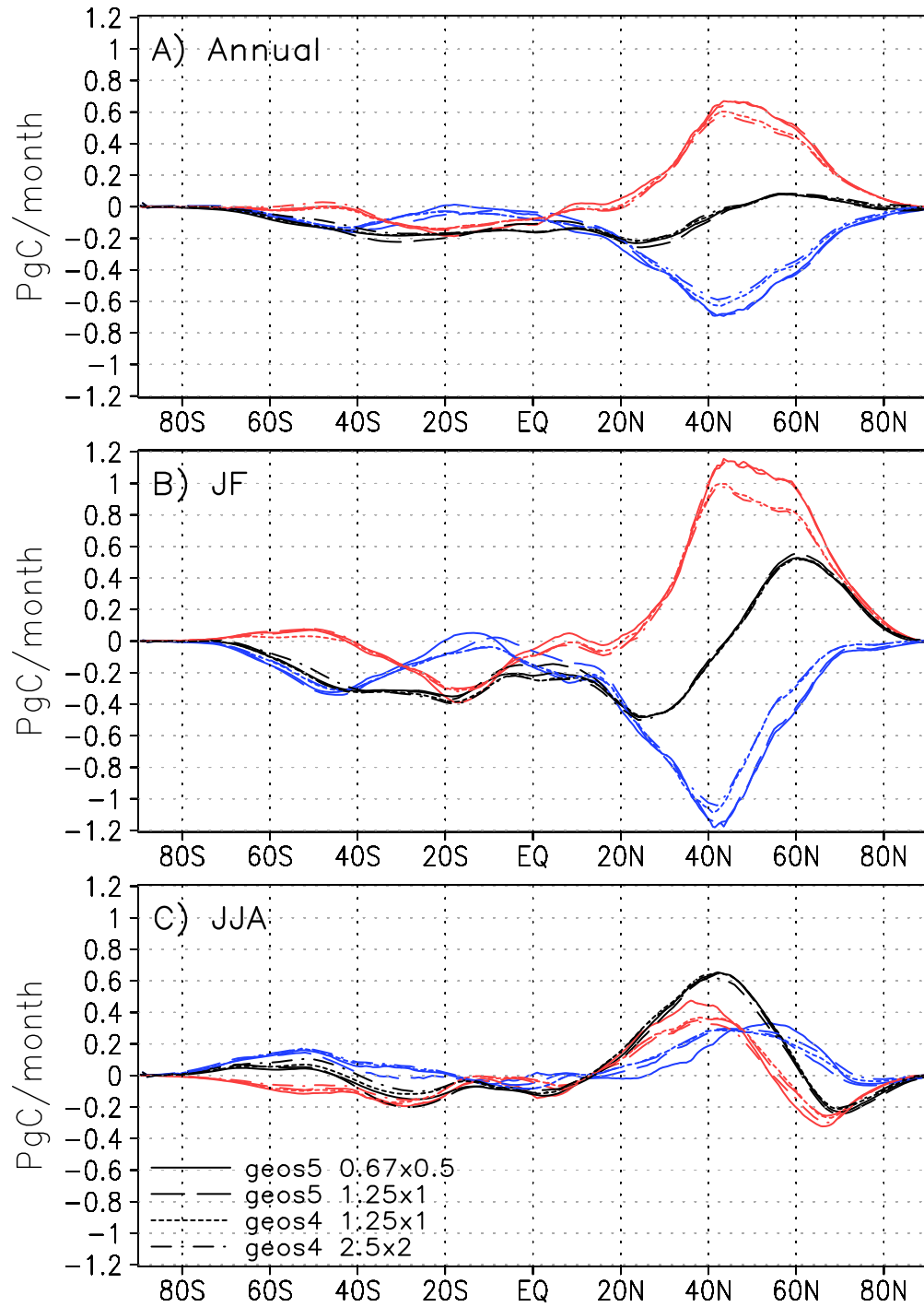


Figure 4.3: Total column integrated CO<sub>2</sub> transport (black) partitioned into transport by the mean circulation (blue) and synoptic eddies (baroclinic + stationary waves, red) in the annual mean (A) and during boreal winter (Jan-Feb, B) and summer (Jun-Jul-Aug, C).

troposphere [Parazoo et al., 2011].

There are important differences in model partitioning of meridional transport into eddy and mean components. Eddy transport differences between GEOS4 and GEOS5 are as large as  $0.1 \text{ PgC month}^{-1}$  from  $40\text{-}60^\circ\text{N}$  in the annual mean (Figure 4.3A), representing an error of 15-20% relative to peak transport. Seasonal differences are sometimes twice as large; for example, differences up to  $0.2 \text{ PgC month}^{-1}$  from  $40\text{-}60^\circ\text{N}$  occur during boreal winter (Figure 4.3B) and  $0.2 \text{ PgC month}^{-1}$  from  $35\text{-}40^\circ\text{N}$  during boreal summer (Figure 4.3C). Eddy transport is weaker in GEOS4 than GEOS5 and weakest in GEOS4-2.5x2. Eddy transport differences between the two GEOS4 resolutions and between the two GEOS5 resolutions are generally small in mid-latitudes. To conserve total meridional transport in the column, increased northward eddy transport (for example in GEOS5) is compensated by increased southward mean transport at the same latitude.

The contribution of eddies to seasonal tendencies of  $\text{CO}_2$  in the Arctic (defined as  $70\text{-}90^\circ\text{N}$ , tendencies have units of  $\text{ppm month}^{-1}$ ) is shown in Figure 4.4A. Since eddy tendencies are calculated for the polar cap, with meridional exchange occurring with mid-latitudes at  $70^\circ\text{N}$ , positive tendencies represent net eddy transport from mid-latitudes into the Arctic (vice versa for negative tendencies). Model differences of  $\sim 1 \text{ ppm month}^{-1}$  occur throughout the year, with all models contributing positively from September-May and negatively from June-August. Eddy exchange between middle and polar latitudes is strongest in GEOS5-0.67x0.5 during all months, and therefore contributes most strongly to seasonality in the Arctic. The weakest eddy exchange occurs in GEOS4 during winter and in GEOS4-1.25x1 and GEOS5-1.25x1 during summer.

Model differences are smaller in magnitude ( $\sim 0.5 \text{ ppm month}^{-1}$ ) in mid-latitudes ( $30\text{-}70^\circ\text{N}$ ) but account for a larger percentage of total eddy transport (Figure 4.4B). Unlike the Arctic, where net meridional exchange occurs only at the southern boundary (i.e.,  $70^\circ\text{N}$ ), mid-latitude exchange occurs at two boundaries ( $30^\circ\text{N}$  and  $70^\circ\text{N}$ ). Caution must therefore be taken when interpreting the cause of mid-latitude tendencies; positive tendencies, for



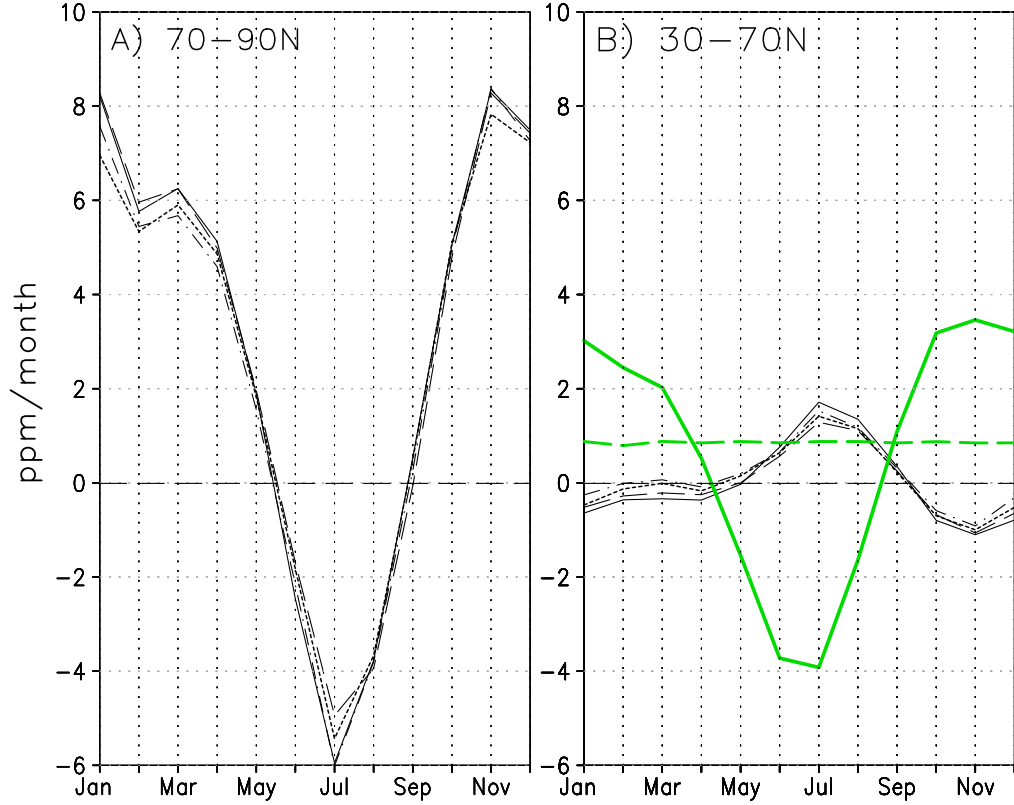


Figure 4.4: Column integrated seasonal CO<sub>2</sub> budget tendencies (ppm month<sup>-1</sup>) for eddies binned into high latitudes (70-90°N, A) and mid-latitudes (30-70°N, B). Tendencies due to total surface flux (green solid) and fossil fuel emissions (green dashed) are shown in (B).

example, can result from equatorward transport out of the Arctic and/or poleward transport out of mid-latitudes. As in the Arctic, eddy transport is strongest in GEOS4-0.67x0.5, with peak positive tendencies in July of  $\sim 1.75$  ppm month<sup>-1</sup> and peak negative tendencies during November of  $\sim 1.1$  ppm month<sup>-1</sup>. In all model cases, eddy transport is weakest from January to April, with tendencies of  $\sim 0.5$  ppm month<sup>-1</sup> in GEOS5-0.67x0.5 but zero net tendency in GEOS4-2.5x2. Put another way, coarse grid models (such as GEOS4-2.5x2) suggest zero net eddy transport out of mid-latitudes (poleward transport at 30°N equals poleward transport at 70°N) during the first third of the year while fine grid models (such as GEOS5-0.67x0.5) suggest up to 0.5 ppm month<sup>-1</sup> net eddy transport out of mid-latitudes.

Model dependence of transport out of mid-latitudes by differential eddy transport at the north and south “borders” is likely to be misinterpreted by regional inverse models.

It is even worse considering that the tendency due to eddies is sometimes twice the fossil fuel signal ( $0.9 \text{ ppm month}^{-1}$ , dashed green line in Figure 4.4B). For example, transport by GEOS5-0.67x0.5 is nearly  $1.75 \text{ ppm month}^{-1}$  during July alone. *The large tendency due to eddies puts a lot of pressure on models, which unfortunately experience eddy transport errors of  $0.5 \text{ ppm month}^{-1}$ , or half the tendency due to fossil fuels, for much of the year.*

Synoptic eddies are composed of transient and stationary waves. The primary difference is that stationary waves are anchored geographically by season while transient waves arise from baroclinic instability within stationary waves. Because baroclinic instability in transient waves leads to fine scale circulations which may be unresolved and moist processes (such as convection) which must often be parameterized, transport by transient waves is more uncertain than stationary waves. To determine which is more uncertain, annual and seasonal mean transport by each component is plotted in Figure 4.5.

Transient and stationary waves transport  $\text{CO}_2$  poleward in the annual mean (Figure 4.5A), with peak transport by transient waves near  $40^\circ\text{N}$  and stationary waves near  $55^\circ\text{N}$ . Transient waves dominate the total eddy signal from  $30\text{--}60^\circ\text{N}$  in the annual mean (Figure 4.5A), with peak transport during winter (Figure 4.5B). During summer (Figure 4.5C) there is convergence of  $\text{CO}_2$  near  $50^\circ\text{N}$  due to poleward transport from the south and equatorward transport from the north. This convergence is coincident with peak biological uptake (see Figure 3.5B). Transport by transient waves from  $40\text{--}60^\circ\text{N}$  increases by  $\sim 0.08 \text{ PgC month}^{-1}$  from GEOS4-2.5x2 to GEOS5-0.67x0.5 in the annual mean, with differences between GEOS4 models and GEOS5 models  $\sim 0.02$  and  $\sim 0.04 \text{ PgC month}^{-1}$ , respectively. Transport differences are as large as  $\sim 0.15 \text{ PgC month}^{-1}$  from  $40\text{--}60^\circ\text{N}$  during winter and  $\sim 0.11 \text{ PgC month}^{-1}$  from  $60\text{--}70^\circ\text{N}$  during summer.

Transport by stationary waves near  $50^\circ\text{N}$  increases by  $\sim 0.05 \text{ PgC month}^{-1}$  from GEOS4-2.5x2 to GEOS5-0.67x0.5 in the annual mean (Figure 4.5A), with differences between GEOS4 models and GEOS5 models trivially small. Transport differences are as large

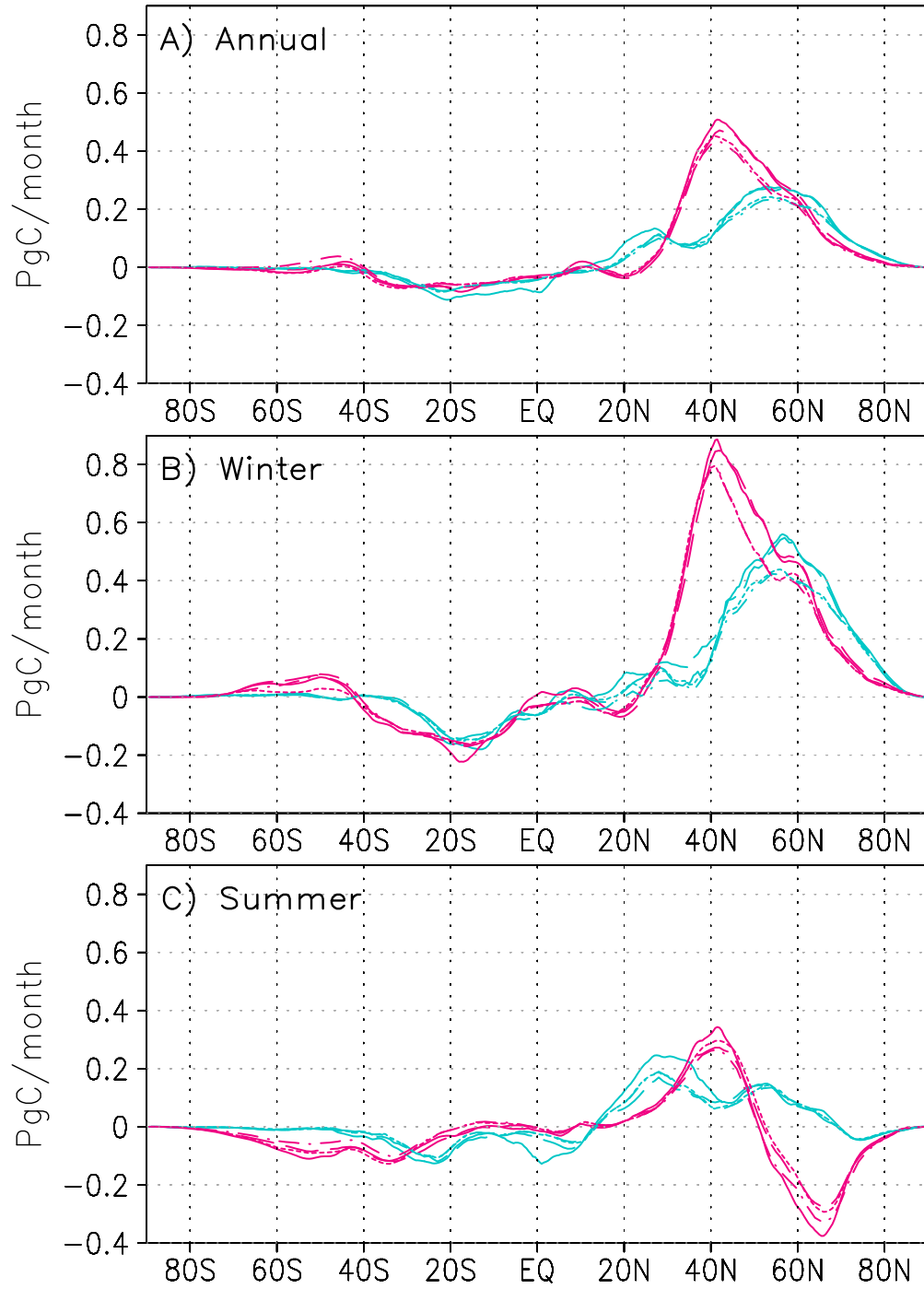


Figure 4.5: CO<sub>2</sub> transport by stationary waves (cyan) and baroclinic waves (magenta) along  $\theta_e$  in the annual mean (A) and during boreal winter (B) and boreal summer (C).

as  $\sim 0.15$  PgC month<sup>-1</sup> from 50-70°N during winter and  $\sim 0.11$  PgC month<sup>-1</sup> near 30°N during summer (Figure 4.5C). *Counter to intuition, stationary waves experience a*

*similar range of uncertainty as transient waves.* Although differences in total eddy transport between the GEOS5 models are small in the annual mean (Figure 4.3A), assessment of the transient and stationary wave components reveals that non-trivial differences exist in transient waves near peak amplitude ( $\sim 40^\circ\text{N}$ ) that largely cancel when added to stationary waves.

Eddy transport is stronger in the GEOS5 models, but comparison of horizontal structure reveals that geographic patterns and magnitudes of transport are very similar between models (Figure 4.6). In all models, patterns of eddy transport emerge over similar spatial scales (1000's km's) and regions. Comparison to a climatology of WCBs in Figure 3 of Eckhardt et al. [2004] suggests that dominant patterns of poleward eddy transport are coincident with climatologically favorable regions for WCBs. Like WCBs,  $\text{CO}_2$  transport tends to follow a northward and eastward track originating in subtropical latitudes, peaking in mid-latitudes, and terminating in the Arctic. The most favorable regions tend to originate over land (e.g., southern part of the United States and Tropical Asia) and terminate over the eastern portion of the adjacent ocean (e.g., Pacific and Atlantic Oceans, respectively). The major difference in models is largely a result of differing magnitudes of transport in favorable synoptic regions, for example over Europe (GEOS5 models dominate) and boreal Canada (GEOS5-1.25x1 dominates).

#### 4.4 Causes of Transport Uncertainty

Comparison of model characteristics helps explain uncertainty in eddy transport. Because GEOS4 models differ only in horizontal grid spacing, the difference between GEOS4-2.5x2 and GEOS4-1.25x1, although small, can be explained purely due to regridding of GEOS4-1.25x1 to the coarser grid of GEOS4-2.5x2. Differences in transport between GEOS5 models can also be explained through differences in grid spacing, although data assimilation procedures (meteorology is reanalyzed in GEOS5-1.25x1 but only analyzed in GEOS5-0.67x0.5) and model updates between 5.1.0 and 5.2.0 (high latitude diurnal cycle)

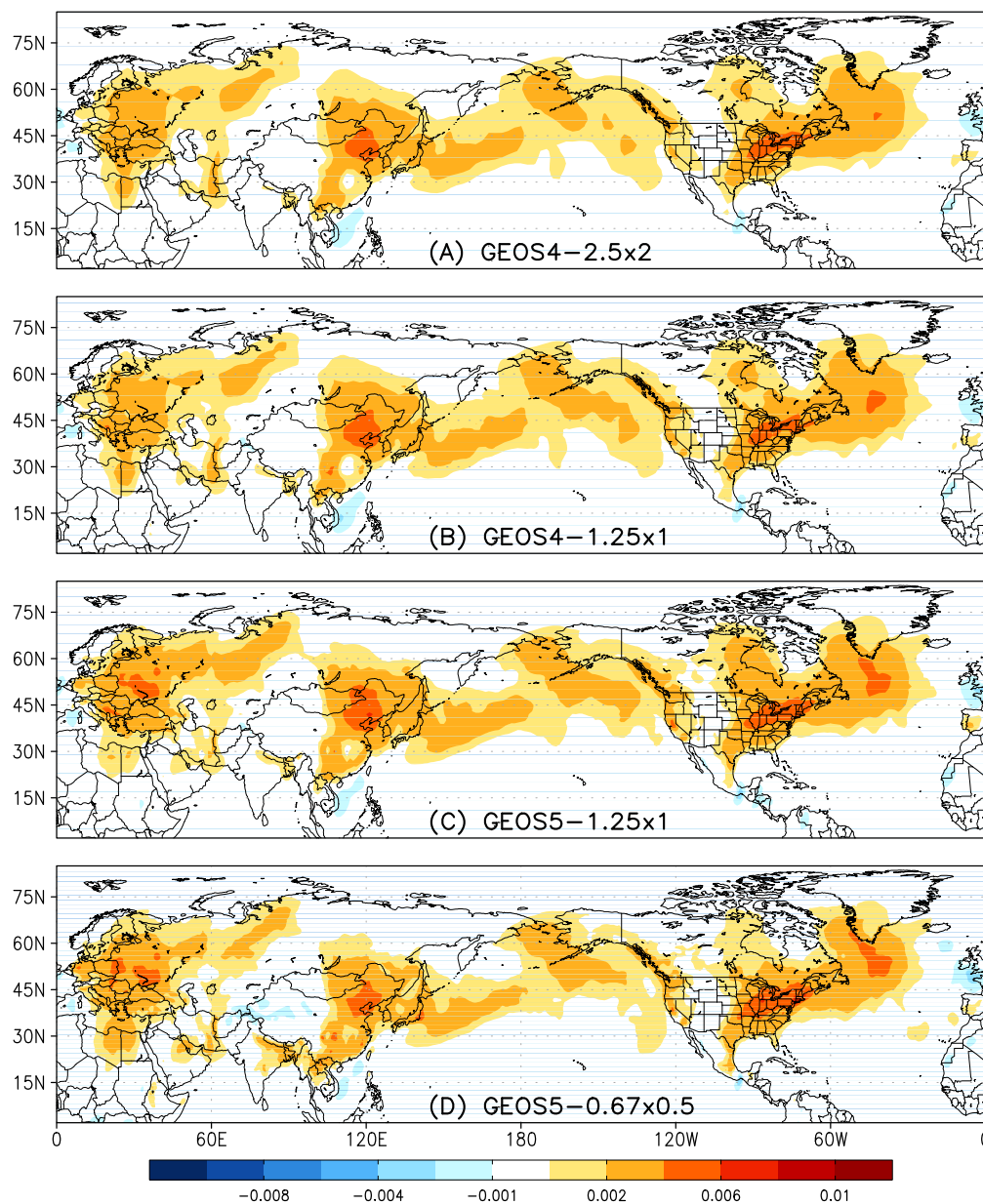


Figure 4.6: Spatial structure of annual mean column integrated  $\text{CO}_2$  transport by synoptic eddies along  $\theta_e$  for GEOS4-2.5x2 (A), GEOS4-1.25x1 (B), GEOS5-1.25x1 (C) and GEOS5-0.67x0.5 (D). The colorbar is shown in units of  $\text{PgC month}^{-1}$ .

may also contribute. Transport differences between GEOS4-1.25x1 and GEOS5-1.25x1 are not as straightforward because several major differences between GEOS4 and GEOS5 exist. First, although grid spacing is the same in GEOS4-1.25x1 and GEOS5-1.25x1, the native grid of GEOS5-1.25x1 is  $0.67^\circ \times 0.5^\circ$  (i.e., GEOS5-1.25x1 is run at  $0.67^\circ \times 0.5^\circ$  with transport

fields saved at reduced resolution). Second, there are differences in cumulus convection and cloud mixing parameterizations. Third, vertical resolution in the troposphere is nearly two times that of GEOS5. Fourth, the number and type of data assimilated per cycle is not consistent between models. Finally, data assimilation schemes are not consistent.

Model differences due to data assimilation are not possible to test in this study because transport simulations are run offline from parent models. Differences due to grid spacing, however, are more straightforward to test. An example of the effect of grid spacing is shown in Figure 4.7. As discussed in the Introduction, large-scale features associated with synoptic waves and “CO<sub>2</sub> weather” show up in both fine and coarse grid models. GEOS5-0.67x0.5 (Figure 4.7A), however, does a much better job resolving finer scale features and magnitudes. The most obvious difference between models in this example is in Western Siberia where wave-like features associated with stronger horizontal gradients in column CO<sub>2</sub> occur in GEOS5-0.67x0.5, but are strongly smoothed in GEOS4-2.5x2. A similar effect is also seen in the N. Atlantic. Additionally, CO<sub>2</sub> tends to build up more in the column in the Amazon in GEOS5-0.67x0.5.

Differences in vertical mixing due to cumulus convection and turbulent mixing are also straightforward to explore. Vertical mass fluxes are plotted as a function of  $\eta$  and averaged in northern mid-latitudes (30-70°N) in Figure 4.8. Turbulent mixing (Figure 4.8A) and cumulus convection (Figure 4.8B) in GEOS5 are consistently weaker than GEOS4 throughout the column in northern mid-latitudes, with cumulus mass flux nearly twice as weak. There are also smaller differences in vertical mixing between GEOS5-1.25x1 and GEOS5-0.67x0.5. Weak vertical mixing in mid-latitudes in GEOS5 is consistent with results published by Ott et al. [2009], who found that single column model application of RAS significantly underestimates convective mass flux relative to CRMs for several case studies of mid-latitude convective storms, resulting in weaker vertical transport of trace gases. The question is whether reduced vertical mixing in GEOS5 is a cause for enhanced meridional CO<sub>2</sub> transport from middle to polar latitudes.

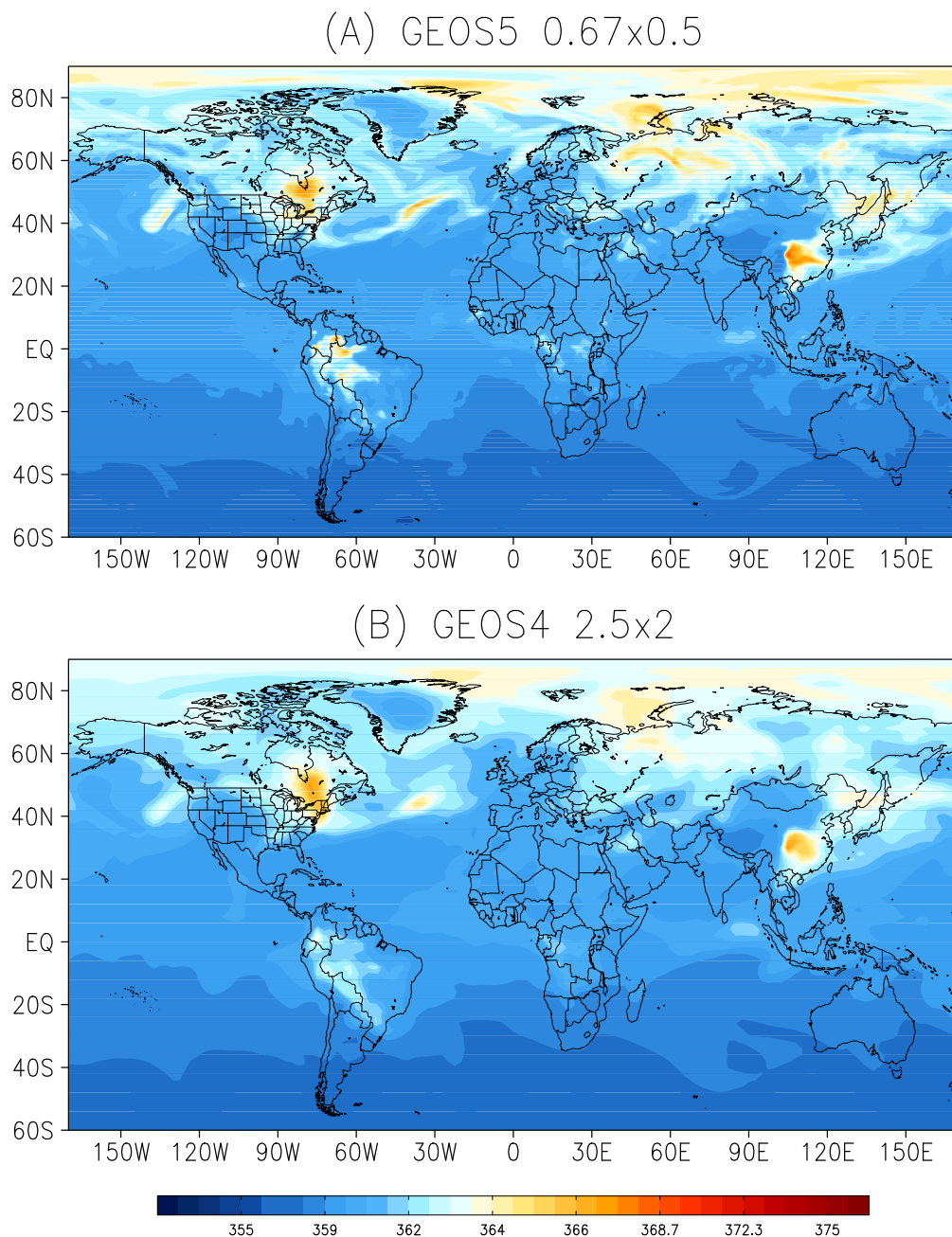


Figure 4.7: Column CO<sub>2</sub> snapshots on 10 January 2005 for GEOS5-0.67x0.5 (a) and GEOS4-2.5x2 (b)

Figure 4.9 shows model differences in annual mean CO<sub>2</sub> along  $\eta$  in northern mid-latitudes. The vertical gradient of CO<sub>2</sub> in the PBL from the surface to about  $\eta = 0.7$  is stronger in GEOS5, where GEOS5 exceeds GEOS4 near the surface by 0.5-1.0 ppm with model convergence of CO<sub>2</sub> in the free troposphere. Weak vertical mixing in GEOS5

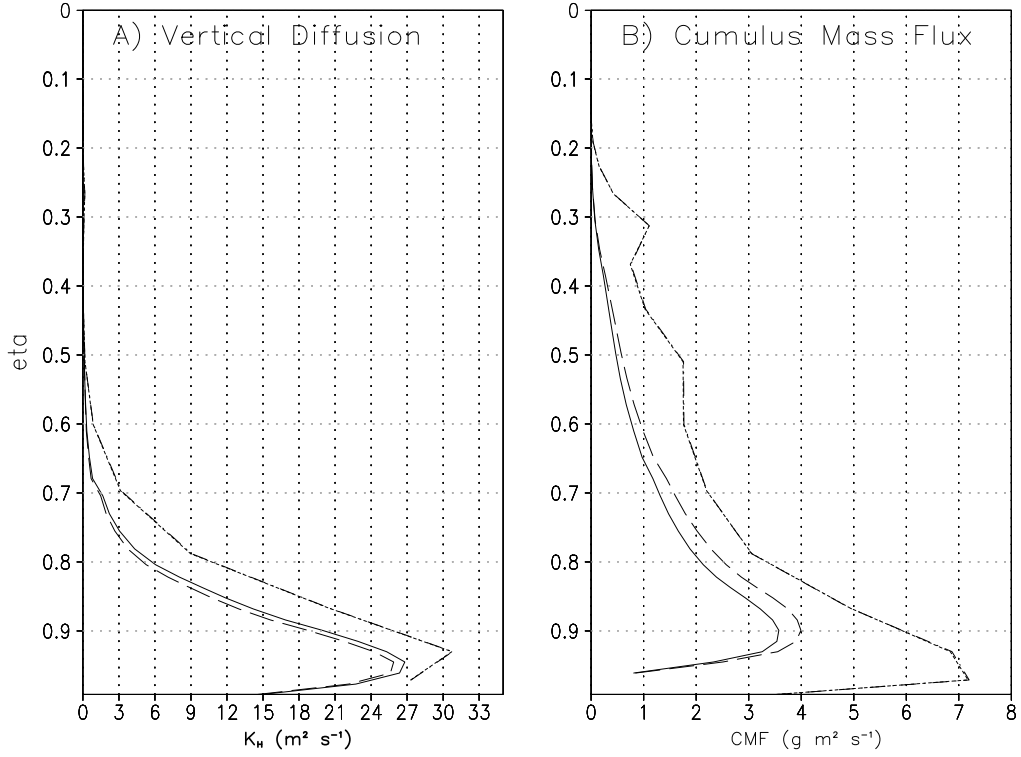


Figure 4.8: Zonal-annual mean values for vertical diffusion (A) and cumulus mass flux (B) binned into mid-latitudes (30-70°N) and plotted as a function of the terrain following coordinate eta ( $\eta$ ) for each reanalysis product. These values are taken directly from the driver files used to force vertical mixing in PCTM. GEOS4 lines are identical and lie directly over one another.

has likely trapped  $\text{CO}_2$  in the PBL and enhanced  $\text{CO}_2$  variability near the surface. This increased variability, in turn, enhances the likelihood of strong temporal and zonal perturbation to  $\text{CO}_2$ , hence driving increased eddy transport out of the PBL. Figure 4.10 shows that the difference in eddy transport between GEOS4-1.25x1 and GEOS5-1.25x1 lies primarily at lower levels, near  $\theta_e = 270$  K, indication that eddy transport is stronger in GEOS5-1.25x1 near the surface. *These results suggest that weak vertical transport by rapid convective events in GEOS5 enhances the probability of baroclinic eddies transporting  $\text{CO}_2$  upward and poleward out of the mid-latitude PBL.*

Differences in poleward flow of warm moist air out of the subtropics may also contribute to enhanced eddy  $\text{CO}_2$  transport in GEOS5. Moist ascent in mid-latitudes is implied in arguments by Pauluis et al. [2008, 2009] that there is a large poleward transport of warm



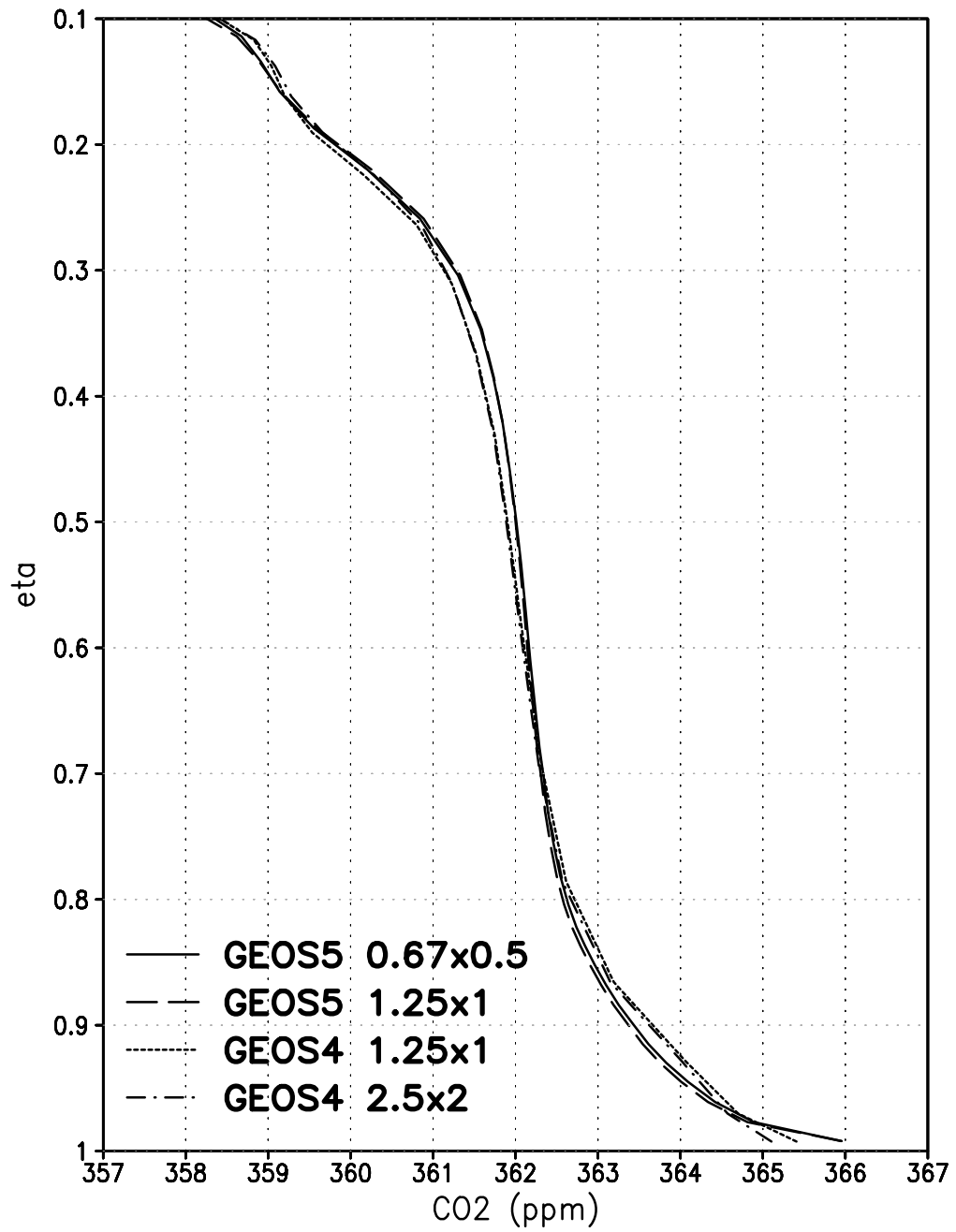


Figure 4.9: Zonal-annual mean CO<sub>2</sub> mixing ratio (ppm) binned into mid-latitudes (30-70°N) and plotted as a function of  $\eta$ .

moist air out of the subtropics into mid-latitudes where air parcels become convectively unstable and ultimately saturated through poleward and upward flow through stormtracks.

Figure 4.11 shows mass transport along three important airstreams of the general circula-

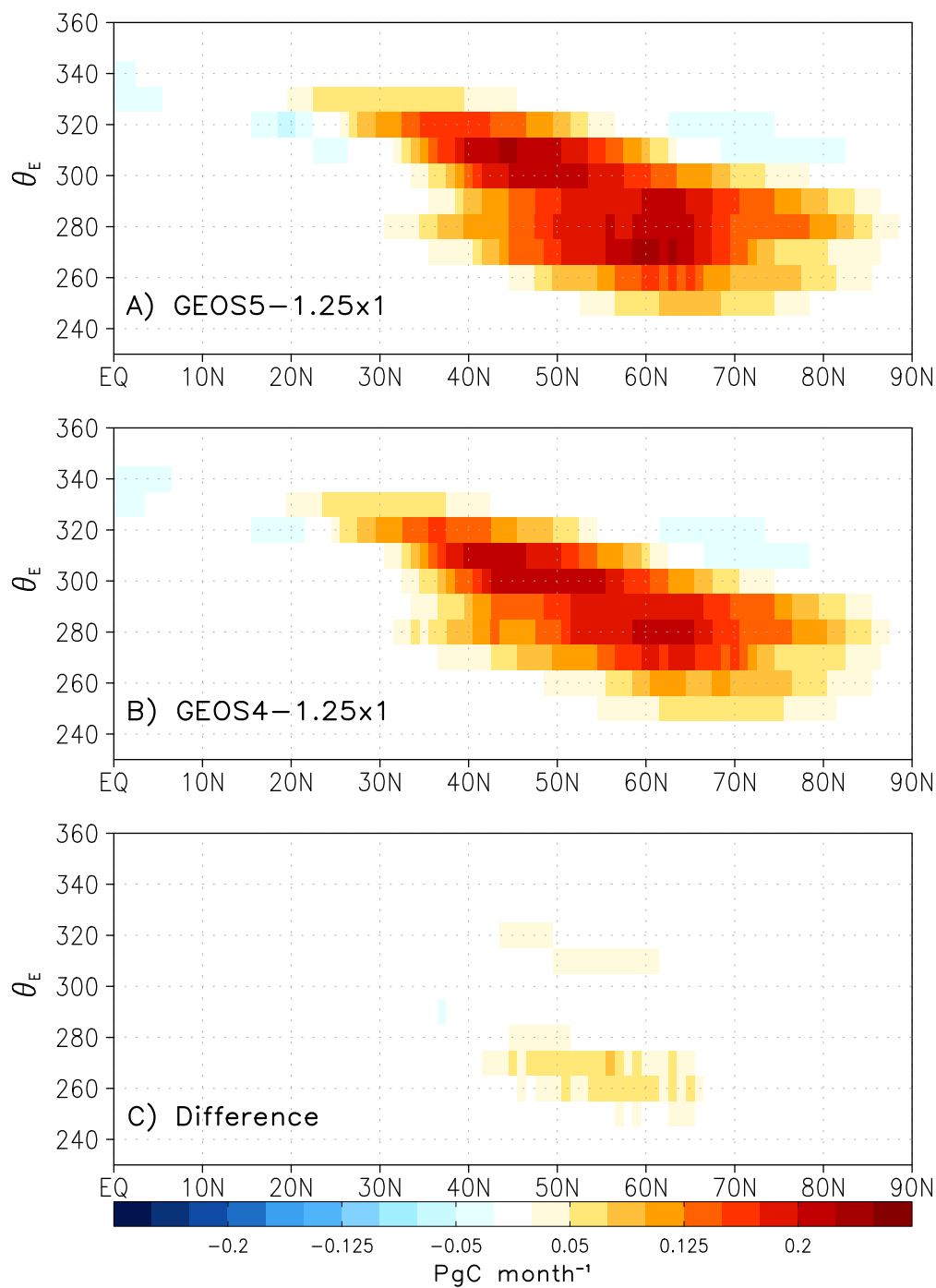


Figure 4.10: Eddy CO<sub>2</sub> transport by GEOS5-1.25x1 (A) and GEO4-1.25x1 (B). Difference in eddy transport between models are shown in (C).

tion: low-level flow of warm moist air (red line), low-level flow of cold dry air (blue line), and upper-level flow of warm dry air (black line). These airstreams are determined using

the joint distribution of mass transport of dry and moist isentropes as described in Eqs. 5a-5c in Pauluis et al. [2009] and add to zero at every latitude.

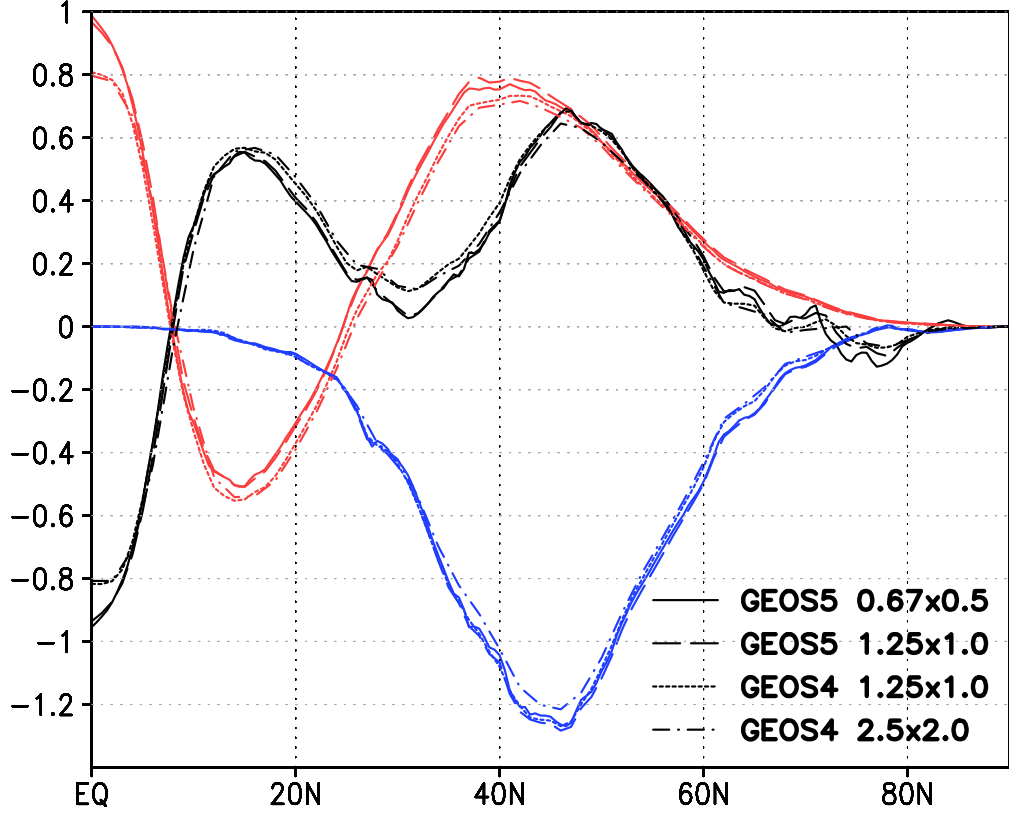


Figure 4.11: Mass transport of low-level cold dry air (blue line), low-level warm moist air (red line), and upper-level warm dry air (black line) as define in Eqs. (5a)-(5c) of Pauluis et al. [2009].

The most important result in Figure 4.11 is the model difference in low-level poleward flow of warm moist air near 40°N. This airstream supplies much of the energy necessary for moist slantwise ascent within baroclinic eddies. This analysis suggests that GEOS5 transports more mass poleward along moist airstreams than GEOS4, and thus the moist circulation is stronger in GEOS5. Strong poleward flow of warm moist air is compensated by strong equatorward flow of cold dry air. Since upward and poleward eddy transport of CO<sub>2</sub> is intimately tied to the moist circulation [e.g., Parazoo et al., 2011], a stronger moist circulation in GEOS5 (a likely result of new data assimilation techniques, number and type of data assimilated, and higher resolution parent grid) helps explain enhanced

eddy transport of  $\text{CO}_2$ . Similarly, the stronger moist circulation in GEOS5-1.25x1 relative to GEOS5-0.67x0.5 (Figure 4.11) given the same mean vertical  $\text{CO}_2$  profile (Figure 4.9) may help explain differences in transport by stationary and transient waves within GEOS5, with the difference in moist circulation explained by reanalysis of more observations in GEOS5-1.25x1 compared to GEOS5-0.67x0.5.

## 4.5 Discussion and Conclusions from Part 2

Eddy  $\text{CO}_2$  transport from a global tracer model is analyzed using identical surface fluxes and transport driven by four reanalysis products. Analysis on  $\theta_e$  reveals much uncertainty at synoptic scales. Differences in eddy transport at northern mid-latitudes reach  $0.1 \text{ PgC month}^{-1}$  in the annual mean and up to  $0.2 \text{ PgC month}^{-1}$  during boreal winter and summer. The weakest overall eddy transport occurs in the coarsest resolution model GEOS4-2.5x2, which reduces transport (relative to GEOS5-0.67x0.5) out of mid-latitudes by as much as  $0.5 \text{ ppm month}^{-1}$  and consequently reduces the eddy component of polar seasonality by as much as  $1 \text{ ppm month}^{-1}$ . The tendency for transport by moist synoptic storms to decrease as grid spacing and fine-scale vertical mixing increase suggests that systematic differences between analysis products may result in non-random tracer transport biases.

Additionally, in the context of fossil fuel emissions these transport biases are significant. Fossil fuels are prescribed using the dataset described by Andres et al. [1996], which has a total global emission of  $6.17 \text{ PgC year}^{-1}$ , or  $\sim 0.5 \text{ PgC month}^{-1}$ . With this dataset as a reference, an annual mean eddy transport bias of  $0.1 \text{ PgC month}^{-1}$  represents 20% of the global fossil fuel inventory and up to 40% when transport biases reach  $0.2 \text{ PgC month}^{-1}$  during winter and summer. The problem is even worse relative to fossil fuel inventories in northern mid-latitudes. This uncertainty in eddy transport puts even more pressure on regional inversions to carefully account for meridional advection at north and south boundaries when prescribing lateral boundary conditions.

In the same way that systematic model underestimates of nocturnal boundary layer mixing degrade the utility of nocturnal continuous measurements of  $\text{CO}_2$  for inverse flux estimates, biases in synoptic transport of  $\text{CO}_2$  might degrade the utility of continuous or satellite measurements of  $\text{CO}_2$  along fronts. Comparison to continuous measurements, for example, has shown that overall models in this study correctly simulate the phase and pattern of synoptic variability, but that the GEOS5 models in particular have the tendency during winter to overestimate the amplitude of synoptic variability. This is no different from nocturnal mixing; ability to simulate timing of mixing, inability to simulate strength. The likely implication for inverse flux estimation, assuming that continuous and/or satellite measurements sample enough points along synoptic weather systems to resolve amplitude and structure of horizontal gradients, is that the inversion will determine the correct distribution of sources and sinks but with the wrong magnitude.

It is more likely, as discussed in Chapter 3, that  $\text{CO}_2$  variations associated with synoptic moist dynamics will be mostly unobserved by satellites.  $\text{CO}_2$  transport at synoptic scales is therefore systematically unresolved *and* unobserved. Given the importance of transport by synoptic weather systems underneath clouds [Parazoo et al., 2008] and the resulting temporal sampling bias in  $\text{CO}_2$  mixing ratio distributions [Corbin et al., 2009], significant flux estimation errors are likely to be incurred due solely to unobserved transport, even if transport were perfect. It would be interesting to see how excluding column  $\text{CO}_2$  data from moist synoptic storm systems using some kind of cloud screening technique might bias inverted surface fluxes given OSSEs with biased and unbiased transport models.

It is important to keep in mind that total meridional transport in northern mid-latitudes is the same in all simulations and therefore reduced transport of  $\text{CO}_2$  by eddies in GEOS4 must be compensated for by reduced transport by mean circulations. This means transport by the mean circulation experiences the same total bias between GEOS4 and GEOS5 as transport by baroclinic waves, albeit for different reasons. Unlike transport by moist baroclinic storms, measurements of transport by the mean circulation do not

experience fair weather bias, and therefore uncertainty in mean transport is alleviated to some degree by continuous and flask observations at the surface, satellite measurements of column  $\text{CO}_2$ , and vertical profiling by aircraft. Because of the large uncertainty in mean transport and the huge opposing flows of  $\text{CO}_2$  throughout the column discussed in Chapter 3, these fair weather measurements are crucial for constraining flow through the dry part of the atmosphere.

A question raised by Pauluis et al. [2009] is whether midlatitude ascent occurs primarily through rapid convective events such as thunderstorms or slower slantwise motions in baroclinic eddies. Eddy decomposition on  $\theta_e$  strongly suggests that slantwise ascent plays a critical role in overall midlatitude ascent. Furthermore, the moist circulation is stronger in GEOS5 despite weaker cumulus mass flux. This raises questions as to what combination of factors sets the moist circulation in GEOS5 apart from that of GEOS4. Data assimilation techniques almost certainly play a role but it is likely finer grid spacing (horizontally and vertically) in GEOS5 is able to better resolve cyclonic circulations, frontal structure, and slantwise ascent along moist conveyors within baroclinic eddies. Eddy transport is only weakly sensitive to winds that have been regridded to coarse resolution (e.g., GEOS4-1.25x1 to GEOS4-2.5x2) or sampled from finer grids (e.g., grid points in GEOS5-1.25x1 are sampled from GEOS5-0.67x0.5). It is therefore more important to run models and generate reanalysis at the finest possible grid spacing in order to resolve fine-scale synoptic features. Future studies should examine sensitivity of eddy transport to grid spacing, with a fixed vertical mixing scheme, by running GCMs at multiple grid spacing from 250 km down to 50 km or finer if possible.

Eddy decomposition is a valuable tool for assessing tracer models via the statistics of eddy  $\text{CO}_2$  transport along the moist synoptic pathway. This technique could almost certainly be applied to other tracers to assess forward or inverse models. It could also be used to investigate the validity of moist convection in mid-latitudes in global models relative to observations and/or cloud resolving models. Even at 50 km grid spacing convective

scales are far from resolved. Ott et al. [2009] show that vertical mixing in GEOS5 is weak compared to a cloud-resolving model, but that parameters within the GEOS5 convection scheme could be optimized to improve vertical transport by convective storms. We have seen that weak vertical mixing in GEOS5 traps CO<sub>2</sub> in the PBL, where it accumulates over time. Eventually moist processes within baroclinic waves, unobserved by satellites, sweeps the CO<sub>2</sub> upward and poleward. Future studies should examine sensitivity of eddy transport within a global mesoscale model such as GEOS5 to vertical mixing, with other factors such as grid spacing and data assimilation held constant. We might find that a combination of schemes is needed to accurately represent vertical mixing in mid-latitudes. Also, the possibility of using a multi-scale modeling framework such as the Goddard-MMF [Tao et al., 2009] to explicitly represent moist convection and vertical mixing of tracer within GCM sized grid cells should not be overlooked.

## Chapter 5

# IMPLICATIONS OF TRANSPORT BIAS FOR INVERSE FLUX ESTIMATION

Part 1 (Chapter 3) showed that moist synoptic storms drive meridional sloshing of CO<sub>2</sub> between middle and polar latitudes, which in turn strongly modulates the seasonality of atmospheric mixing ratios of CO<sub>2</sub> in northern latitudes, and that much of this transport is unobserved by satellites. Part 2 (Chapter 4) showed that CO<sub>2</sub> transport by moist synoptic weather systems is significantly biased when driven by different weather products that are architecturally very similar but have key differences in horizontal grid spacing and the strength of fine-scale vertical mixing. Part 3 extends analysis in Parts 1 and 2 by using a top-down inversion approach to test for flux estimation errors due to systematic differences in transport between GEOS4 and GEOS5.

### 5.1 Introduction

The primary goal of CO<sub>2</sub> inversions is quantification of source and sink patterns at fine spatial and temporal scales that are as consistent as possible with CO<sub>2</sub> mixing ratio observations. Such inverse estimates provide independent information on flux distributions that can then be compared with “bottom-up” estimates. Inversions may also be used in an OSSE framework to provide specific information about instrument design needed to meet certain science objectives. For example, future satellite missions use OSSEs to determine specifics about orbit, averaging kernels (i.e., vertical weighting functions), and measurement precision that may be needed to measure certain flux signals (e.g., increasing emissions due



to permafrost melt) produced by the Earth system. Baker et al. [2010] is an excellent example of an OSSE designed to quantify carbon source and sink information that could be provided by column CO<sub>2</sub> measurements from OCO-2.

Unfortunately, inversion of observed or synthetic data must often assume perfect transport; that is, winds supplied to the atmospheric transport model are assumed to perfectly mimic the real atmosphere. Although not realistic, this assumption is often justified because there is no simple or standard way for inversions to account for transport errors. Furthermore, the assumption is reasonable if certain criteria are met. For example, transport errors that are more random than systematic could be assumed to average out over long enough time scales. If, however, transport errors are biased in some kind of systematic way (underestimated vertical mixing, phase difference of day-to-day variability due to synoptic processes, poorly represented PBL heights) then corresponding errors in flux estimation should also be expected (assuming other components of the inversion system are not biased in a different direction).

Chapter 4 set out to determine whether transport errors between GEOS4 and GEOS5 existed and to what degree errors were systematic and biased or just purely random. Comparison to surface based CO<sub>2</sub> mixing ratio observations showed good agreement between model and observations at synoptic time scales, with small spread between models. Net meridional transport by moist synoptic storms, however, was shown to be significantly degraded from GEOS5 to GEOS4. This so-called “eddy” transport component of the mid-latitude circulation was shown to be most strongly sensitive to (1) grid spacing, with enhanced transport on finer grids, and (2) the representation of fine scale vertical mixing by PBL turbulence and moist convection, with enhanced transport out of polluted boundary layers associated with weaker fine-scale vertical mixing. Transport errors at synoptic scales are significant, on the order of 0.1 PgC  $yr^{-1}$  in the annual mean and close to 0.2 PgC  $yr^{-1}$  in northern mid-latitudes during boreal winter and summer. Given that these reanalysis products are similar in architecture (similar parent GCM, data assimilation schemes, finite-

differences algorithms, etc.) and run through a common transport model, it is likely that these are conservative estimates for the types of transport bias to be expected.

The obvious question at this point is whether transport bias is aliased into errors in flux estimation vis-a-vis CO<sub>2</sub> inversion. Baker et al. [2010] tested the effect of transport errors on flux estimation by generating synthetic satellite retrievals with one transport model and inverting the retrievals with the same model but with stored meteorology offset by 18 hours. The result was significant degradation of 7-day flux improvements in the optimization. Houweling et al. [2010] also tested for the effect of transport errors on inversion of satellite data, but instead of using the same model offset in time, entirely different transport models were used to generate satellite data and to run the inversion. This technique better quantifies errors due to biased transport or systematic differences in transport. Although seasonal differences in column CO<sub>2</sub> were shown to be small ( $\sim 0.5$  ppm), the errors were correlated in space and caused flux estimation errors large enough that high precision satellite missions such as A-SCOPE might be limited in their ability to reduce uncertainties in flux estimates and meet science objectives of high accuracy at regional scales.

In this chapter, inversion techniques are utilized as in Houweling et al. [2010] to test for flux estimation errors that may result from transport bias. The primary objective is to generate synthetic satellite measurements of CO<sub>2</sub> using a model at fine resolution and inverting the synthetic data using a similar model at relatively coarse resolution. Control experiments are also needed to make sure that (1) the optimization technique, in this study MLEF (discussed in more detail in Chapter 2.6.3), is accurate and robust to random measurement errors, (2) enough satellite observations are available after cloud screening to provide a reasonable solution given realistic conditions, and (3) temporal sampling biases due to cloud screening of column CO<sub>2</sub> (discussed in Corbin et al. [2009]) don't significantly bias flux estimates. Control experiments will use the same transport model to generate and invert synthetic satellite data to eliminate possible errors due to transport bias. Given

the simplicity of these control cases, so-called “sink recovery” experiments are added to determine whether the inversion can detect realistic sinks hidden within satellite data.

In addition to transport bias, the accuracy of any inversion technique is limited by the optimization technique. Any given optimization algorithm searches for the best combination of surface fluxes that is consistent with observations by minimizing a cost function (see Equation 2.23). In ensemble data assimilation, the solution that best minimizes the cost function is obtained from a sample of the total range of possible solutions. The ensembles are therefore an approximation to the full analytical solution. More ensembles facilitate convergence of this “reduced rank” solution to the full solution but at greater computational expense. Since a major benefit of ensemble data assimilation is computational savings and the ability to process large amounts of data, tradeoff is needed between computational expense and accuracy. Zupanski et al. [2007] show that approximating the full solution does not have obvious detrimental impact on the solution and is worth the saved computational cost. The concepts of an “information matrix” and the DFS were discussed in Chapter 2.6.3 as guides for determining an appropriate number of ensembles. Based on this criteria and considering the large number of satellite data in this experiments, 200 ensembles and assimilation windows of 14 days were deemed appropriate for this experiment.

For details regarding the inversion framework, including prior estimates, the observation operator, description of the optimization technique and bias estimation strategy, covariance smoothing, flux and uncertainty calculations, and basic description of terminology, please refer to Chapter 2.6. Note that results in this chapter are presented either as net flux over the land (terrestrial NEE) or ocean (air-sea exchange). There are a total of six experiments in this study, with experiments becoming progressively more challenging for the inversion. The first two (Experiment 1 and 2, Section 5.2) are control experiments, designed to test for flux errors that arise given perfect transport and unbiased truth. The next two (Experiments 3 and 4, Section 5.3) are designed to test whether hypothetical sources and sinks can be recovered from satellite data, given perfect transport and biased

truth. The last two (Experiments 5 and 6, Section 5.4) test for flux errors that arise given biased transport and two different scenarios of truth.

## 5.2 Control Experiments - Perfect Transport

For simplicity, true fluxes are defined with  $\beta_{GPP} = \beta_{RESP} = \beta_{OCEAN} = 0$  so that truth is unbiased. Satellite data is generated using forward runs of GEOS4-2.5x2, and then inverted using ensembles of GEOS4-2.5x2 transport. In the absence of any errors in the ensemble approximation or significant temporal sampling biases after cloud screening, the inversion will know the satellite data is unbiased and return analyzed fluxes that are equal to the priors. Two experiments are run: one in which all satellite data is retained (Experiment 1), and one in which satellite data is retained only during sunny conditions (Experiment 2). In a real inversion clouds and aerosol in the atmospheric column contaminate measurements of column  $\text{CO}_2$  and these points must be discarded. Cloud contamination is assessed using cloud optical depth ( $\tau$ ) from MERRA, with data thrown out if  $\tau > 0.3$ . So, by introducing cloud screening in Experiment 2, the OSSEs are tested for sensitivity to (1) less observations and (2) observations that experience fair weather bias.

### 5.2.1 *Experiment 1: Perfect Transport, Unbiased Truth, and no Cloud Screening*

Annual mean errors in terrestrial NEE are shown at pixel scale for Experiment 1 in Figure 5.1. Since SiB is annually balanced ( $GPP = RESP$ ), prior estimates of terrestrial NEE are zero in the annual mean. Therefore, in a perfect inversion or the case that the ensemble solution has converged to the full Kalman solution, analyzed fluxes should also be zero in the annual mean. As can be seen, the inversion is not perfect, and non-zero, but nevertheless small, flux errors emerge. It is unlikely that these flux errors are due to data constraints considering the large size of the observation vector in this control case; they are more likely a result of the ensemble approximation. Fortunately, annual mean flux errors do have any obvious coherent spatial structure and could be interpreted as random in this

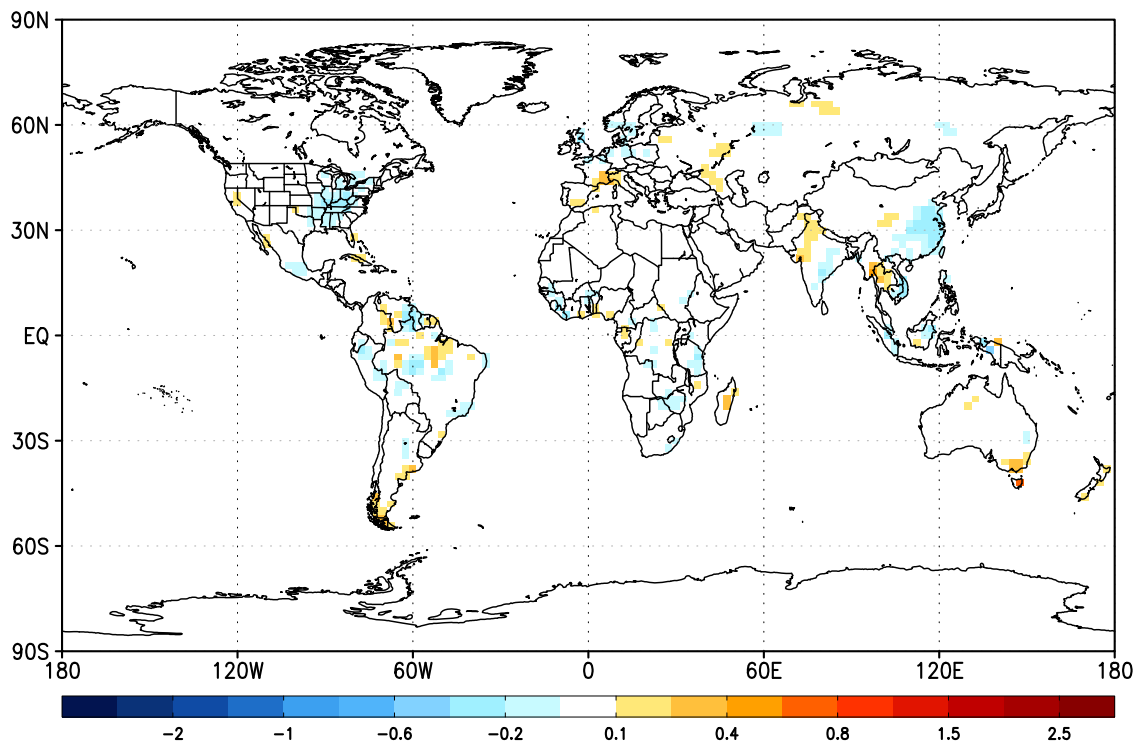


Figure 5.1: Annual mean errors in terrestrial NEE in units of  $\mu\text{moles m}^{-2} \text{s}^{-1}$  at pixel scale ( $2.5^\circ \times 2.5^\circ$ ) for Experiment 1.

experiment. These results are helpful confirmation that the inversion is stable and doesn't produce extremely biased flux estimates.

It could be argued that small these small annual mean flux errors are a convenient result of cancellation of larger seasonal errors. Seasonal flux errors are therefore plotted in Figure 5.2. For comparison to annual mean fluxes, seasonal fluxes are plotted on the same scale as in Figure 5.1. Larger differences between true and analyzed fluxes exist at seasonal time scales than in the annual mean, especially during boreal summer. The major feature that stands out in the seasonal plot is summer efflux out of boreal Asia and uptake in eastern N. America. These are clearly a non-random features that largely cancel in the annual mean, which is a slight cause for concern for seasonal inversions. Flux errors during the remaining seasons are much smaller and seemingly less systematic.

To get a sense of the relative size of flux errors, seasonal errors are compared directly

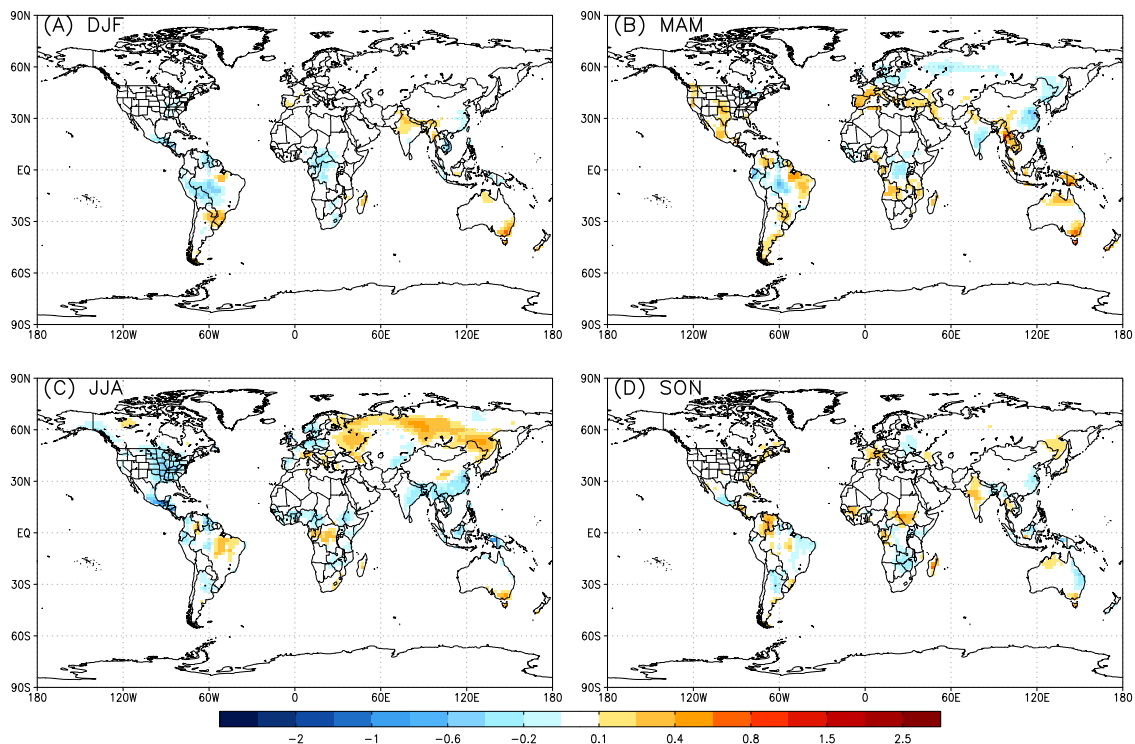


Figure 5.2: Similar to Figure 5.1 except plotted as a function of season for Experiment 1: (a) boreal winter, or December-January-February (DJF), (b) boreal spring, or March-April-May (MAM), (c) boreal summer, or June-July-August (JJA), and (d) boreal fall, or September-October-November (SON). The colorbar has the same scale as in Figure 5.1.

to priors (Figure 5.3) and divided by the priors (Figure 5.4) as a measure of percent error, where blue colors (values close to zero) indicate small errors relative to the priors and red colors (values close to one) indicate large errors. Together, Figures 5.3 and 5.4 demonstrate that flux errors are mostly small and trivial (blue shading), especially in northern mid-latitudes where seasonal NEE is strong. Comparison of red values in Figure 5.4 to priors in Figure 5.3 indicates that the largest relative errors tend to occur where priors are weak (deserts, savanna, mountains).

Prescribed and analyzed uncertainties for terrestrial NEE, along with uncertainty reductions, are calculated using the covariance matrix according to Equation 2.27 and plotted in Figure 5.5. The prescribed error covariance is calculated by assuming prior *GPP* and *RESP* components have uncertainties of  $\sigma = 0.2$ . This prior uncertainty on the component

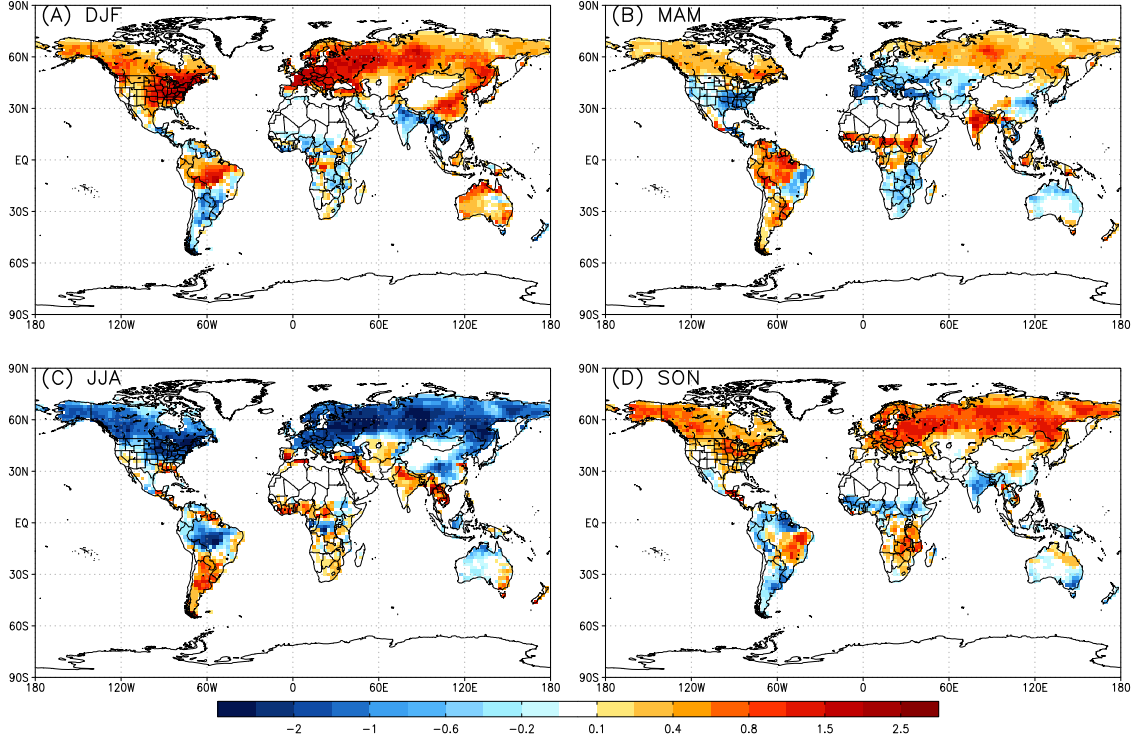


Figure 5.3: Seasonal distribution of prior estimate of terrestrial NEE, plotted on the same scale as Figure 5.1. These priors are used for all experiments in this chapter.

fluxes is somewhat generous, but is chosen such that the vast amounts of satellite data can help steer the optimized fluxes away from the priors. These large prior uncertainties will be especially useful when the inversions are confronted with real satellite observations. As seen in Figure 5.5A, the prescribed error covariance produces reasonable looking uncertainties over land regions that scales with terrestrial NEE. As expected, the largest uncertainties occur over tropical regions where component fluxes of *GPP* and *RESP* are largest. Posterior uncertainty for a single grid point is large (relative to the mean net flux) relative to uncertainty aggregated up to larger regions. This is because pixels tend to be anti-correlated over larger regions, the effect of which is to reduce overall uncertainty.

As discussed in Lokupitiya et al. [2008], cross-correlations in the covariance matrices become diagonally dominant after a few cycles, despite strong smoothing at the first cycle, as error covariances learn from the observations. Lokupitiya et al. [2008] also showed that

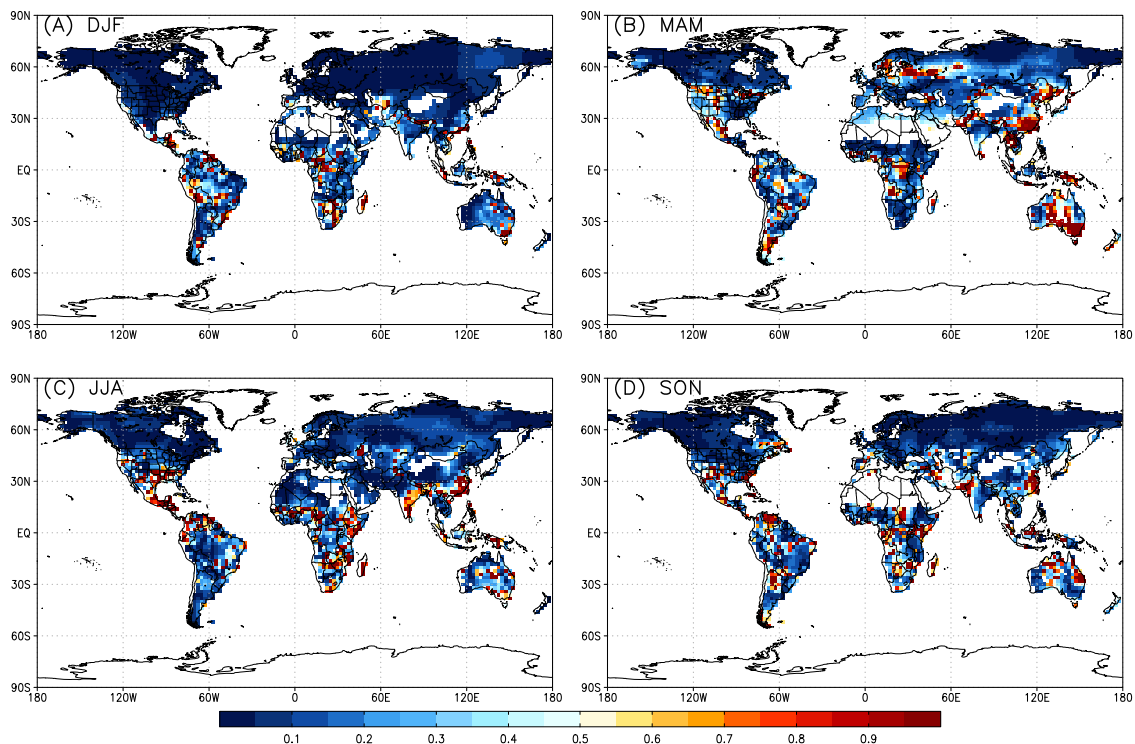


Figure 5.4: Seasonal flux errors for Experiment 1 plotted as a fraction of the prior.

variance (diagonal terms of covariance matrix), and hence uncertainty, is most strongly reduced in regions where observations are dense. When constrained by surface observations, the analyzed error covariance is most strongly reduced in northern mid-latitudes. As can be seen in Figure 5.5 B and C, however, uncertainties are strongly reduced at all latitudes in the repeating process of assimilation of satellite observations and propagation of error covariance forward in time to the next assimilation cycle. According to Figure 5.5C, uncertainty is reduced by about 70-80% globally. Keeping in mind the simplicity of the first experiment, this is a very encouraging result.

Annual mean flux errors over the ocean are shown in Figure 5.6. Unlike in SiB, ocean fluxes are not annually balanced and thus priors and truth are non-zero in the annual mean (see Figure 5.7). As was the case for terrestrial NEE, inversion for air-sea exchange is not perfect, but errors are small and tend to scale with annual mean ocean flux priors. It is possible that small ocean flux errors result from ocean fluxes being too weak to be



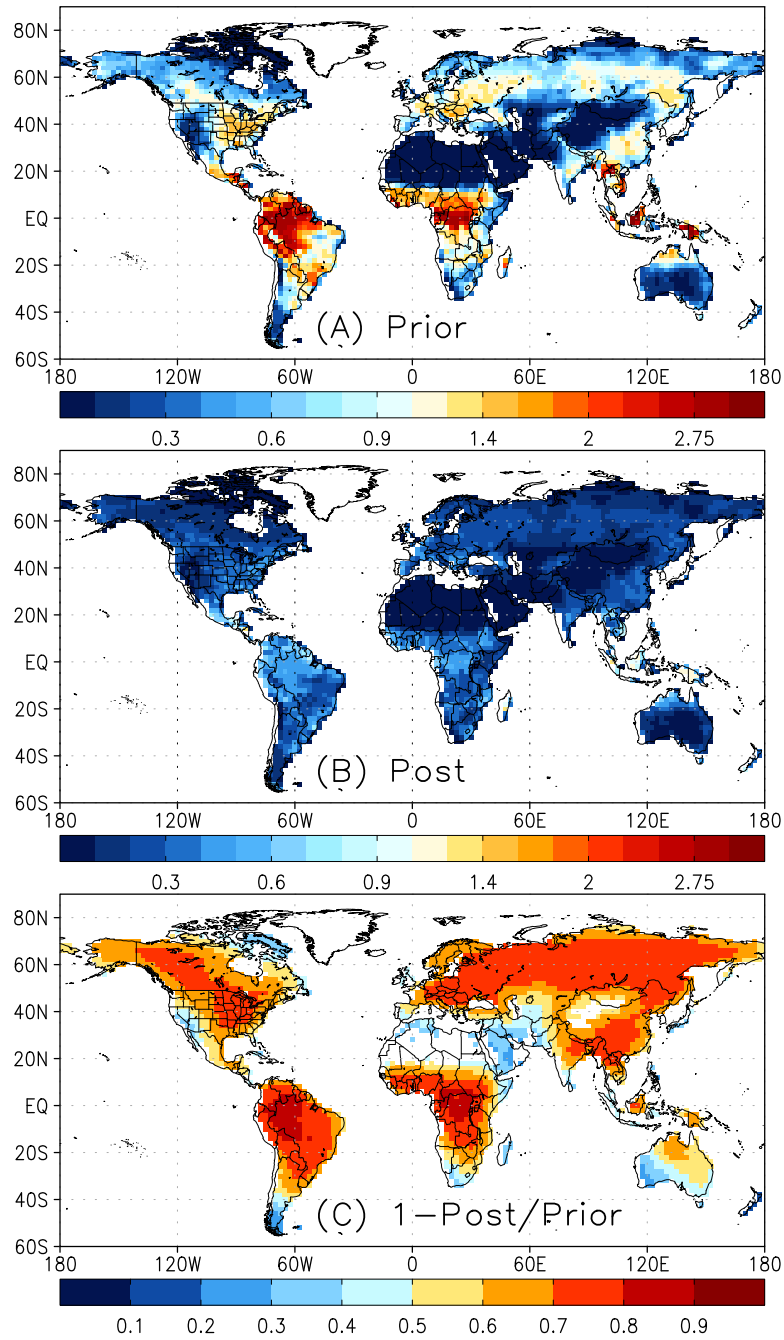


Figure 5.5: Annual mean uncertainties in Experiment 1 for (a) prior and (b) analyzed flux. Uncertainty reduction for terrestrial NEE is shown in the bottom plot (c).

detected in the satellite data by the inversion, and that analyzed fluxes over the ocean are constrained more by the priors than satellite data. This phenomenon will be discussed in more detail later in Section 5.3 when the inversion is forced to recover hypothetical biases

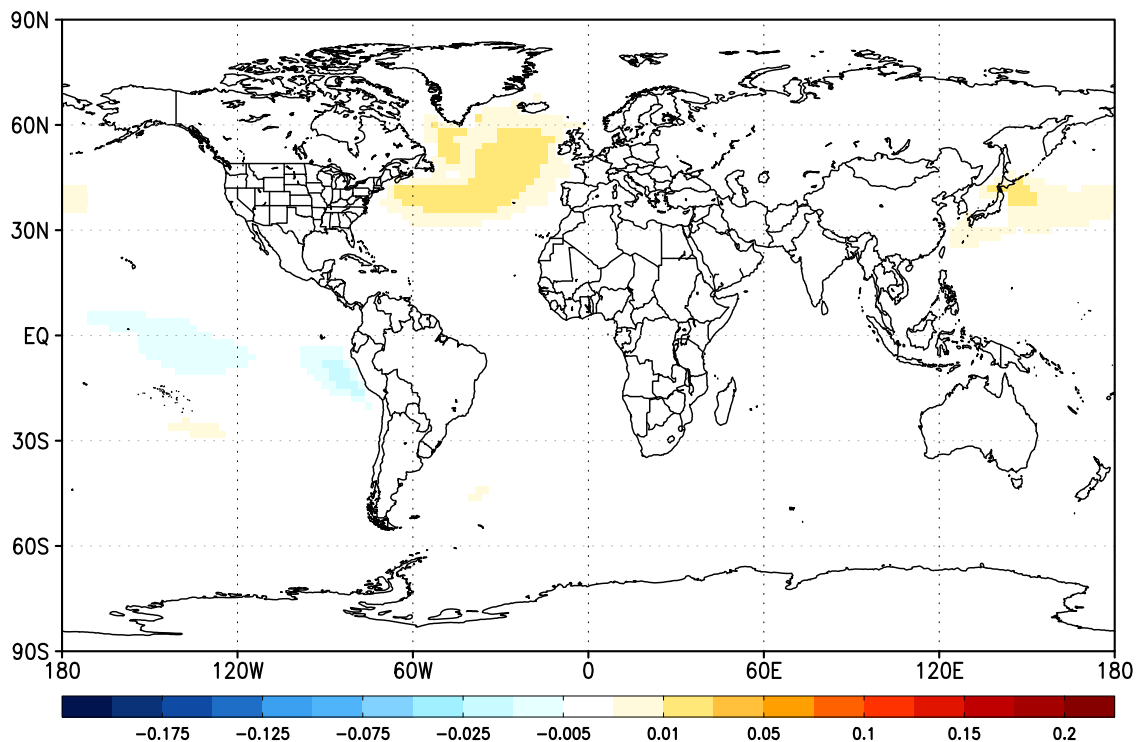


Figure 5.6: Same as Figure 5.1 but for air-sea exchange for Experiment 1. Note difference in scale between land and ocean.

in air-sea exchange.

As with terrestrial NEE, it is possible that small residual fluxes in the annual mean are a result of cancellation of larger seasonal fluxes. Seasonal flux errors are therefore plotted in Figure 5.8. No large cancellation between seasons is apparent, most likely because seasonality of air-sea exchange is very weak (i.e., there is little change in magnitude or sign of fluxes within a given region). These results show that annual mean errors are dominated by net efflux in the northern Pacific and Atlantic during boreal winter and spring. Comparison to seasonal prior air-sea exchange, plotted in Figure 5.9, shows that the winter and spring errors in the north Pacific and Atlantic scale slightly with an uptake feature in the Gulf Stream. Air-sea exchange errors are divided by ocean priors in Figure 5.10. Ocean flux errors are mostly small and trivial (blue shading), with the largest errors approximately 20% of the prior. Errors are smallest where ocean exchange is strongest, and largest where ocean exchange is weakest.

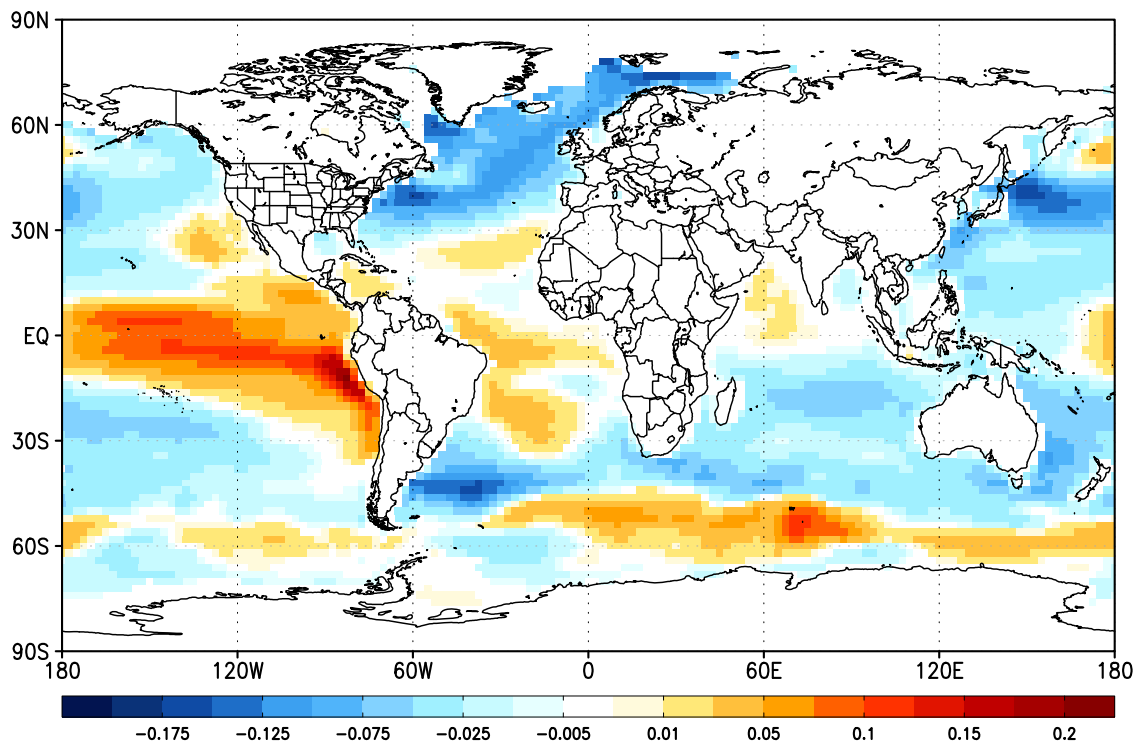


Figure 5.7: Same as Figure 5.6 but for prior air-sea exchange.

Prescribed and analyzed uncertainties for air-sea exchange, along with uncertainty reductions, are calculated using the covariance matrix according to Equation 2.29 and plotted in Figure 5.11. The prescribed error covariance is calculated by assuming ocean priors have an uncertainty of  $\sigma = 0.1$ . The prescribed uncertainty over the ocean is half that of *GPP* and *RESP* because ocean fluxes are much weaker than terrestrial NEE and therefore more difficult to detect in satellite data. As a consequence, and more importantly by design, ocean flux estimates tend to be more heavily weighted towards the priors. The prescribed error covariance produces reasonable looking uncertainties over ocean regions that scales with the largest air-sea exchanges (Figure 5.11A). Although not comparable in reduction to terrestrial NEE, and despite the weaker flux signal, uncertainties over the ocean are still strongly reduced at global scale in the process of assimilation of satellite observations (Figure 5.11B). According to Figure 5.5C, uncertainty is reduced by 40-50% globally. This result is encouraging, but not as much so as the uncertainty reduction for terrestrial NEE.

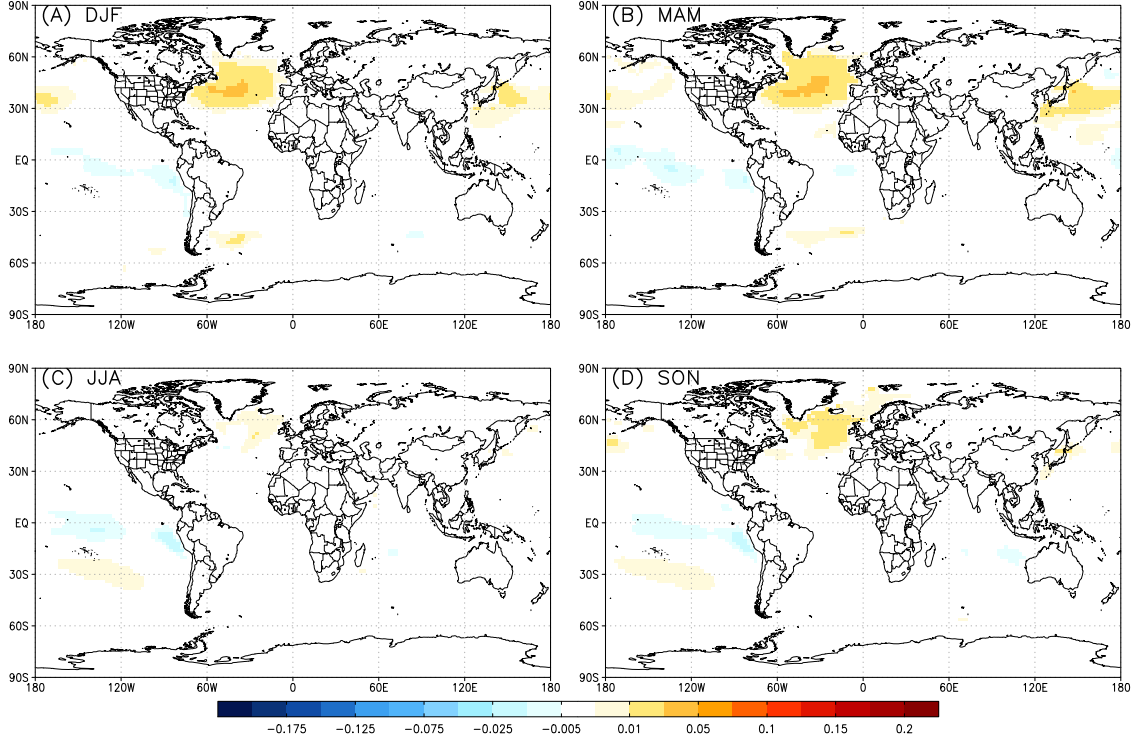


Figure 5.8: Similar to Figure 5.6 for Experiment 1, except plotted as a function of season (a) DJF, (b) MAM, (c) JJA, and (d) SON. The colorbar has the same scale as in Figure 5.6.

### 5.2.2 Experiment 2: Perfect Transport, Unbiased Truth, and Cloud Screening

Experiment 2 is the same as Experiment 1 except that satellite data is screened for clouds. More importantly, column  $\text{CO}_2$  data from moist synoptic storm systems is excluded. This means that a large number of points are discarded and the observation constraint is greatly reduced from Experiment 1 (e.g., see Figure 2.3). Cloud screening is performed by sampling forward model runs only when  $\tau_{\text{cld}} < 0.3$ . While it is more realistic to use observed cloud distributions, here  $\tau_{\text{cld}}$  is prescribed from MERRA to be consistent with weather driver data used in this study (see Chapter 2.6.2). Note that aerosols also contaminate satellite measurements but are not considered in this study.

Annual mean errors in terrestrial NEE are shown at pixel scale for Experiment 2 in Figure 5.12. The largest errors are incurred over the tropics where cloud screening eliminates most of the data (see Figure 2.2B and 2.3). Flux errors are actually reduced

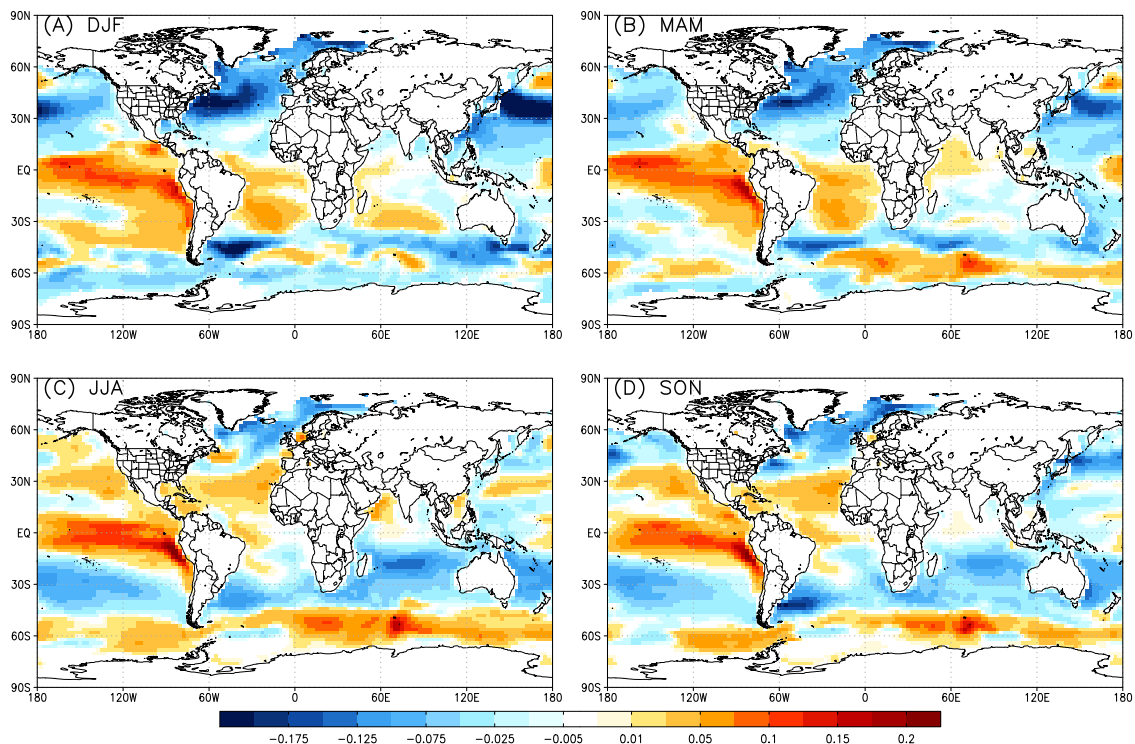


Figure 5.9: Same as Figure 5.8 except for air-sea exchange priors.

over N. America relative to Experiment 1, but become more positively biased in Europe. Flux errors due to cloud screening are trivial in the annual mean.

Seasonal errors (Figure 5.13) are by comparison much larger and therefore nontrivial. Tropical regions experience large errors during all seasons. Europe experiences large positive bias during boreal spring and summer. Large seasonal errors occur over N. America (positive bias during summer, negative bias the remainder of the year) that cancel in the annual mean. Boreal and temperate Eurasia continue to experience very little error throughout the year. In general, seasonal errors in Experiment 2 are more spatially coherent than in Experiment 1. Data reduction hasn't simply relaxed the solution to the priors. *Fair weather bias in measurements have steered the solution away from internal errors due to ensemble approximation, most notably toward positive bias during summer in northern temperate and boreal latitudes.*

Uncertainty is not as strongly reduced in Experiment 2 (50-60% globally, see Fig-

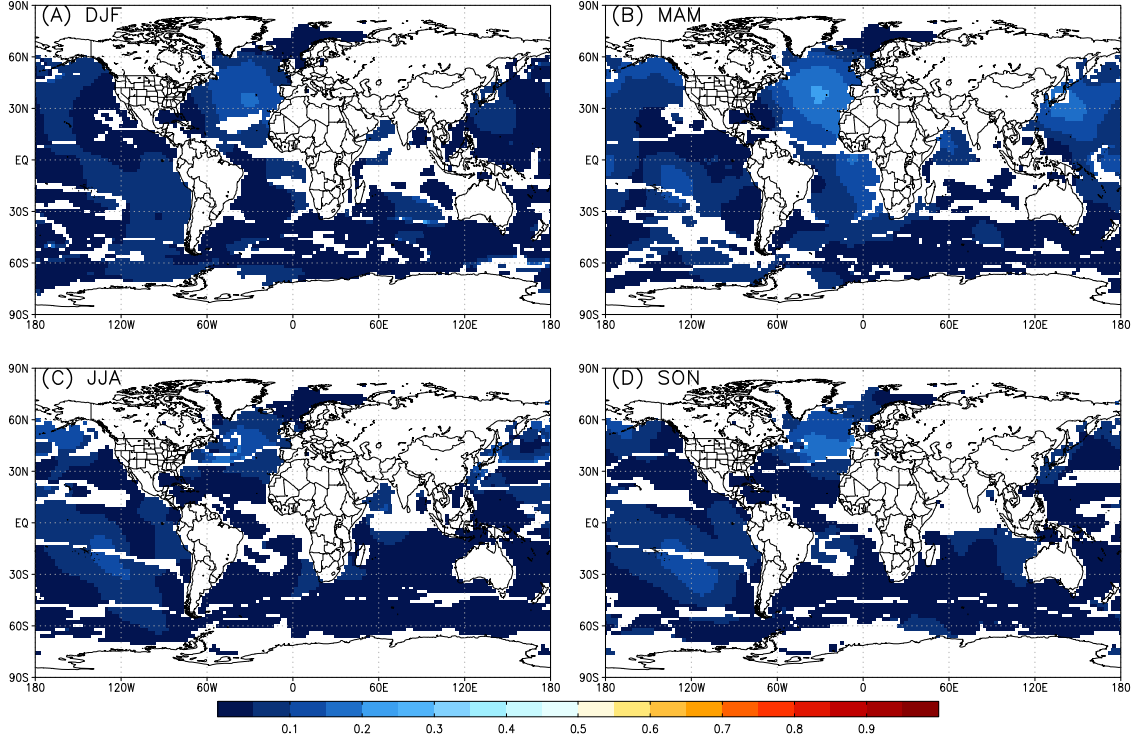


Figure 5.10: Seasonal errors in air-sea exchange in Experiment 1 plotted as a fraction of the prior.

ure 5.14B and C) because there is much less data for the information matrix in Equation 2.25 to learn from. However, considering the huge reductions of satellite data, especially in the amazon where clouds are prevalent all year, substantial reductions in uncertainty are nevertheless achieved. Flux errors over the ocean are actually much weaker than when cloud screening isn't applied (also not shown). This is most likely because there is so little information over the ocean (e.g., Figure 2.2) that the inversion is forced to relax to the priors.

### 5.2.3 Aggregation of Experiments 1 and 2 up to TransCom Regions

Flux errors (and uncertainties) are often correlated in space due to (1) spatial smoothing applied to the covariance matrix during the first assimilation cycle, (2) transport patterns that emerge from the prevailing wind, and (3) spatial coherence of vegetation type.

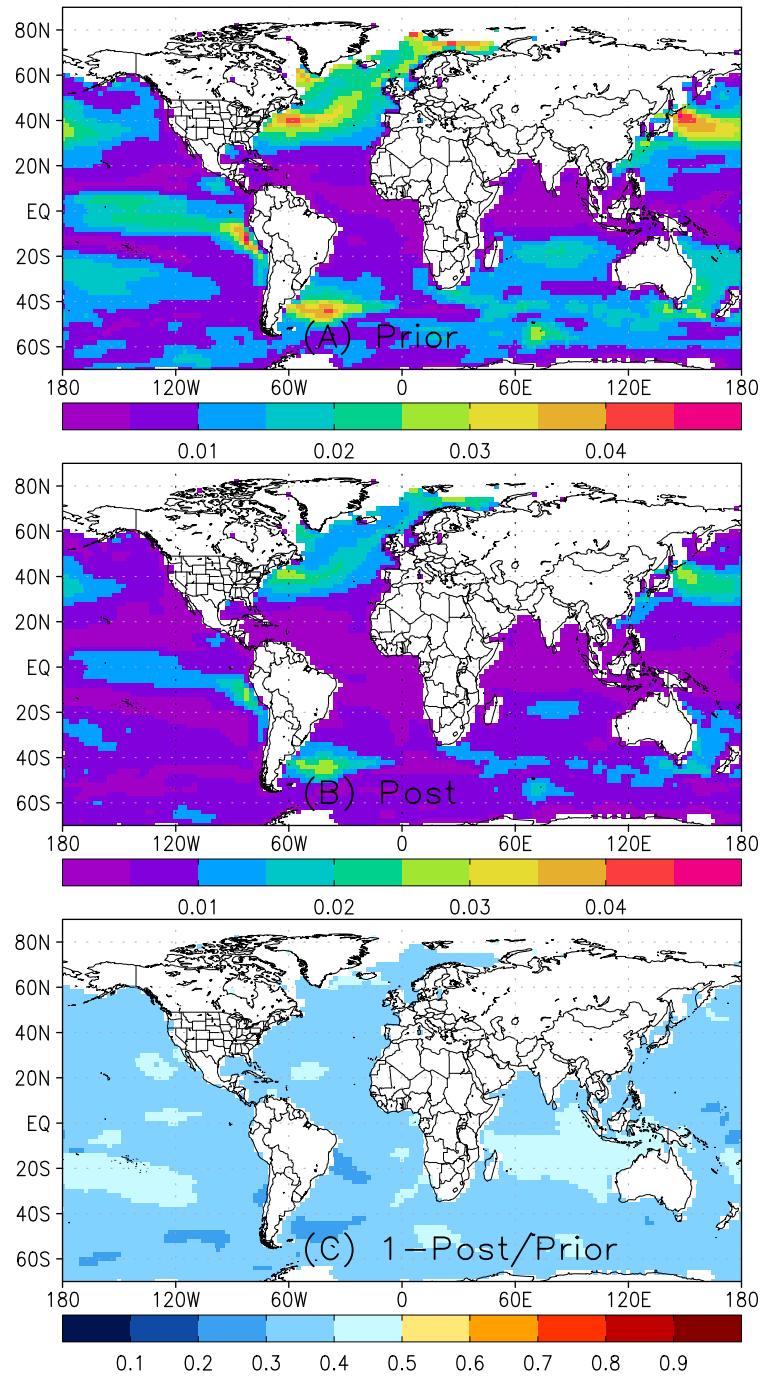


Figure 5.11: Annual mean uncertainties in Experiment 1 for (a) prior and (b) analyzed air-seas exchange. Uncertainty reduction is shown in the bottom plot (c).

Uncertainties are also highly anti-correlated in space so it is useful to aggregate up to larger regions to reduce the uncertainty of flux errors. A common practice for global scale inver-



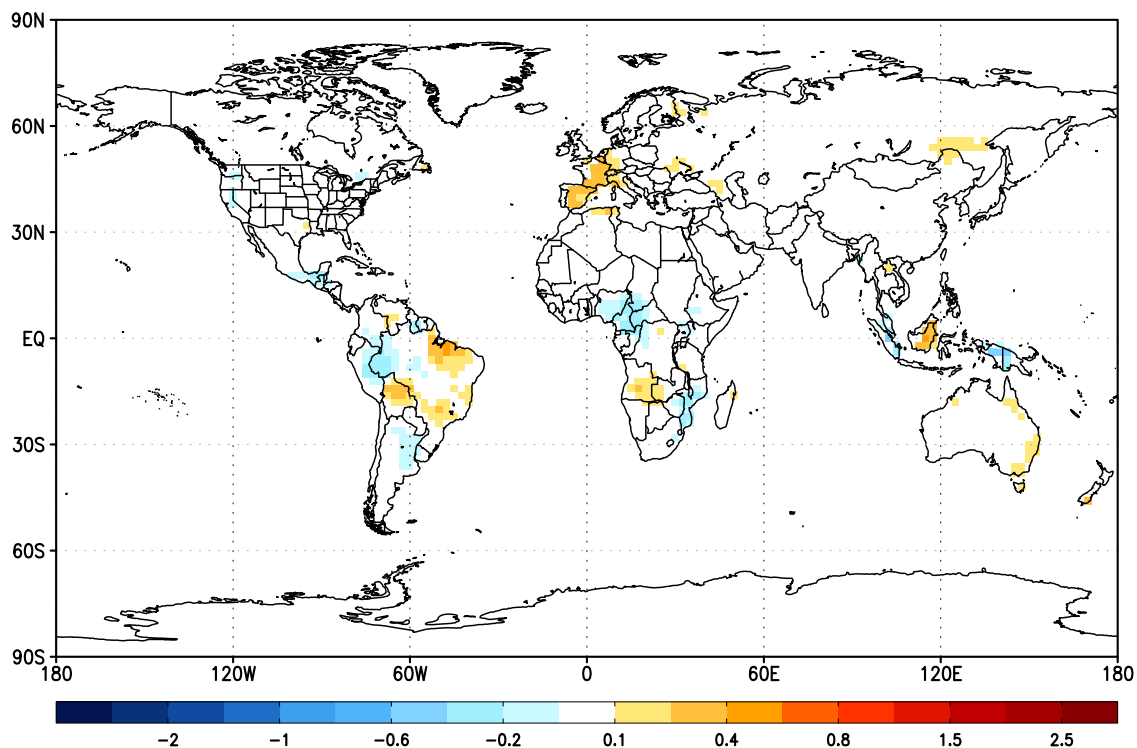


Figure 5.12: Same as Figure 5.1 except for Experiment 2.

sion experiments is to aggregate fluxes up to TransCom-sized basis functions that reflect both geographical and mechanical elements (see Figure 5.15). In this way, flux errors over large regions can be related to realistic land cover type classification.

Annual flux and uncertainty estimates for Experiments 1 and 2 are aggregated up, binned into TransCom regions (11 land and 11 ocean), and converted to  $\text{GtC year}^{-1}$  in Figure 5.16 according to Equations 2.30-2.33. Note that since total annual prior (also truth in this case) terrestrial NEE is zero at every pixel (SiB3 balances *GPP* and *RESP* in an annual cycle), aggregation up to TransCom regions also results in zero net flux. Total annual prior air-sea exchange is non-zero at pixel and TransCom scale.

In Experiment 1, recovered fluxes of terrestrial NEE are close to zero when aggregated up, never more than  $0.15 \text{ GtC year}^{-1}$  away from zero and, except for Australia, always within  $1 \sigma$  of the truth. This is good; the simplified inversion does not create fluxes that are significantly biased from the truth. The same is true for air-sea exchange, which has



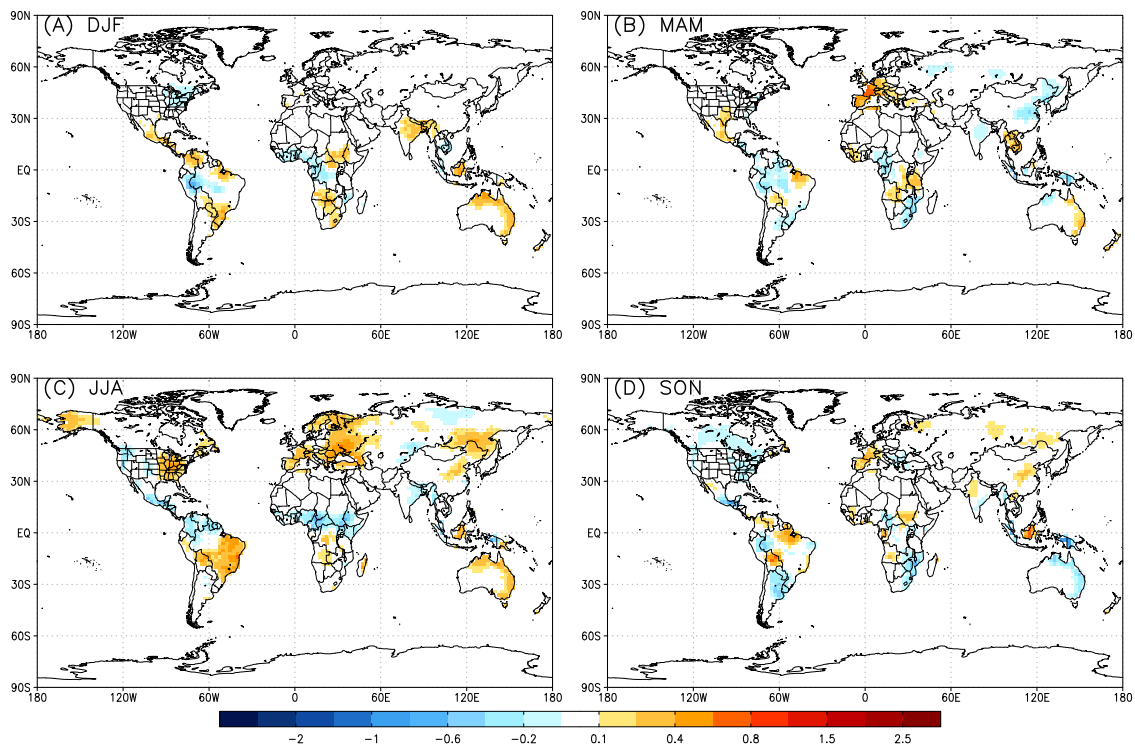


Figure 5.13: Flux errors plotted as a function of season for Experiment 2: (a) boreal winter, or December-January-February (DJF), (b) boreal spring, or March-April-May (MAM), (c) boreal summer, or June-July-August (JJA), and (d) boreal fall, or September-October-November (SON). The colorbar has the same scale as in Figure 5.12.

an even tighter uncertainty. Overall, there tends to be negative bias in terrestrial NEE in Experiment 1, which is dominated by negative biases in the tropics and Temperate N. America.

Cloud screening in Experiment 2 doesn't change annual TransCom sized fluxes significantly, but there are some very noticeable and important differences from Experiment 1. First, reduction of satellite data increases posterior uncertainty of flux estimates. While fluxes in most regions are within  $1\sigma$  of the truth, flux errors in Australia and Europe reach  $\sim 0.1 \pm 0.115$  and  $0.25 \pm 0.245$  GtC year<sup>-1</sup>, respectively, and are  $1\sigma$  away from the truth. N. Africa is within  $1\sigma$  of the truth but incurs an error of nearly  $0.2 \pm 0.350$  GtC year<sup>-1</sup> due to cloud screening. Ocean flux estimates are closer to the truth in Experiment 2 than Experiment 1 because data reduction has given more weight to priors. The total land flux

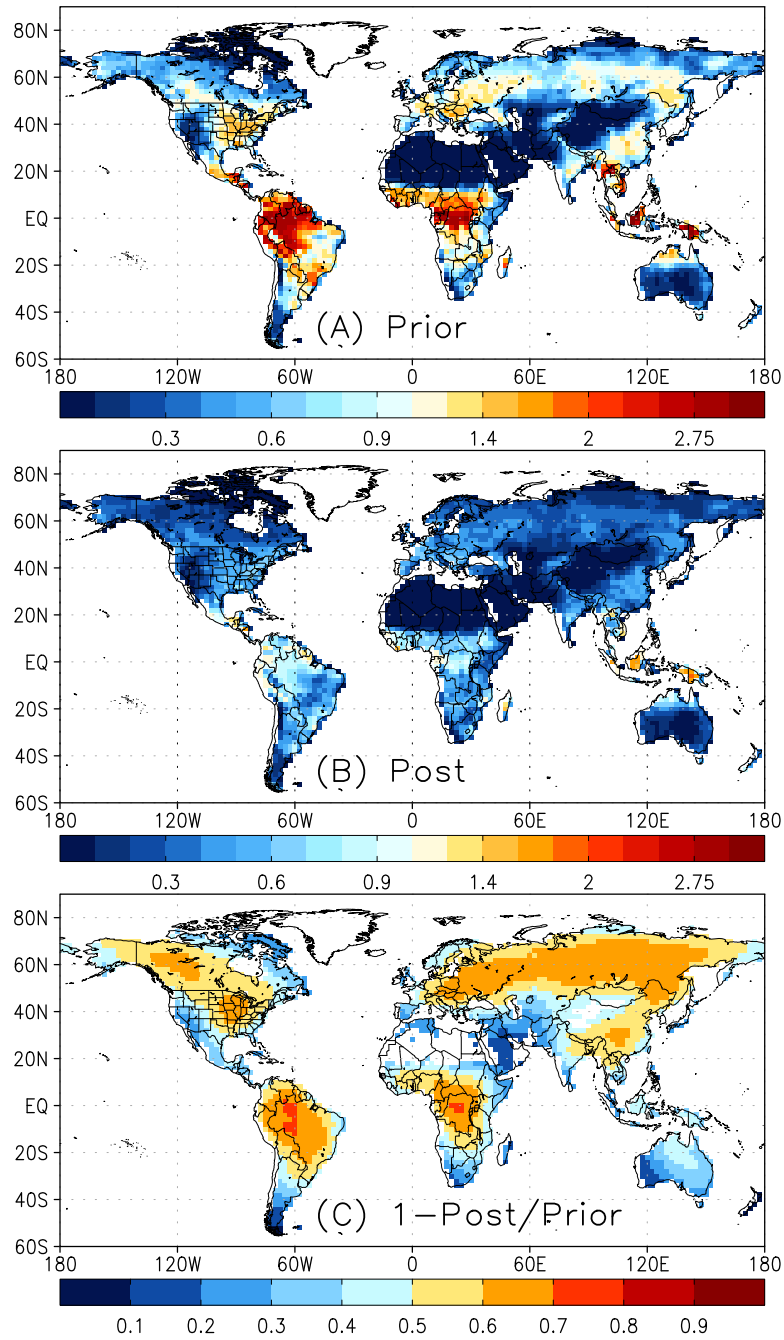


Figure 5.14: Annual mean uncertainties in Experiment 2 for (a) prior and (b) analyzed flux. Uncertainty reduction for terrestrial NEE is shown in the bottom plot (c).

has a positive bias of  $0.3 \pm 1.2 \text{ GtC year}^{-1}$ , which is about 10% of the total annual global sink. The total land bias is dominated by positive bias in Europe, Eurasian Boreal, and Australia.

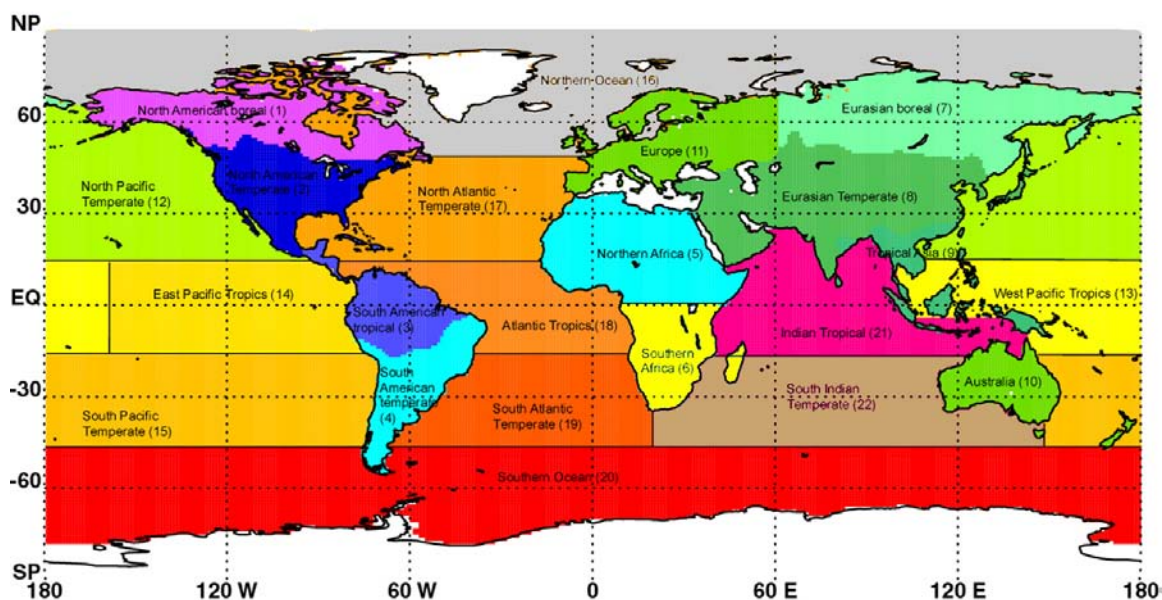


Figure 5.15: TransCom basis functions.

The effect of removing column  $\text{CO}_2$  data from moist synoptic storms in mid-latitudes and day-to-day weather in the tropics through cloud screening is more obvious in seasonal plots of flux errors from Experiments 1 and 2 (Figure 5.17), where Experiment 1 is plotted in green and Experiment 2 in blue to be consistent with Figure 5.16. The pattern for northern boreal and temperate latitudes (Boreal and Temperate N. America, Eurasian Boreal and Temperate, and Europe) is for positive bias during summer (consistent with Figure 5.13) relative to Experiment 1 and to net zero error, negative bias during shoulder seasons, and little to no bias during winter. The strongest positive bias, 0.1 PgC, occurs during summer in Europe. The largest errors at northern latitudes also tend to occur during summer, suggesting that cloud screening becomes more detrimental for flux estimation during the growing season when component fluxes are larger. While there are no obvious patterns for flux errors in the tropics, flux errors are nevertheless larger on average at monthly time scales when cloud screening is applied.

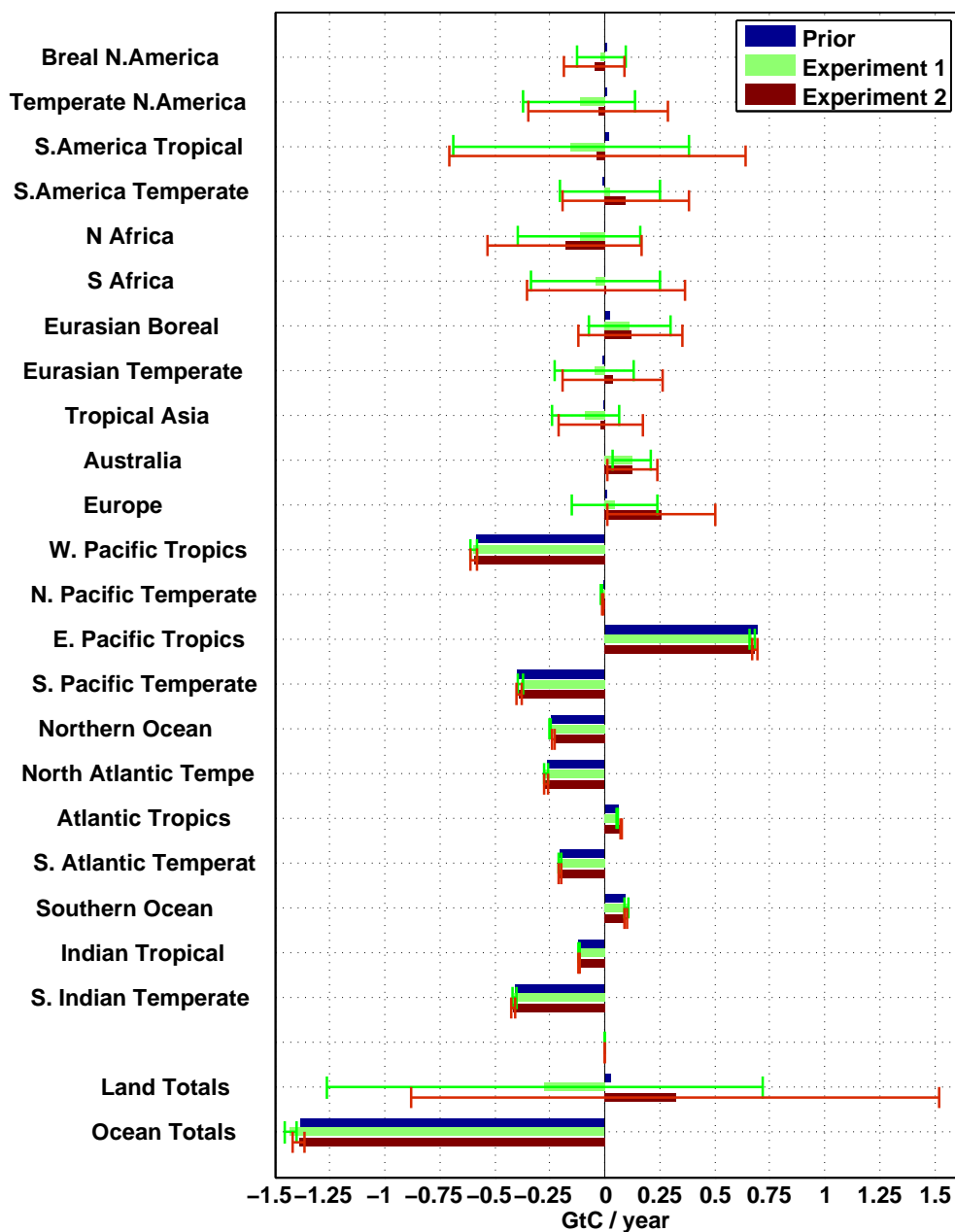


Figure 5.16: Bar plot of total annual fluxes from prior (dark blue), analysis from Experiment 1 (green), and analysis from Experiment 2 (red) with corresponding posterior uncertainty estimates (standard deviation), aggregated up to TransCom Regions (see Figure 5.15), in  $\text{GtC year}^{-1}$ . The top 11 regions correspond to land, the next 11 regions correspond to ocean, and the bottom two regions correspond to total land and ocean.

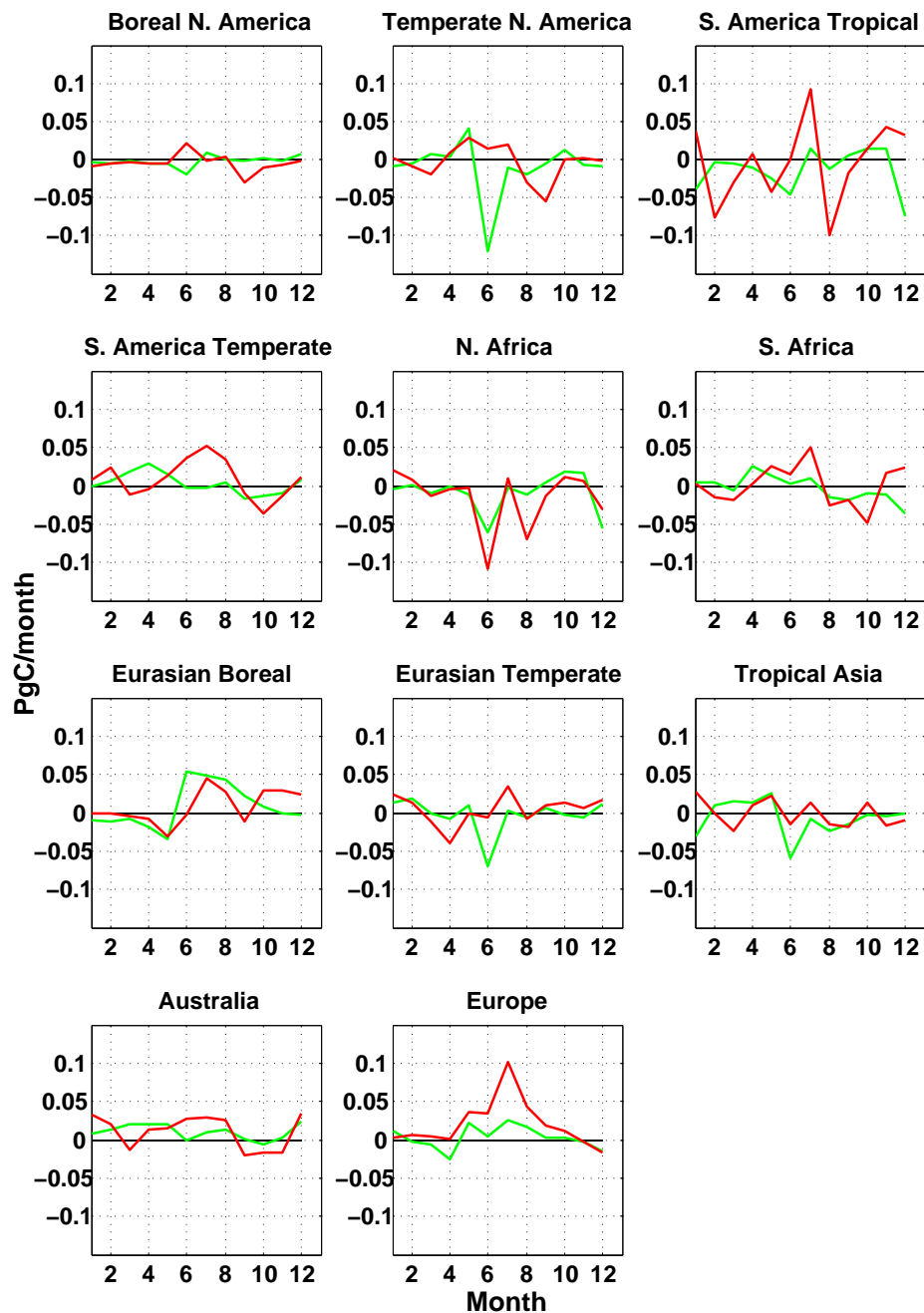


Figure 5.17: Seasonal flux errors for TransCom land regions 1-11 for Experiment 1 (green) and Experiment 2 (red), plotted in  $\text{GtC month}^{-1}$ .

### 5.3 Bias Recovery and Perfect Transport

The previous experiments tested the inversion framework for systematic errors in flux retrieval due to (1) internal problems with the optimization that might arise, for example,

due to the ensemble approximation (Experiment 1) and (2) reductions of observation constraint due to cloud contamination (Experiment 2). The results showed that the inversion is fairly robust to random measurement errors, but not necessarily data reductions that introduce fair weather bias into the measurements. These experiments are extremely simplified in the sense that the true fluxes are unbiased, and therefore the same as the priors.

The next step is to test whether the inversion can recover hypothetical truths hidden within cloud-screened satellite data and therefore learn from satellite data that contains signals due to fluxes that differ from the priors. To simulate a real world situation that the inversion might encounter, true fluxes are prescribed from realistic carbon processes that cause a net annual sink of  $\sim 3$  GtC over land. The catch is that, while such a prescribed global land sink is realistic, the underlying processes are slowly varying and produce fluxes that are small relative to background fluxes. For example, terrestrial NEE varies strongly in time and space due in large part to variations of solar radiation, which varies in time at diurnal, synoptic, and seasonal scales and in space as a function of latitude and season. Prescribed sinks are assumed to be small but persistent such that, over time, the inversion sees through background noise and recovers low signal-to-noise information.

As discussed in Chapter 2.6.1, hypothetical truths are generated by assuming that background fluxes are systematically biased by processes not modeled well in SiB. Processes that bias model estimates of *GPP* include underestimation of available nitrogen, forest management, agricultural land use, and CO<sub>2</sub> fertilization. An example of a process that biases *RESP* is forest regrowth. These persistent processes are represented by the multiplicative correction factors  $\beta_{GPP}$  and  $\beta_{RESP}$ , which were set to zero in control experiments. To generate hypothetical truth from these processes,  $\beta_{GPP}$  and  $\beta_{RESP}$  are prescribed according to the maps in Figure 5.18. *GPP* (Figure 5.18A), for example, is enhanced ( $\beta_{GPP} > 0$ ) in this hypothetical world in (1) northern mid-latitude terrestrial forests in the northeastern temperate N. America and Europe and (2) tropical regions of S. America, Africa, and Asia. Processes that control long-term and persistent enhancement of *GPP* include nitrogen de-

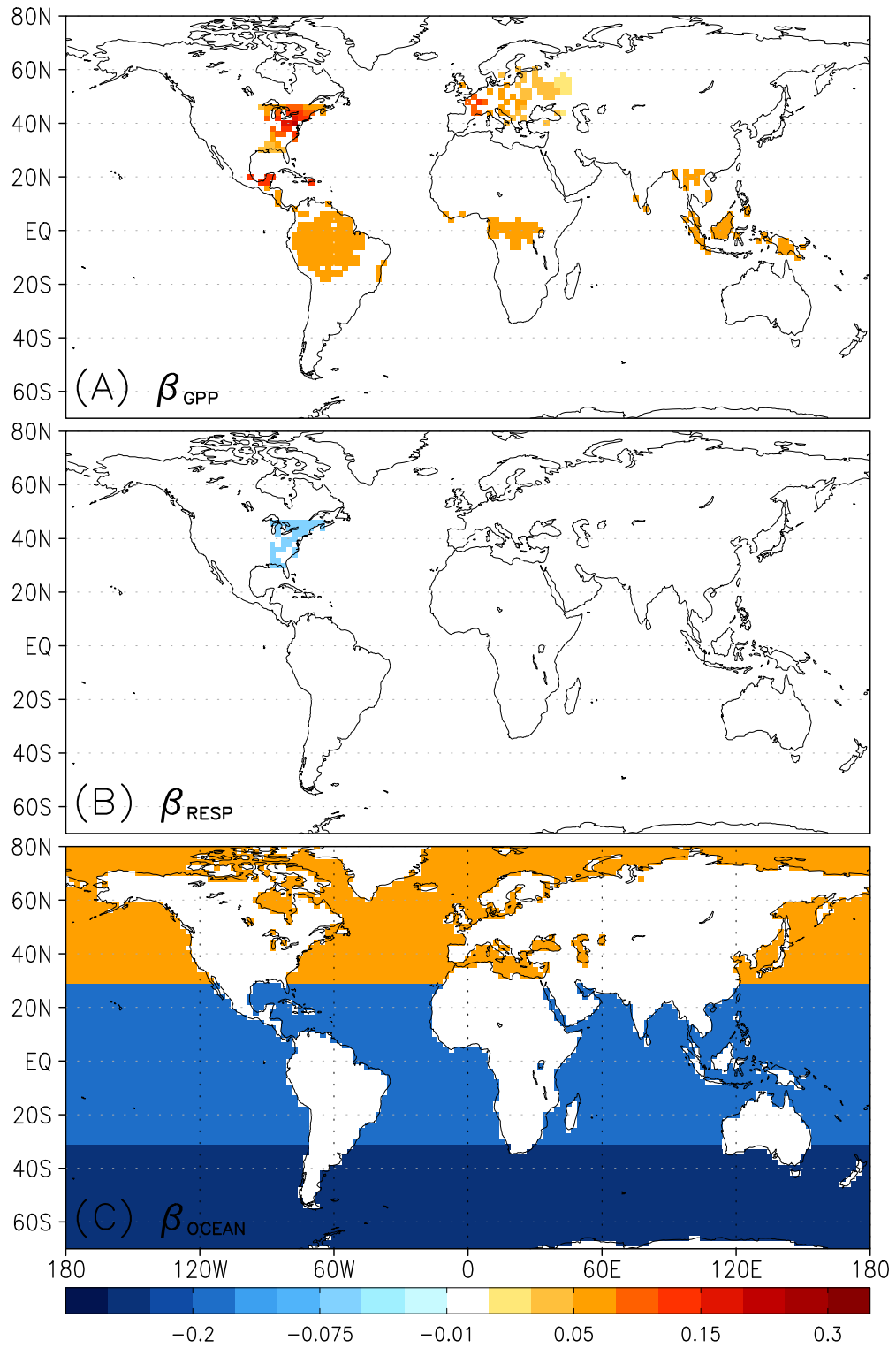


Figure 5.18: Multiplicative correction factors  $\beta_{GPP}$  (a),  $\beta_{RESP}$  (b), and  $\beta_{OCEAN}$  (c).

position in the eastern N. America, agricultural land use in Europe, and CO<sub>2</sub> fertilization in the tropics. To simulate regrowing forests in northeastern Temperate N. America that result from forests still recovering from agricultural abandonment and forests harvested before 1980, respiration (Figure 5.18) is reduced by 5% ( $\beta_{RESP} = -0.05$ ), since by definition a regrowing forest is one in which *RESP* lags *GPP*.

Biases in air-sea exchange are also included in this experiment (Figure 5.18C) as a function of latitude, with enhancement of flux in northern middle and high latitudes, reduction of flux at tropical latitudes, and strong reduction of flux at southern middle and high latitudes. Relative to annual mean air-sea exchange shown in Figure 5.7, these biases enhance uptake in the north Pacific and Atlantic Oceans (deepening mixed layers), reduce outgassing in the tropics and especially in the eastern Pacific (enhanced upwelling during La Nina), and reduce outgassing in the Southern Ocean (enhanced circulation during climate change relaxes back to conditions typical of a pre-industrial climate).

Like in the control experiments, the bias recovery experiments assume that  $\beta$ 's are initially unknown in the priors by setting  $\beta_{GPP} = \beta_{RESP} = \beta_{OCEAN} = 0$ . Experiment 3 assumes that true  $\beta$ 's are constant throughout the year. Experiment 4 assumes  $\beta$ 's vary in space and time and are perturbed with a random 10% noise. This added complexity is necessary considering that  $\beta$ 's, which represent slowly varying processes in the real world, may be persistent but are likely not constant in time or smooth in space.

### 5.3.1 *Experiment 3: Perfect Transport, Constant and Unperturbed Biases, and Cloud Screening*

Maps of total annual true and recovered fluxes are shown at pixel scale in Figure 5.19. By design, the truth (Figure 5.19A) contains net sinks in the annual mean in temperate N. America, Europe, and tropical S. America, Africa, and Asia. ***Recovered sinks are remarkably similar in both pattern and magnitude***, most notably in the eastern United States and throughout the tropics (Figure 5.19B). The seasonal cycle of terrestrial NEE is



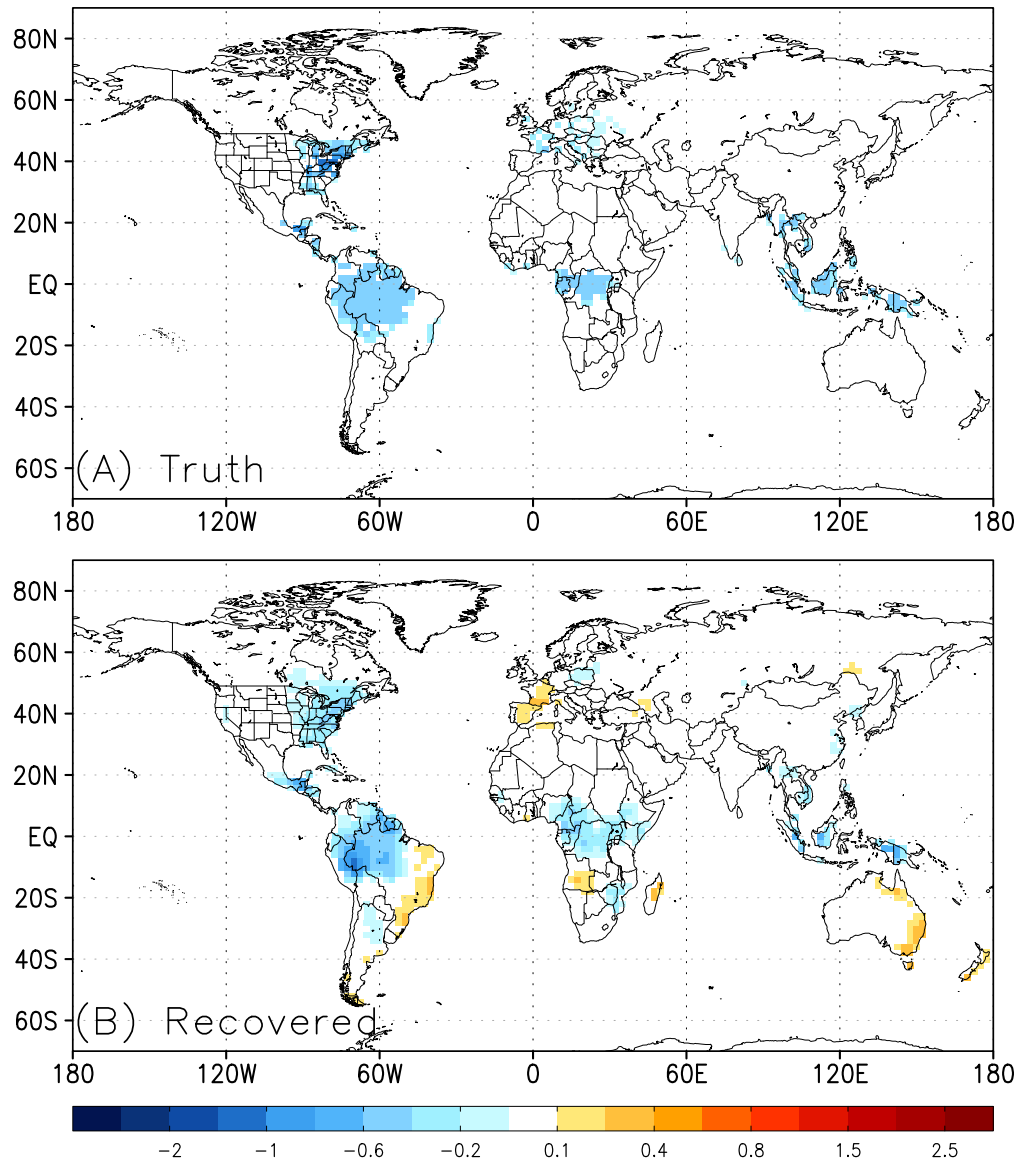


Figure 5.19: True (a) and recovered (b) annual mean terrestrial NEE in  $\mu\text{moles m}^2 \text{s}^{-1}$  for Experiment 3.

also recovered well, as demonstrated by a plot of seasonal prior, true, and recovered NEE at a location from the middle of the forest in northeastern United States (see Figure 5.20). This shows that the sink at this location is partially recovered at monthly time scales. The recovered flux has a tendency towards the true flux but is pulled back by the prior weighting.

According to Figure 5.19, there are several differences between the true and recovered fluxes, most noticeably in western Europe, the east coast of S. America, S. Africa, and

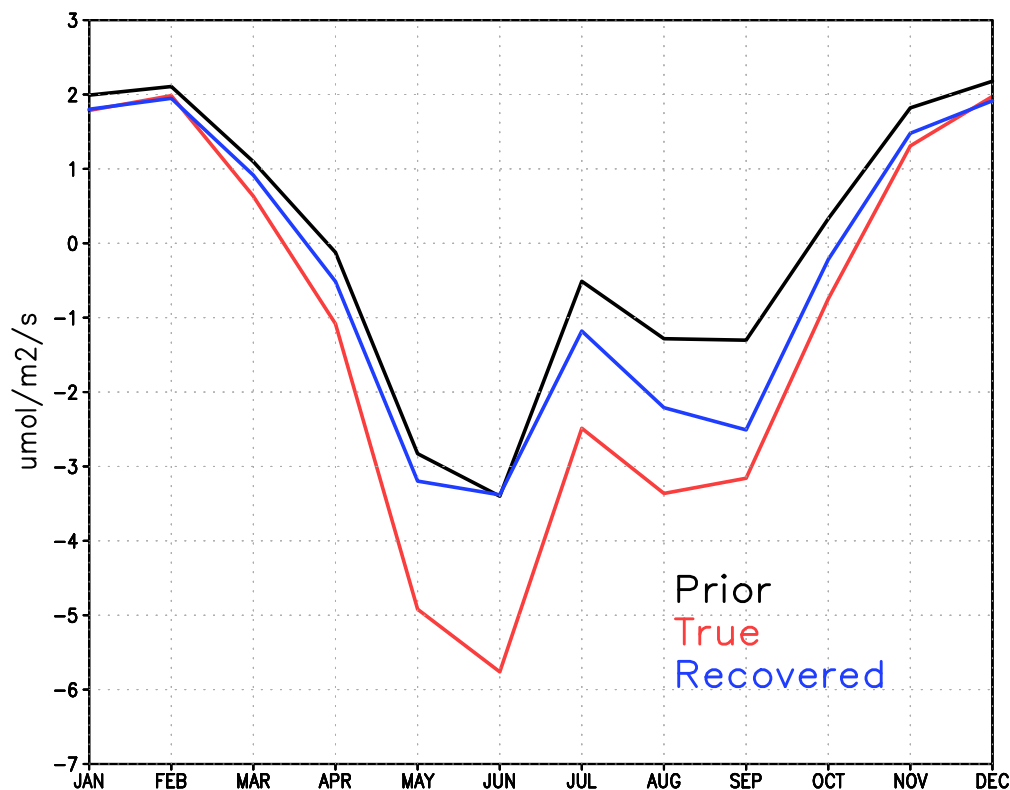


Figure 5.20: Seasonal cycle of prior (black line), true (red line), and recovered (blue line) terrestrial NEE in  $\mu\text{moles m}^2 \text{s}^{-1}$  at a forested location in Temperate N. America with coordinates  $[80^\circ\text{W}, 40^\circ\text{N}]$ .

eastern Australia. These errors are especially apparent in a map of the difference between recovered and true fluxes shown in Figure 5.21. Flux errors are largest where artificial sinks are prescribed.

Total annual recovered sinks, along with prior and true fluxes as reference, are aggregated up to TransCom regions and shown as bar plots with posterior uncertainty estimates in Figure 5.22. Agreement at the continental scale of TransCom regions is remarkably good. The most impressive results are in recovery of terrestrial NEE, especially in Temperate N. America, S. America Tropical, S. Africa, and Tropical Asia. The recovered flux in Temperate N. America is nearly indistinguishable from the true flux. Comparison of bias recovery in Experiment 3 with unbiased recovery in Experiments 1 and 2 (see Figure 5.16) shows that recovered fluxes for the most part stay true to the satellite data. It is very encouraging,

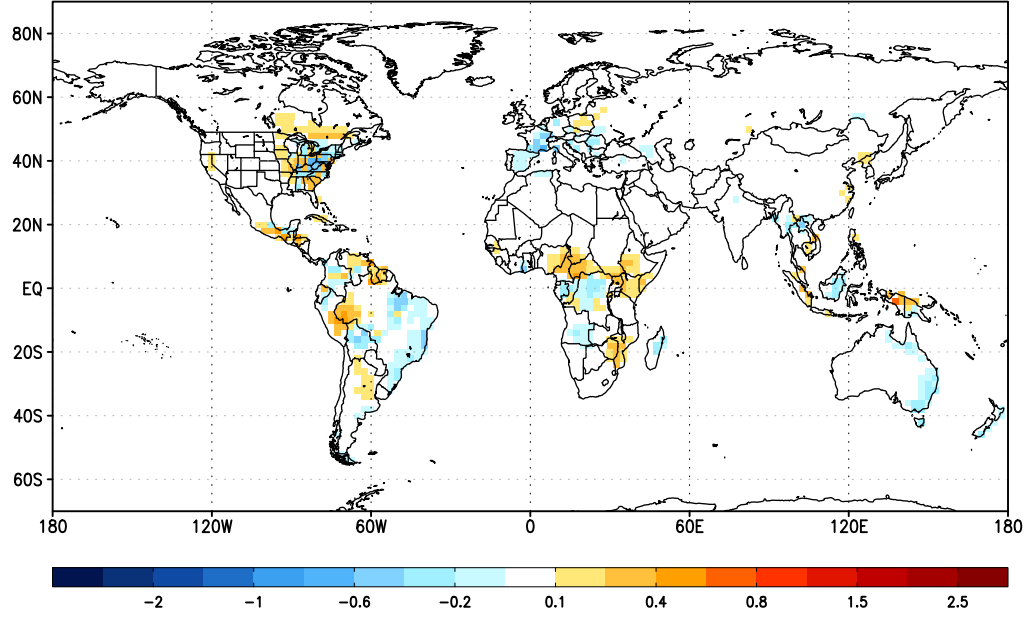


Figure 5.21: Difference between true and recovered annual mean flux, in  $\mu\text{moles } m^2 \text{ s}^{-1}$ , for Experiment 3.

for example, that recovered fluxes are closer to the truth than the prior. Furthermore, with the small exception of Tropical Asia, recovered fluxes are always within  $1 \sigma$  of the truth. In regions where substantial sinks are prescribed such as Temperate N. America, S. America Tropical, and N. Africa, flux estimates are certain to within  $1 \sigma$  of having the correct sign (i.e., source vs sink). These posterior uncertainty estimates need to be interpreted carefully, however, as they are strongly dependent on prior uncertainty, which at 20% of component fluxes is somewhat generate.

The seasonality of recovered fluxes in TransCom regions is also impressive (Figure 5.23). As in Figure 5.21, prior fluxes in Temperate N. America are eventually pulled towards the truth in late summer and fall. The large sink in S. America Tropical is  $\sim 100\%$  recovered from the onset of data assimilation. The same is true for Tropical Asia.

Some noteworthy errors are apparent in recovery of TransCom scale sinks. Like in Experiments 1 and 2, recovered fluxes have a strong positive bias relative to the truth in Australia and Europe. The recovered flux over Europe suggests a net source when actually

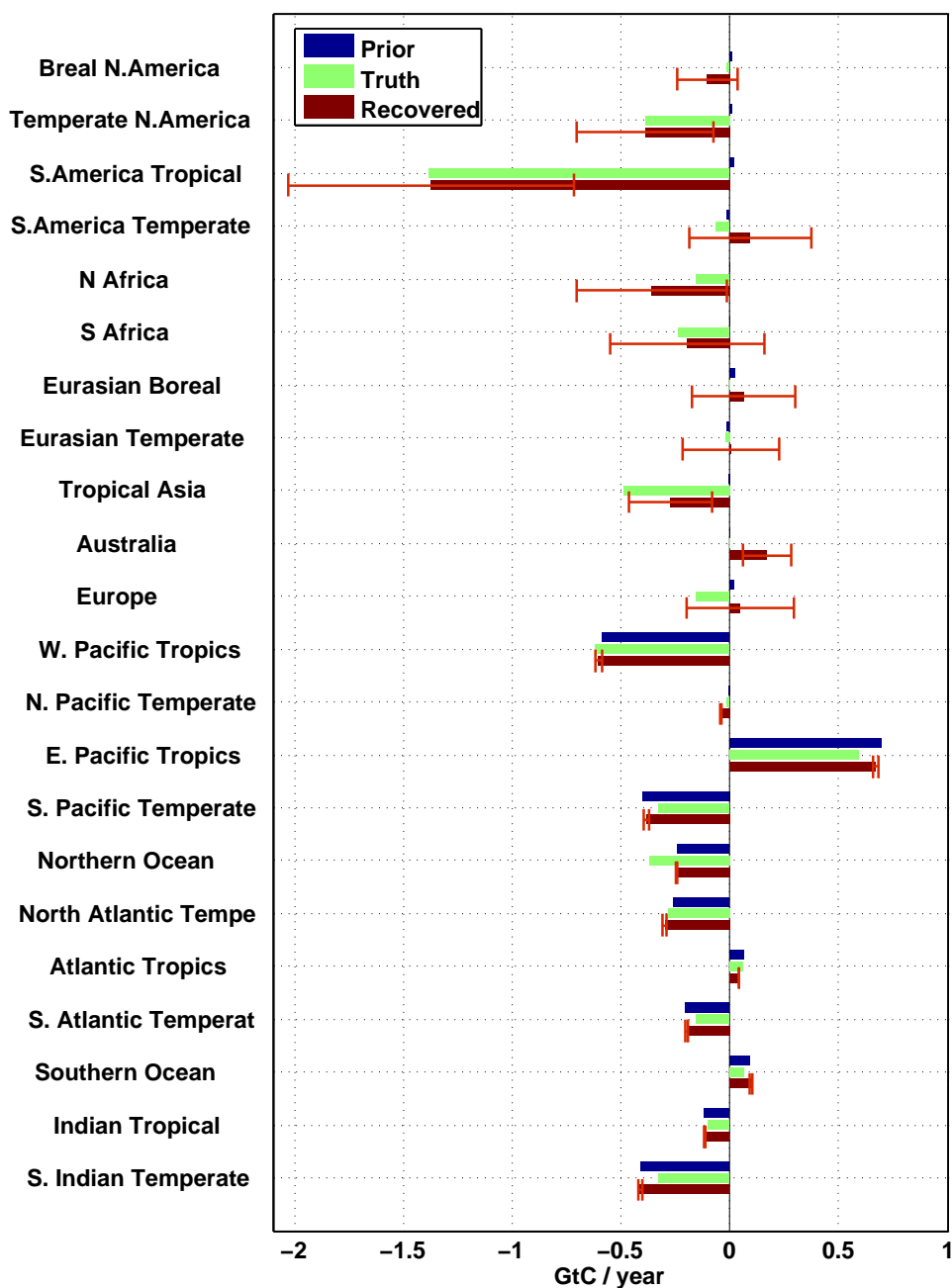


Figure 5.22: Bar plot of annual prior flux (dark blue), true flux (green), and recovered flux (red), aggregated up to TransCom regions like Figure 5.16, and plotted in  $\text{GtC year}^{-1}$  for Experiment 3. Posterior uncertainty is plotted in red.

the true flux is a sink. The seasonal distribution of the European sink recovery shows that Europe stays close to the prior for almost the entire year, and is at least  $1 \sigma$  away from the truth for much of the year. The recovered solution should always fall somewhere

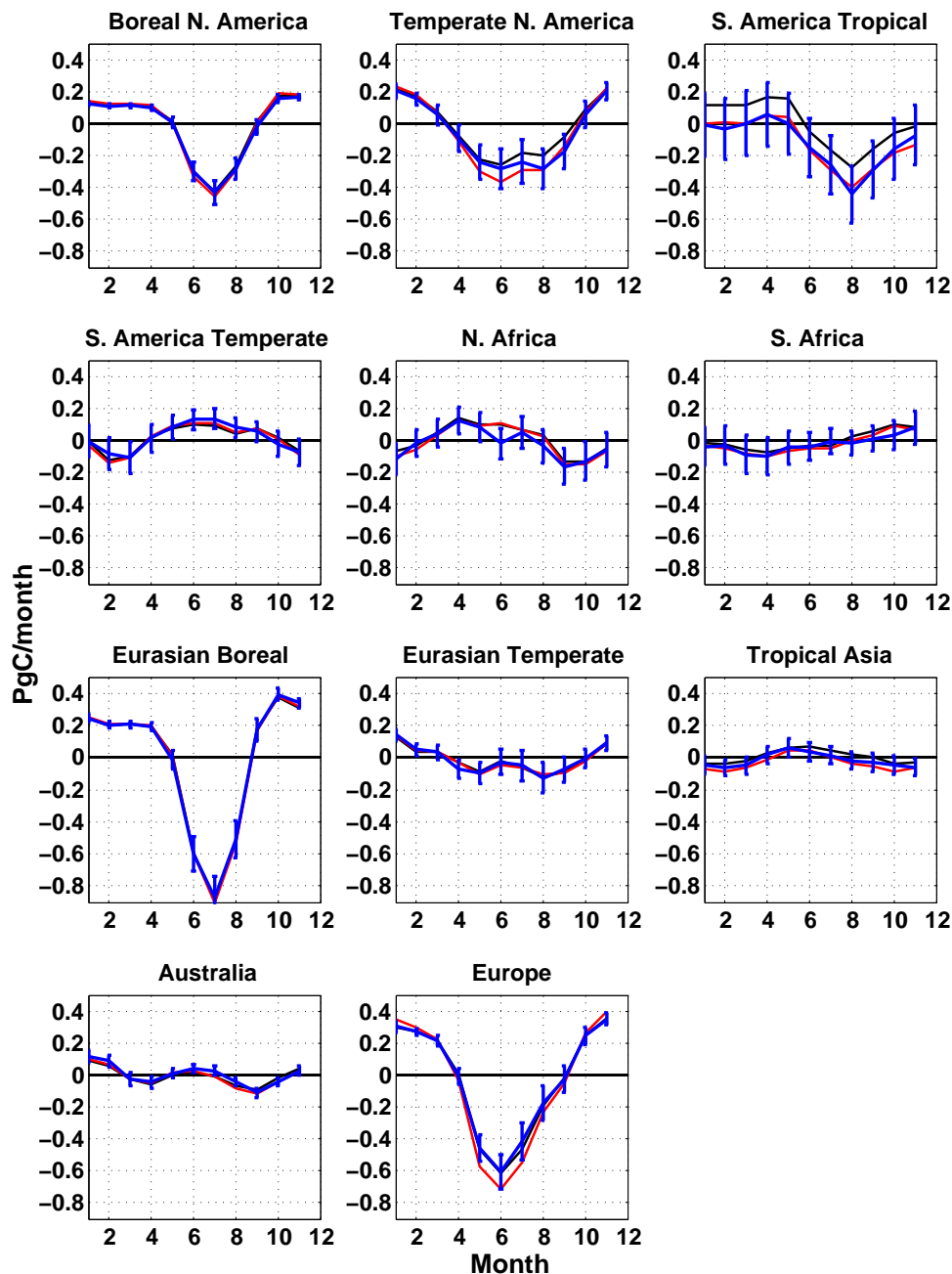


Figure 5.23: Seasonal cycle of prior (black), true (red), and recovered (blue) TransCom fluxes for Experiment 3 (biased truth) for land regions 1-11, plotted in  $\text{GtC month}^{-1}$ . Posterior (red) uncertainties are plotted as error bars.

between the prior and true flux, but this is not always the case (e.g., N. Africa in the annual mean and S. Africa at the end of the year) due to errors discussed in Experiments 1 and 2. The total recovered land sink is within 20% of the truth (not shown), which isn't bad

but leaves room for improvement. Recovered ocean fluxes aren't significantly different from priors. This is expected because it is assumed that ocean models are able to better simulate slowly varying processes than process-based ecosystem models of terrestrial NEE. That being said, the inversion was given the opportunity to recover biases in air-sea exchange. Overall, recovered fluxes fall somewhere between priors and truth.

### 5.3.2 *Experiment 4: Perfect Transport, Seasonal and Noisy Biases, and Cloud Screening*

Seasonality is added to  $\beta$ 's in Figure 5.18 through a sinusoidal function with peak amplitude during boreal summer such that fluxes in northern latitudes are most strongly biased during the active part of the growing season. To represent this seasonal dependence, constant  $\beta$ 's are divided into 12 monthly values and then multiplied by the function  $1 + 0.5\cos(x + \pi)$ . This has the effect of changing regional sinks by a few percent, but increasing the total land sink only very slightly from 2.253 to 2.305 PgC month<sup>-1</sup>.  $\beta$ 's are then multiplied by white gaussian noise with  $\mu = 0$  and  $\sigma = 0.1$  using the *rnorm(.)* function in *FORTRAN*. Maps of true and recovered fluxes are shown at pixel scale in Figure 5.24. Recovered sinks are again similar in both pattern and magnitude, and overall not much different from Experiment 3. The recovered fluxes aren't as noisy as the true fluxes due to (1) covariance smoothing in the first run and (2) smoothing out of noisy flux patterns by atmospheric mixing. It is likely, however, that small scale variations in the  $\beta$ 's could be recovered more accurately with surface measurements that are close, but not too far downstream, to the sources and sinks.

As was done for the case of constant  $\beta$ 's, total annual prior, true, and recovered fluxes are aggregated up to TransCom regions in Figure 5.25. Agreement at the continental scale of TransCom regions continues to be remarkably accurate despite biases that change with time and aren't as smoothly distributed as in Experiment 3. These  $\beta$  recovery runs have shown that the inversion is can detect small and slowly varying signals in the satellite data

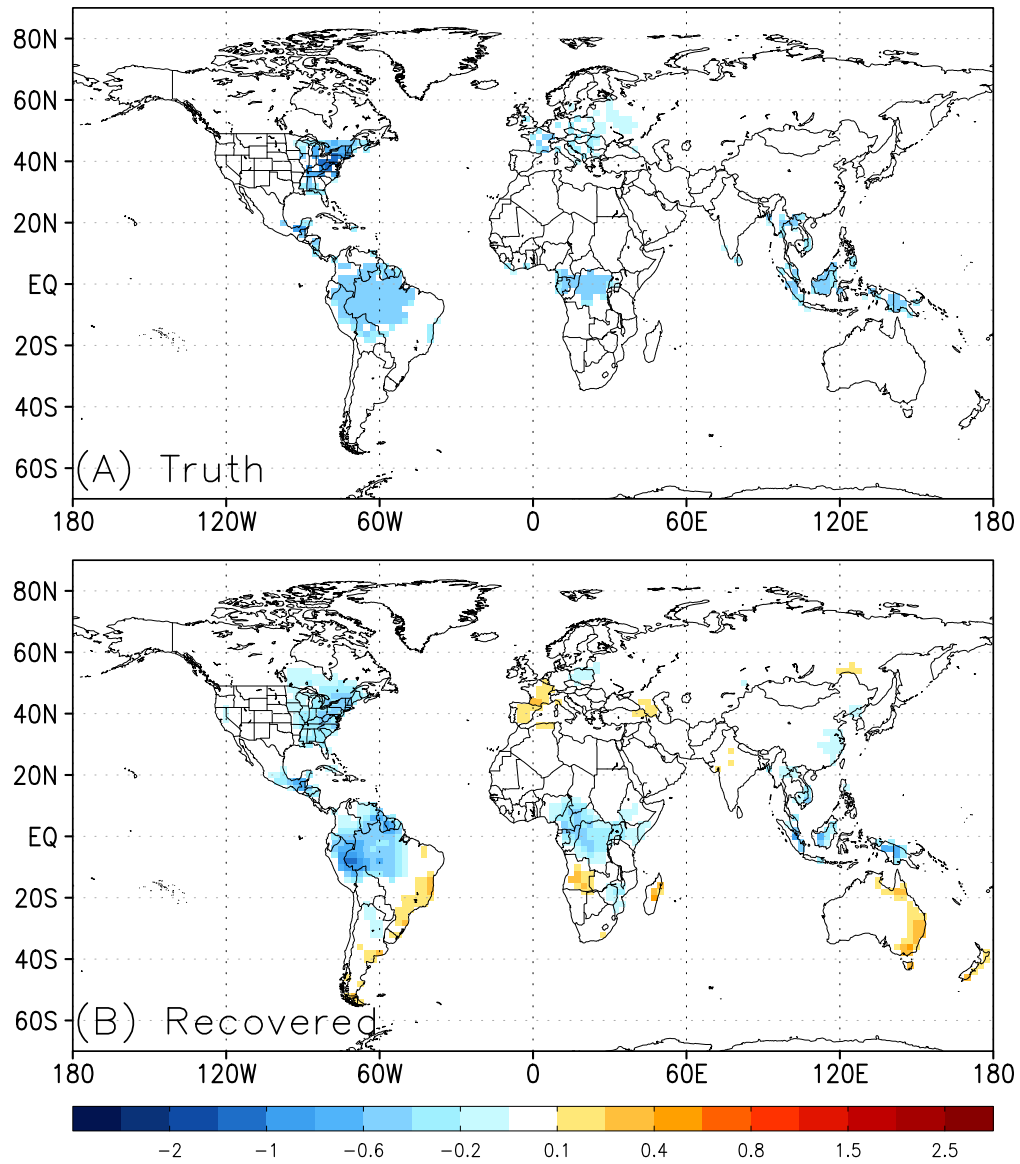


Figure 5.24: True (a) and recovered (b) annual mean terrestrial NEE in  $\mu\text{moles m}^2 \text{s}^{-1}$  for Experiment 4.

despite random noise in the measurements and true fluxes.

## 5.4 Biased Transport

The control runs in Section 5.2 illustrate the degree of accuracy of the MLEF inversion strategy given perfect transport and unbiased truth. Flux errors are generally small (although biased in northern summer latitudes) despite 3 ppm random measurement noise.

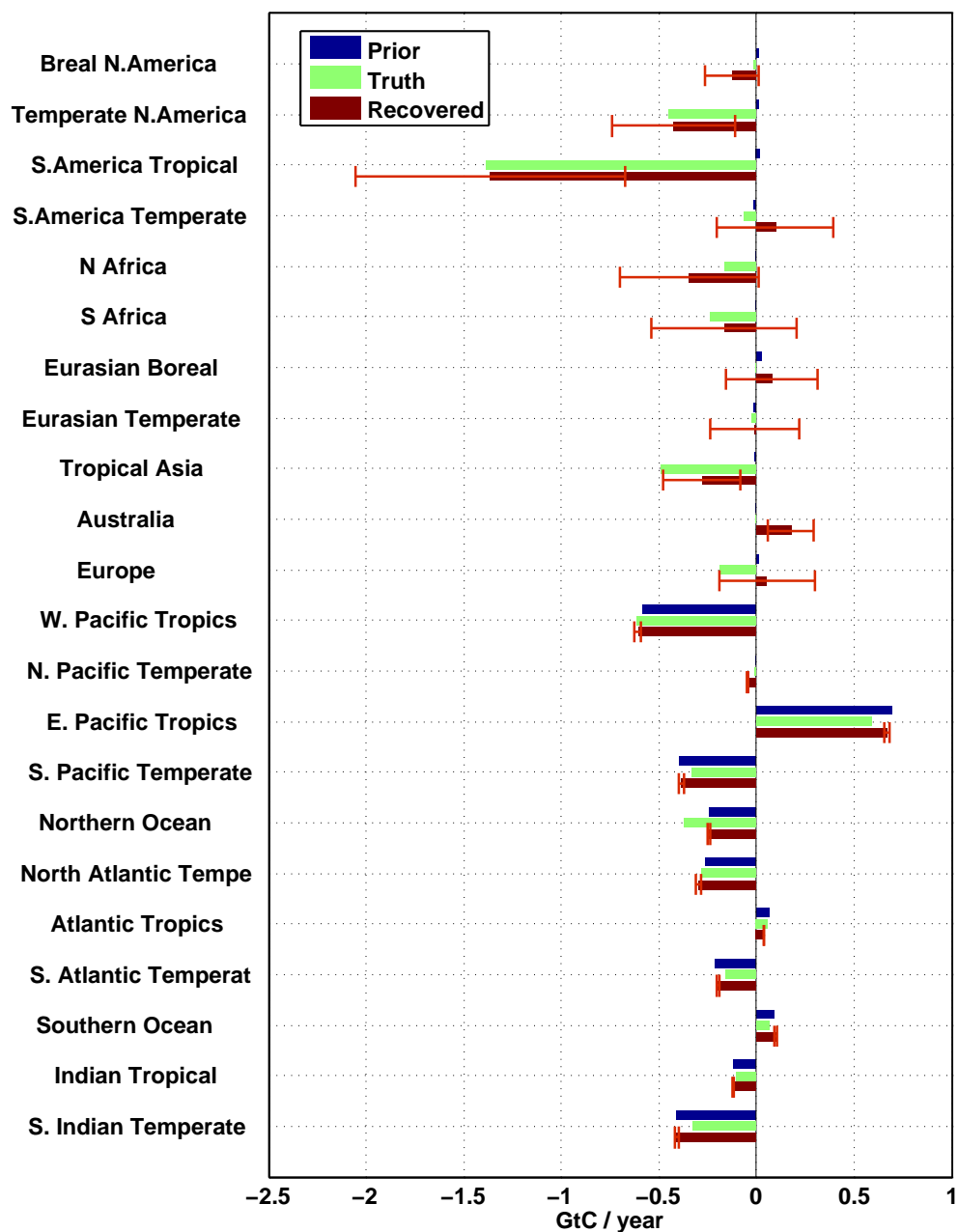


Figure 5.25: Bar plot of annual prior flux (dark blue), true flux (green), and recovered fluxes (red) for Experiment 4.

With a couple exceptions, true fluxes fall within  $1\sigma$  of the recovered flux. Realistic sources and sinks are then hidden within satellite data in Section 5.3 to test whether the inversion could recover small sinks superimposed on large background fluxes. These  $\beta$  recovery runs



illustrate the ability of MLEF to recover realistic sinks from satellite data given perfect transport and biased truth that differs from prior estimates by multiplicative correction terms. Together, the previous experiments have set the stage to test whether biased transport becomes aliased into optimized fluxes. To test for flux estimation errors due to biased transport, it is imperative that the inversion assumes perfect transport. One could, for example, account for biased transport in the observation covariance matrix. It is possible, however, that transport errors, whether purely random or biased, do not cause flux errors. In this case, there would be no need to account for biased transport. Perfect transport must therefore be assumed to determine whether flux errors different previous experiments arise.

The following experiments are carried out in the sense that synthetic satellite measurements are generated using one model and inverted using another model. Forward model runs with PCTM driven by GEOS5-0.67x0.5 are used to produce synthetic measurements, where realistic four-dimensional CO<sub>2</sub> output is sampled using the same GOSAT orbital ephemeris in the previous experiments. Due to the global mesoscale resolving capacity of GEOS5-0.67x0.5 meteorology, the synthetic measurements can thought of as a very good representation of actual satellite observations. These synthetic retrievals are then inverted using MLEF inversion of PCTM transport driven by the coarse grid of GEOS4-2.5x2. As discussed in Chapter 4, these models produce similar looking variations of CO<sub>2</sub> at synoptic timescales using the same basic eddy transport mechanisms, but horizontal gradients of CO<sub>2</sub> are much stronger in simulations driven by GEOS5-0.67x0.5. The goal here is to sample column CO<sub>2</sub> mixing ratios from the higher resolution simulations and then assess how these mixing ratios are interpreted by an inversion that then uses coarse resolution transport to translate synthetic measurements from another model back to surface flux distributions. If variations of column CO<sub>2</sub> are systematically different between transport models, surface fluxes backed out of the inversion will be different from surface fluxes going into synthetic measurements, and flux estimation errors will result.

Two experiments are run: Experiment 5, where unbiased truth (priors with  $\beta$ 's set to

zero) is run through GEOS5-0.67x0.5 to generate synthetic satellite data, and Experiment 6, where biased truth (priors with monthly varying non-zero  $\beta$ 's added to generate annual sinks) is run through GEOS5-0.67x0.5. Experiment 5 is therefore related to Experiment 2 in that the same true fluxes are used, but different in that true fluxes are run through a different transport model. Experiment 6 is related to Experiment 4 for the same reason.

#### 5.4.1 Experiment 5: Biased Transport, Unbiased Truth, and Cloud Screening

Annual mean errors in terrestrial NEE are shown at pixel scale in Figure 5.26. *Large,*

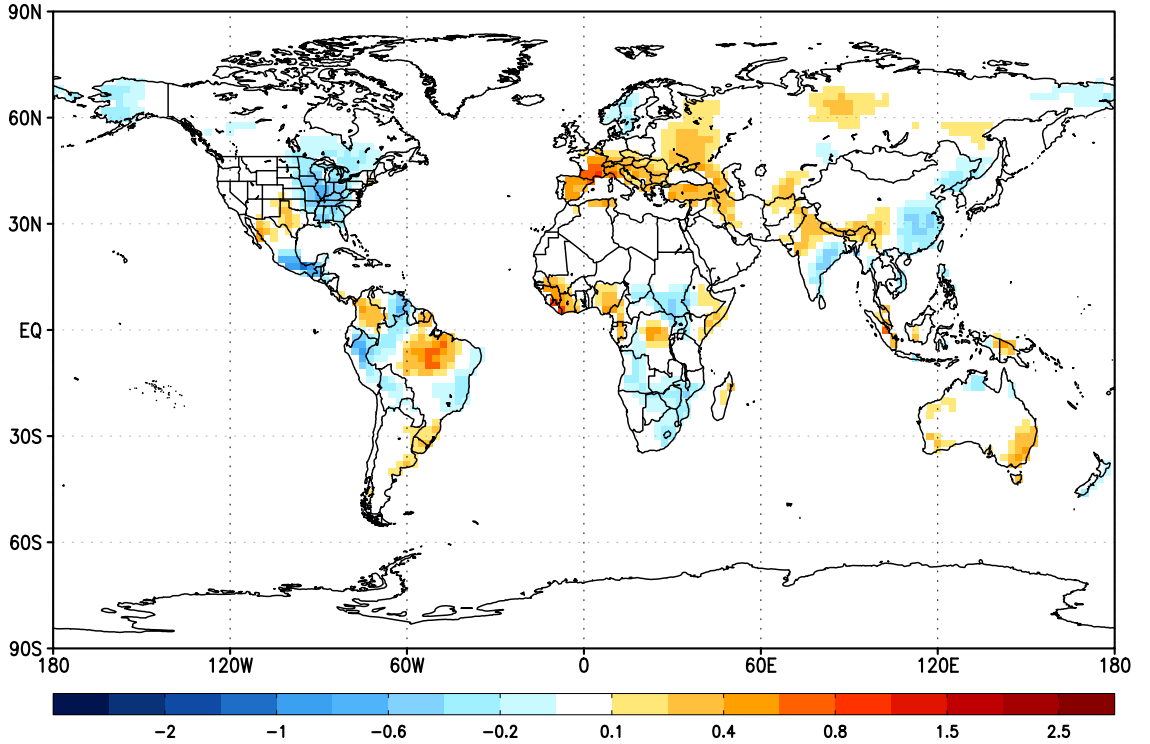


Figure 5.26: Annual mean flux errors over land for Experiment 5 (unbiased truth), plotted in  $\mu\text{moles m}^{-2} \text{s}^{-1}$  at pixel scale ( $2.5^\circ \times 2^\circ$ ).

*spatially coherent flux errors emerge on a global scale due to differences in transport between GEOS4 and GEOS5.* In particular, sinks appear over N. America and tropical Asia while net sources emerge over Europe and temperate and boreal Asia. S. America and S. Africa have strongly opposing sources and sinks. Annual errors in Experi-

ment 5 are significantly larger than those in Experiment 2.

It is even worse on seasonal timescales (Figure 5.27), especially during boreal summer. In most regions, similar patterns are seen throughout the year. For example, sinks

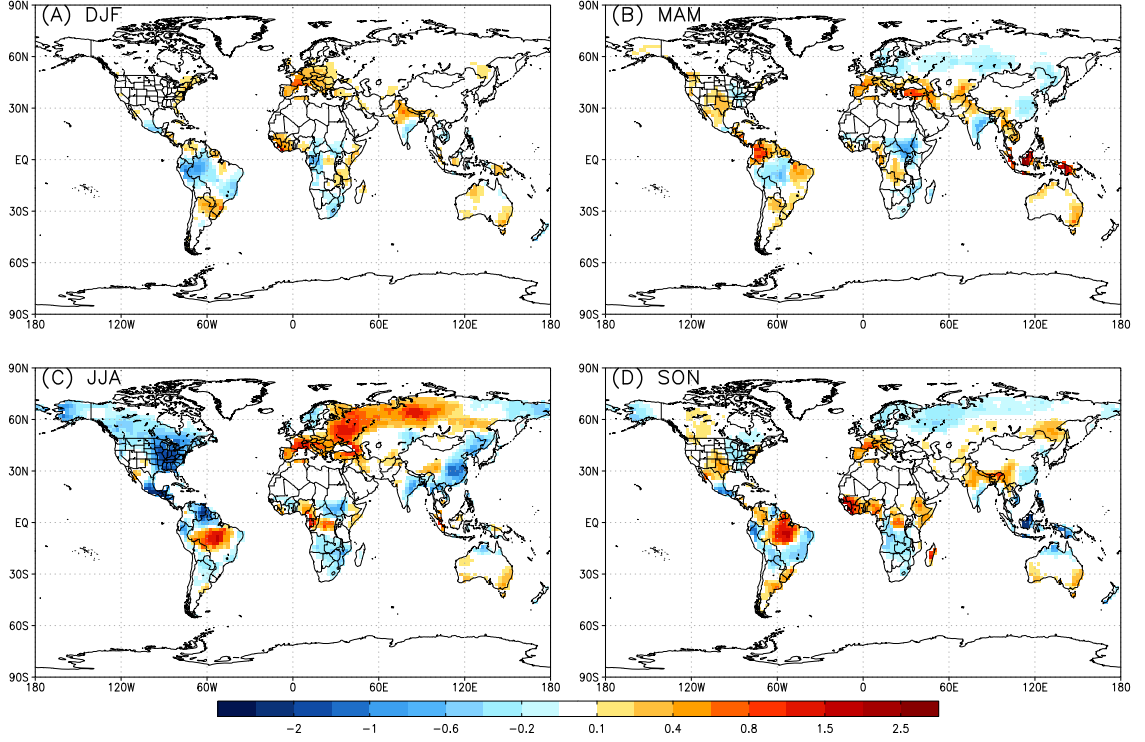


Figure 5.27: Seasonal flux errors over land for Experiment 5.

consistently arise in N. America, sources in Europe, and strongly opposing sources and sinks in S. America and Africa, regardless of season. *The summer sink in N. America in particular is in stark contrast to the summer source created when cloudy points were removed in Experiment 2.* There tends to be seasonal variations of sources and sinks in boreal Eurasia, with zero net flux during winter, sinks during spring, a strong source during summer, and back to a sink during fall.

These results are aggregated up to TransCom regions and compared to experiments with perfect transport in Figure 5.28. The key result here is that flux errors are small at continental scale, with truth contained within the  $1\sigma$  bounds of the posterior uncertainty, when transport is unbiased and all satellite points are included (Experiment 1). When

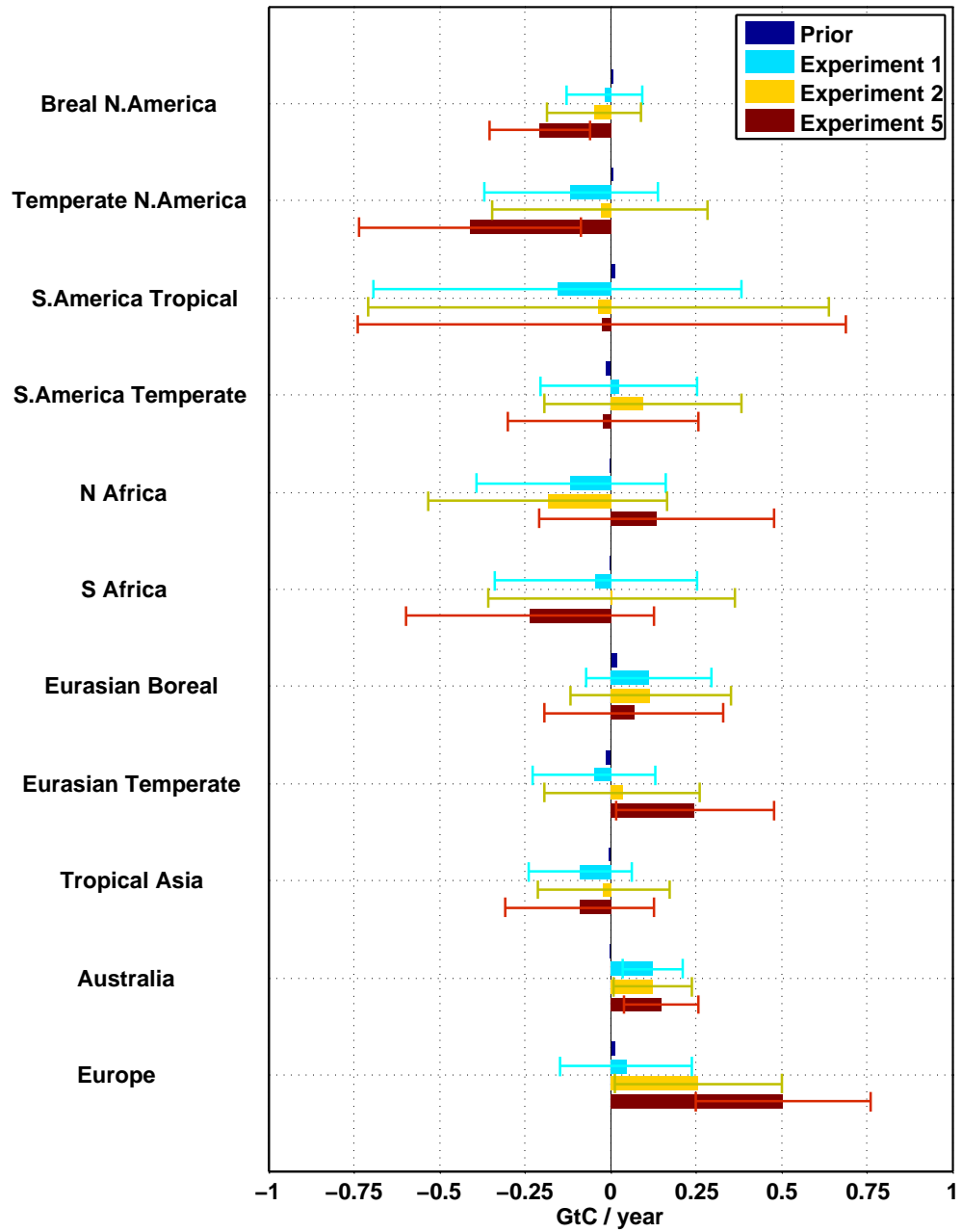


Figure 5.28: Bar plot of annual prior (dark blue), flux errors from Experiment 1 (light blue, perfect transport and no cloud screening), Experiment 2 (yellow, perfect transport and cloud screening), and Experiment 5 (red, biased transport and cloud screening), aggregated up to TransCom land regions and plotted in  $\text{GtC year}^{-1}$ .

cloud screening is applied similar errors occur, except flux errors grow large enough in Australia and Europe that the truth is no longer contained within  $1 \sigma$  (Experiment 2).

With the exception of Europe and Australia, the inversion incurs mostly small errors when given realistic satellite data and a perfect transport model. When transport is biased, however, much larger continental scale flux errors occur. In the cases of Boreal N. America, Temperate N. America, Eurasian Temperate, Australia, and Europe, flux errors are large and certain enough that the truth is no longer contained within  $1\sigma$  of the recovered flux. *The largest flux errors occur in Europe and Temperate N. America, amounting to  $0.51 \pm 0.256$  and  $-0.41 \pm 0.324$  GtC year<sup>-1</sup>, respectively.*

The seasonal cycles of true and recovered fluxes for TransCom land regions are shown in Figure 5.29. Recovered fluxes have the same seasonality as the truth, but are sometimes characterized by fluctuations about the truth. While flux errors accumulate in the annual mean as shown in Figure 5.28 over certain regions such as Europe, Eurasian Temperate, Temperate N. America, and Boreal N. America, other regions, such as S. America Tropical (and the tropics in general) have large seasonal errors that cancel in the annual mean. Finally, like in previous experiments without transport bias, posterior uncertainty estimates continue to decrease in time relative to the prior. *Introduction of transport bias into the MLEF framework causes the inversion to become more certain about the wrong answer.*

#### 5.4.2 Experiment 6: Biased Transport, Biased Truth, and Cloud Screening

Annual errors due to biased transport and biased truth (i.e., monthly varying sinks) are also significant (Figure 5.30), but not much different from errors due to biased transport and unbiased truth. Posterior uncertainty estimates are actually larger in a few regions when transport is biased. This is because the uncertainty is a function of the variance of  $\beta$  times component flux estimates. Flux estimates in most regions are slightly larger (or more wrong) in the case of biased transport, and so uncertainty also increases slightly.

The seasonal cycle of true and recovered fluxes for TransCom land regions for the case of biased truth is shown in Figure 5.31. Like in Experiment 5, recovered fluxes have

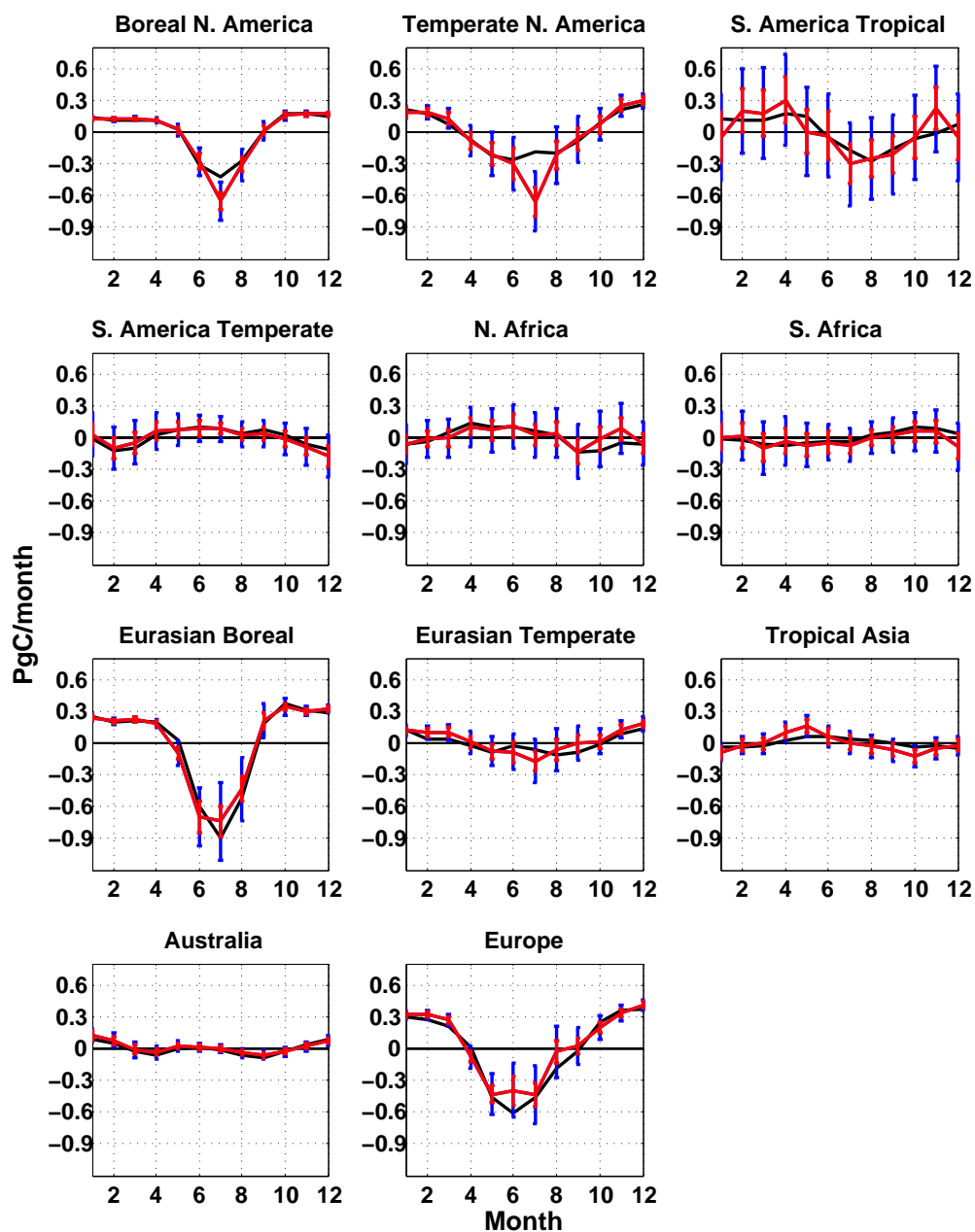


Figure 5.29: Seasonal cycle of true (black) and recovered (red) TransCom fluxes for Experiment 5 (unbiased truth) for land regions 1-11, plotted in  $\text{GtC month}^{-1}$ . Prior (blue) and posterior (red) uncertainties are plotted as error bars.

the same seasonality as the truth with some fluctuation superimposed by transport bias. Also like Experiment 5, posterior uncertainty decreases in time relative to the prior. In most cases, the truth is contained within  $1\sigma$  of the posterior uncertainty. Flux estimates in

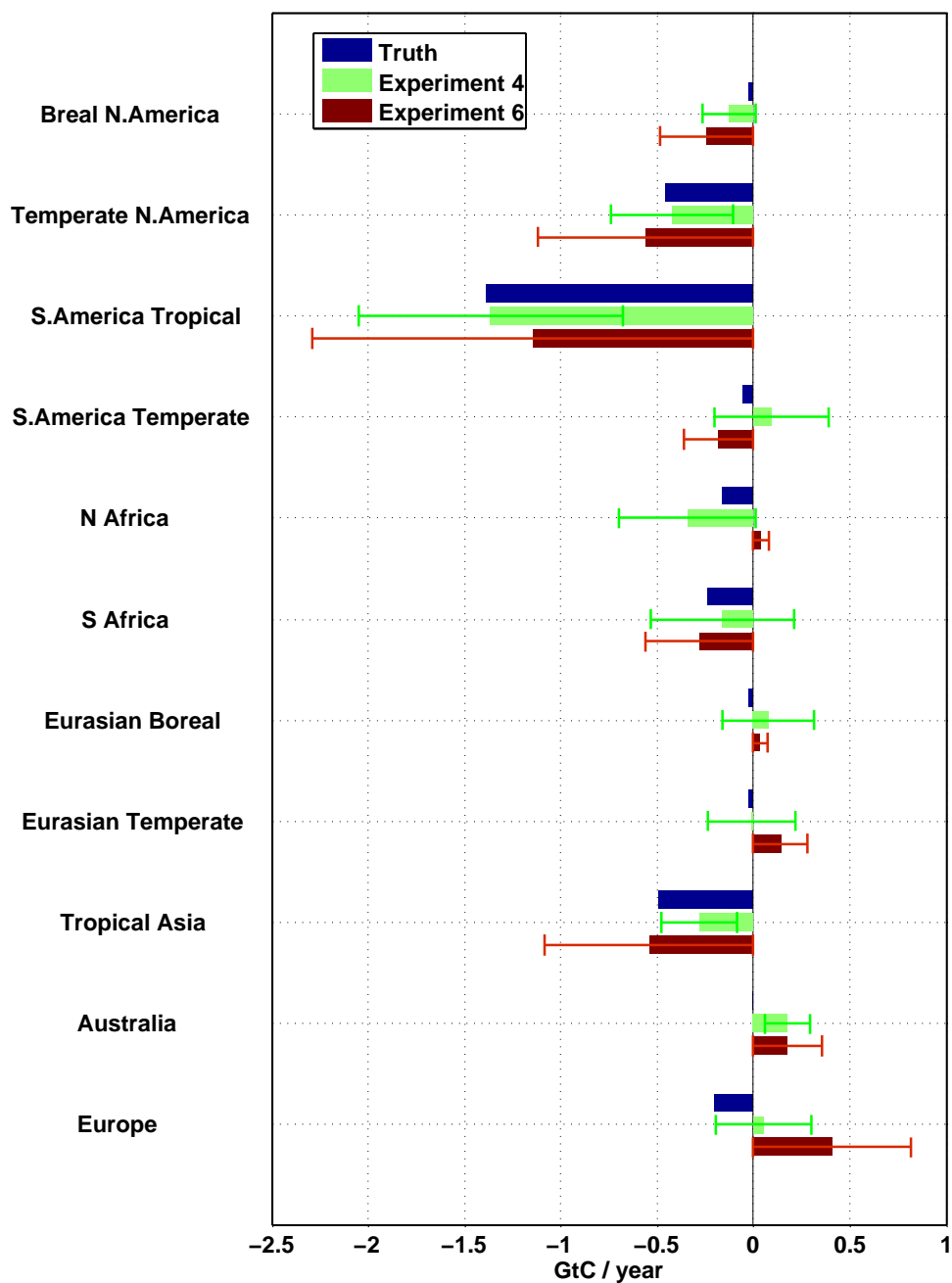


Figure 5.30: Bar plot of annual prior (dark blue), flux errors from Experiment 4 (green, perfect transport and monthly varying sinks), and Experiment 6 (red, biased transport and monthly varying sinks), aggregated up to TransCom land regions and plotted in  $\text{GtC year}^{-1}$ .

some regions such as Tropical Asia, Boreal N. America, and Eurasian Temperate are just outside the truth. Europe is a fairly extreme outlier in this experiment and is close to  $2 \sigma$

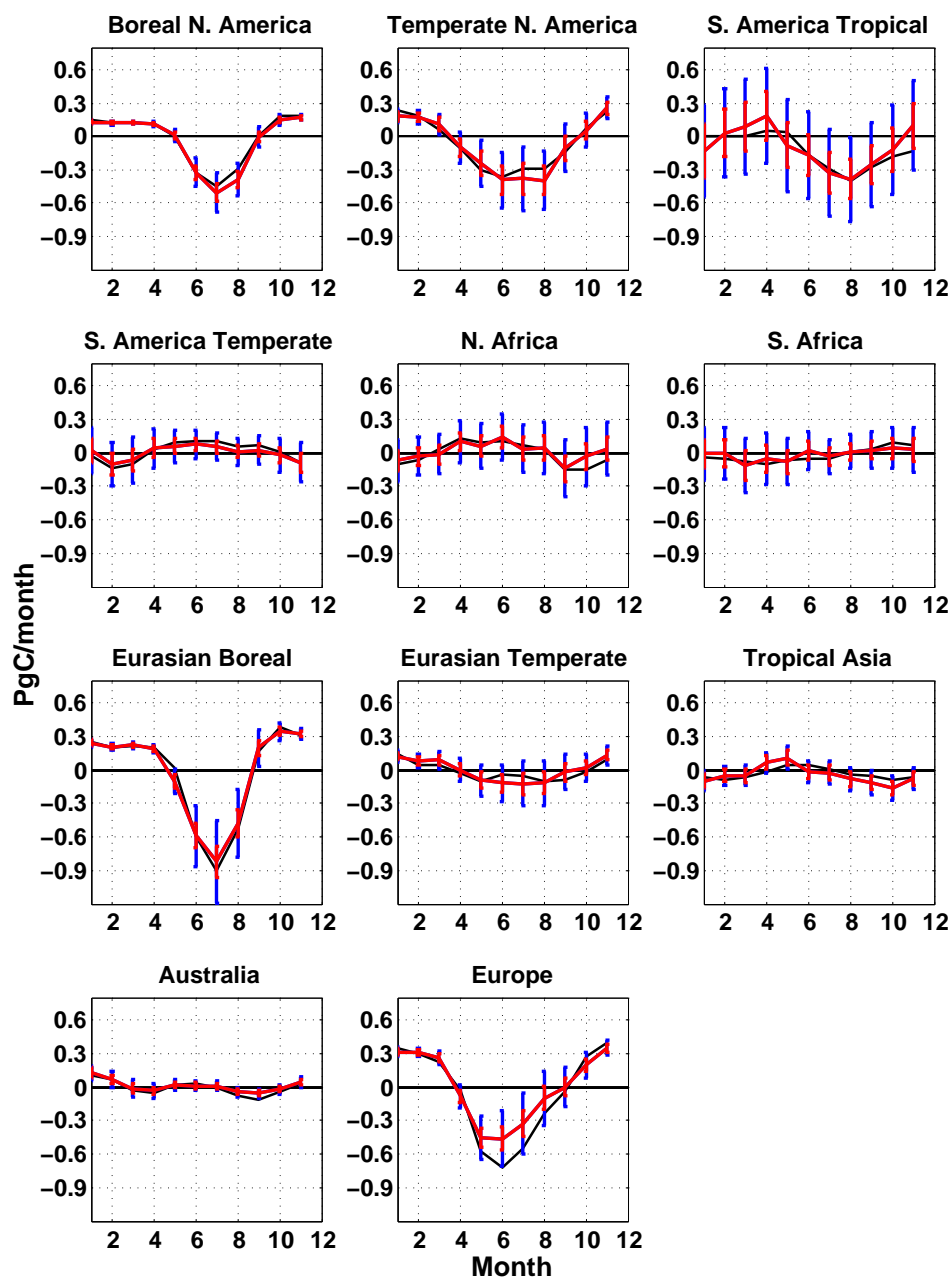


Figure 5.31: Seasonal cycle of true (black) and recovered (red) TransCom fluxes for Experiment 6 (biased truth) for land regions 1-11, plotted in  $\text{GtC month}^{-1}$ . Prior (blue) and posterior (red) uncertainties are plotted as error bars.

away from the truth during the summer.

Seasonal flux errors (recovered minus true) for TransCom land regions are plotted for Experiments 5 and 6 in Figure 5.32. With the exception that Experiment 5 thinks strong



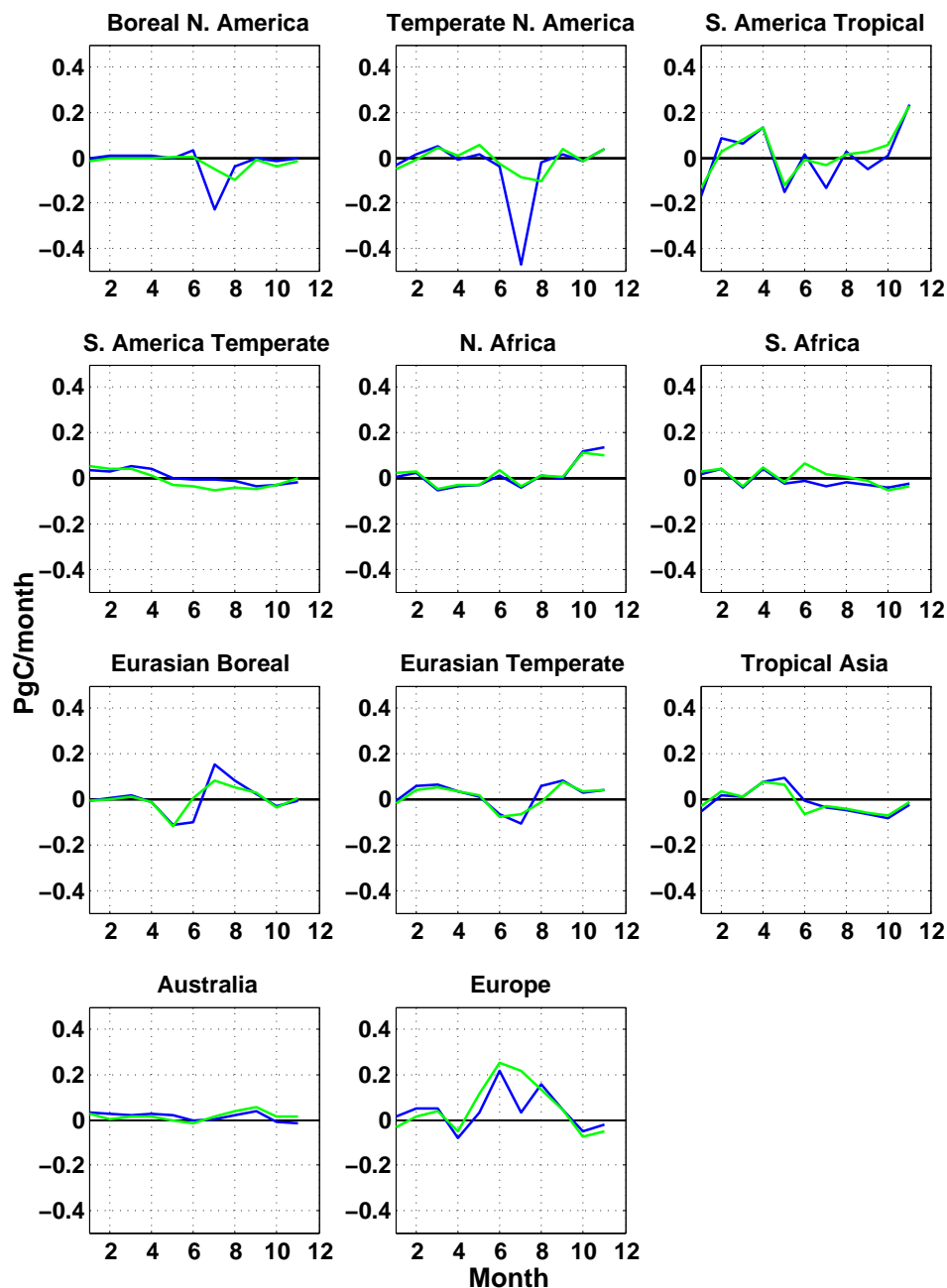


Figure 5.32: Seasonal flux errors for TransCom land regions 1-11 for Experiment 5 (blue) and Experiment 6 (green), plotted in  $\text{GtC month}^{-1}$ .

summer drawdown occurs in Boreal and Temperate N. America, the seasonality of flux errors is consistent between biased transport experiments. It seems the need for a strong sink in the biased transport model is satisfied by the N. American sink imposed in the

satellite data in Experiment 6.

## 5.5 Discussion and Conclusions from Part 3

This chapter consists of a comparison of end-to-end OSSEs with perfect and biased transport to test for flux estimation errors in the MLEF inversion framework. All flux estimates were optimized using MLEF inversion of synthetic satellite data with 3 ppm random measurement error. In the so-called perfect transport OSSEs, which consisted of experiments with and without hypothetical carbon sinks hidden within satellite data, the same transport model (PCTM driven by GEOS4-2.5x2) was used to generate and invert synthetic data. Because of the possibility that systematic errors in transport, such as those discussed in Chapter 4, are aliased into errors in inverse estimates of surface flux, a third set of so-called biased transport OSSEs were run in which synthetic satellite data was generated using a global mesoscale model (PCTM driven by GEOS5-0.67x0.5) and inverted using a coarse resolution global model (PCTM driven by GEOS4-2.5x2).

The perfect transport OSSEs are robust to random measurement errors and are able to recover small but reasonable looking sinks hidden within satellite data. The robustness of the inversion to random measurements errors results from the tendency for random errors to average out over time (in contrast to biased measurement errors, which are not discussed in this study). Small flux estimation errors are incurred in the perfect transport OSSEs, however, when column CO<sub>2</sub> data is excluded due to cloud contamination. There is a tendency in northern latitudes, most notably in Europe, for positive surface flux biases during the growing season, suggesting that synthetic satellite measurements must also have positive bias when data from moist synoptic storms is excluded. This is consistent with temporal sampling errors in atmospheric CO<sub>2</sub> mixing ratio calculated by Corbin et al. [2008] who show, along with Parazoo et al. [2008], that satellite data becomes enriched with CO<sub>2</sub> during the growing season primarily due to masking of large horizontal CO<sub>2</sub> gradients from satellites by clouds. These simple perfect transport OSSEs have demonstrated the tendency

for fair weather bias in satellite measurements in northern mid-latitudes to be aliased in flux inversions as spatially coherent positively biased surface fluxes of up to  $0.1 \text{ PgC month}^{-1}$  but at the very worst no more than  $0.257 \pm 0.245 \text{ PgC year}^{-1}$ .

Nevertheless, the inversion recovers realistic looking sinks hidden in satellite data despite (1) errors that result from excluding data in moist synoptic storms, (2) large and high frequency background noise, (3) random measurement errors of 3 ppm, and (4) monthly varying sources and sinks that contain random spatial noise. The most remarkable sink recovery occurs in the Amazon, where satellite data is all but eliminated by clouds. Further investigation is needed, but two possible explanations are (1) weak Coriolis and slow mixing timescales enhance atmospheric memory of nearby sources and sinks, thus increasing the odds that a satellite will sample a non-cloudy and information containing pixel, and (2) information over S. America is eventually propagated offshore where the chance of cloud contamination is reduced. Because the sinks are slowly varying and persistent throughout the year, and the MLEF framework is designed to propagate state vectors and error covariance matrices forward in time between assimilation cycles, optimized fluxes are able to learn from a years worth of satellite data and converge within reasonable distance of the true fluxes. Overall, this system produces excellent grid-scale and TransCom scale fluxes, assuming perfect transport.

Biased transport leads to substantial errors in flux estimation at all latitudes. The largest annual errors occur in northern temperate and boreal latitudes, including Boreal and Temperate N. America, Eurasian Temperate, and Europe. For example, the inversion estimates a  $0.5 \pm 0.256 \text{ GtC year}^{-1}$  source in Europe (nearly double the error due to cloud screening) and  $0.41 \pm 0.324 \text{ GtC year}^{-1}$  sink in Temperate N. America (almost no error due to cloud screening) even though no such sources or sinks are prescribed in the satellite data. Other regions that don't experience large annual errors actually experience huge opposing seasonal errors that cancel. Flux errors are equally large when sinks are added to the satellite data. The inversion is confident to  $1 \sigma$  that 5 out of the 11 land regions

contain sources or sinks of CO<sub>2</sub> that don't exist in the true fluxes. These flux error estimates are likely to be conservative given that the (1) two analysis products used in the biased transport OSSE are similar in architecture and (2) same chemistry transport model is used to generate and invert synthetic satellite data.

Flux errors caused by biased transport present a real challenge for inversion modelers and, as discussed in a similar study by Houweling et al. [2010], puts stringent requirements on the performance of transport models. Theoretically, transport errors could be accounted for in the observation covariance matrix. This is currently done in in-situ inversions in response to high uncertainty of nocturnal boundary layer dynamics by strongly reducing the precision of nocturnal continuous measurements [e.g., Lokupitiya et al., 2008]. This solution is less than ideal, however, as these measurements are essentially filtered from the inversion and therefore contribute almost no information to flux optimization. A far more desirable solution is to improve model representation of boundary layer mixing through some combination of development of model dynamics and data assimilation techniques.

This raises the question as to whether there are certain aspects of model development that could and/or should receive priority. With regard to this study, are there certain systematic differences between GEOS4-2.5x2 and GEOS5-0.67x0.5 that cause large pixel- and continental- scale flux errors? For example, Chapter 4 discussed two important differences between these models: (1) systematically weak vertical mixing in GEOS5-0.67x0.5 and (2) finer grid spacing and therefore better resolution of baroclinic waves. Although eddy transport differences are significant during winter, results from this chapter indicate that flux errors in northern mid-latitudes are very small during winter; thus, winter transport bias winter might not matter for flux inversion. However, large eddy transport errors also occur during summer, and we saw that annual mean flux errors are dominated by errors during summer when component fluxes and fine-scale vertical mixing by deep convection peak. It is therefore possible that weak vertical mixing in GEOS5-0.67x0.5 during summer contributes significantly to flux errors. This is certainly consistent with many previous studies [e.g.,

Stephens et al., 2007] who find that vertical mixing and surface flux estimation are strongly related. Future studies could do a much better job isolating the primary source of error by inverting satellite data with a model that has the capability of running a multiple grid spacing with vertical mixing held constant, and vice-versa.

## Chapter 6

### CONCLUSIONS AND FUTURE WORK

Mass transport along moist isentropic surfaces on baroclinic waves represents an important component of the atmospheric heat engine that operates between the equator and poles. This is also an important vehicle for tracer transport, and is correlated with ecosystem metabolism because large-scale baroclinicity and photosynthesis are both driven seasonally by variations in solar radiation. This research has pursued a dynamical framework for explaining atmospheric transport of CO<sub>2</sub> by synoptic weather systems at middle and high latitudes. A global model of atmospheric tracer transport, driven by meteorological analysis in combination with a detailed description of surface fluxes, is used to create time varying CO<sub>2</sub> distributions in the atmosphere. Simulated mass fluxes of CO<sub>2</sub> are then decomposed into a zonal monthly mean component and deviations from the monthly mean in space and time. Mass fluxes of CO<sub>2</sub> are described on moist isentropic surfaces in order to include transport along frontal systems in the eddy-terms rather than in the mean.

The resulting transport by CO<sub>2</sub> mass fluxes in moist conveyors and dry intrusions accounts for significant exchanges of CO<sub>2</sub> of up to 1 PgC month<sup>-1</sup> between middle and high northern latitudes. Through seasonal covariance with ecosystem metabolism, synoptic eddies modulate the seasonality of CO<sub>2</sub> mixing ratios in northern latitudes by strongly damping seasonality in the biologically active mid-latitudes by 50% of that implied by NEE while strongly amplifying seasonality in the Arctic. Meridional fluxes of CO<sub>2</sub> are of comparable magnitude as biological and anthropogenic surface exchange of CO<sub>2</sub> and thus require careful consideration in (inverse) modeling of the carbon cycle. An additional

complication arises in that transport along stormtracks is correlated with rising, moist, cloudy air along frontal clouds and moist conveyors, which systematically hide this CO<sub>2</sub> transport from satellites. Continuous in-situ records can supplement airborne and remotely observed measurements during inclement weather, but only at a few locations. This fair-weather bias in measurements puts stringent requirements on models of moist transport.

Tracer transport by frontal systems and moist processes occurs on scales that are typically poorly represented in global models and may therefore be a source of error for inverse estimates of CO<sub>2</sub> flux. Uncertainty in meridional transport is investigated by calculating model spread in eddy transport by a global model driven by four analysis products from the Goddard EOS Data Assimilation System for 2005 in combination with identical surface fluxes. Eddy transport is found to be highly variable between simulations, with significant seasonal biases of up to 0.2 PgC, representing up to 50% of fossil fuel emissions. The analysis products used are architecturally very similar and these bias estimates are therefore likely to be conservative. Differences in grid spacing and vertical mixing by moist convection and PBL turbulence are the primary culprits for the large model spread. Enhanced grid spacing is shown to dramatically improve the fine-scale structure of baroclinic waves, but strong sensitivity of CO<sub>2</sub> mixing ratio distributions, and therefore eddy transport, to uncertain vertical mixing parameters makes it difficult to gauge improved fidelity of CO<sub>2</sub> transport at global mesoscale resolution relative to the coarse grid analysis forcing.

To test for aliasing of transport bias into inverse flux estimates, synthetic satellite data is (1) generated by sampling forward simulations from the global mesoscale model using the GOSAT orbital ephemeris and (2) inverted using the coarsest grid model. The MLEF ensemble filtering method is used to optimize fluxes. When the same transport model is used to generate and invert satellite retrievals, the inversion framework is able to recover realistic looking sinks hidden within the data. When different transport models containing systematic biases due to vertical mixing and grid spacing are used, flux estimates are highly biased at pixel and continental scale at all latitudes. The most significant errors are the 0.5

$\pm 0.256 \text{ GtC year}^{-1}$  source in Europe and a  $0.41 \pm 0.324 \text{ GtC year}^{-1}$  sink in Temperate N. America. A portion of the total error in the biased transport experiment is attributed to temporal sampling biases that occur in northern temperate and boreal latitudes during the growing season as a result of exclusion of column  $\text{CO}_2$  data from moist synoptic storm systems. There are also large seasonal errors in tropical regions. Errors due to transport bias are likely to be conservative, assuming flux errors scale with transport bias.

Key results are highlighted below...

- The prominence of tracer transport by synoptic storms is illustrated more clearly when written on moist isentropic surfaces
- Synoptic storms strongly damp the seasonality of  $\text{CO}_2$  mixing ratio in northern mid-latitudes to about half of the seasonality implied by net ecosystem exchange, while amplifying seasonality in the Arctic
- Meridional fluxes of  $\text{CO}_2$  by synoptic storms are of similar magnitude as anthropogenic surface fluxes of  $\text{CO}_2$
- $\text{CO}_2$  transport along the east side of baroclinic waves is systemically unobserved by satellites due to correlations of rising condensing air, precipitation,  $\text{CO}_2$ , and heat transport
- $\text{CO}_2$  transport by synoptic storms is sensitive to modeling of frontal systems, moisture transport in warm conveyors, and vertical mixing
- Model differences in meridional fluxes of  $\text{CO}_2$  by synoptic storms is significant, equal to nearly half of fossil fuel emissions
- Inversion of synthetic satellite data using ensemble like approximation and bias parameter estimation recovers realistic looking sinks, assuming underlying transport is unbiased



- Nearly 100% of sinks are recovered in the Amazon despite large reductions in column CO<sub>2</sub> data due to cloud contamination
- Temporal sampling errors in column CO<sub>2</sub> in northern mid-latitudes cause small but spatially coherent flux estimation errors during summer
- Transport biases are strongly aliased into inverse estimates of sources and sinks

Sensitivity of eddy transport to factors such as storm track position, frontal weather systems, seasonal tendencies in CO<sub>2</sub> mixing ratio and the pattern of seasonal change in surface CO<sub>2</sub> flux over the globe poses a challenging task for inversion modelers. This study has addressed implications of frontal weather systems for the atmospheric carbon cycle and demonstrated the need to represent these systems with high fidelity, but could not link uncertainty of transport by frontal weather systems to a single factor, making it difficult to make specific recommendations for future modeling efforts. Future studies would therefore do extremely well to use a global model capable of running at multiple grid spacings and/or with multiple vertical mixing schemes to examine sensitivity of eddy transport to (1) vertical mixing with grid spacing held constant and (2) grid spacing with vertical mixing held constant. Such a model could be applied rather easily in an ensemble inversion framework, in which an adjoint is not required, to test for aliasing of a single factor into surface flux errors.

This work has touched on, but not yet begun to fully address, the influence of storm track position or seasonal change in surface CO<sub>2</sub> flux on transport by moist synoptic storms. It is observed, for example, that storm track position shifts north and south with annular modes and quite possibly climate change. Using a dry dynamical core of an AGCM, Butler et al. [2010] show that heating in the tropical troposphere might lead to a poleward shift of extratropical storm tracks. Enhancement of moist isentropic circulations with global warming are also possible, as shown in work by [Laliberte and Pauluis, 2010].

Large meridional fluxes of CO<sub>2</sub> as discussed in this research are currently possible

due in large part to the favorable positioning of baroclinic waves with the most seasonally active biological part of the world. Somewhat remarkably, most of the global land area resides in this part of the world, making the northern mid-latitudes one of the most productive regions as well. It seems that the tilt of the planet has provided a mechanism that simultaneously regulates geophysical fluid dynamics and plant metabolism, while plate tectonics has arranged land mass in a way that is favorable not only for the carbon cycle, but people. Responses of the Earth system to dynamical forcing from global warming will almost certainly cause a shift in the balance between geophysical fluid dynamics and surface fluxes of  $\text{CO}_2$ .

## BIBLIOGRAPHY

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. **4**, 135–162.