UNIVERSITY OF CALIFORNIA,
IRVINE

Land-atmosphere feedbacks in the energy, water, and carbon cycles of Earth system models

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Earth System Science

by

Paul Alexander Levine

Dissertation Committee:
Professor James T. Randerson, Chair
Assistant Professor Michael S. Pritchard
Professor Michael L. Goulden

2019

DEDICATION

In loving memory of Soraya Vorster; November 17, 1977 – January 11, 2011

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Funding for this work was provided in part by the National Aeronautics and Space Administration (NASA) Earth and Space Science Fellowship (NESSF) 16-EARTH16F-0196. I also received funding support from the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation Scientific Focus Area (RUBISCO SFA), which is sponsored by the Regional and Global Model Analysis (RGMA) Program in the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research (BER) in the U.S. Department of Energy (DOE) Office of Science. The research in Chapters 3 and 4 used resources from Project m2467 of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility (DE-AC02-05CH11231). Additional funding for Chapter 4 was provided by the European Research Council (ERC) for the 14Constraint project.

I would like to thank my co-authors, past and present. David M. Lawrence, Sean C. Swenson, and James T. Randerson were instrumental in assisting my efforts that culminated in Chapter 2 of this dissertation, and they share with me the copyright to the manuscript published in *Hydrology and Earth System Sciences*, which is adapted here under the CC Attribution 3.0 License. Michael S. Pritchard, Yang Chen, Min Xu, Forrest M. Hoffman, and James T. Randerson worked with me to produce Chapter 3 of this dissertation, which was published in the *Journal of Climate*, and was adapted here with permission from the American Meteorological Society, which holds the copyright. William J. Riley, Qing Zhu, Alison M. Hoyt, Susan E. Trumbore, Steven D. Allison, Zheng Shi, Forrest M. Hoffman, and James T. Randerson have been assisting me as I continue the work that comprises Chapter 4 of this dissertation, and I look forward to their involvement as I complete a manuscript for submission.

I am thankful for the support I received from the Department of Earth System Science at the University of California, Irvine. All of the faculty, staff, and students I have interacted with have collectively contributed to an incredible sense of support and inspiration. In particular, I would like to thank François Primeau and Isabella Velicogna for all of their extra office hours, which I greatly appreciated.

I would also like to thank my Ph.D. committee, Michael L. Goulden and Michael S. Pritchard, whose support and guidance have benefited me greatly. I am grateful for the many hours of assistance provided by our resident Keck AMS radiocarbon experts, Xiaomei Xu, Guaciara Dos Santos, and John Southon. I also owe special thanks to Forrest M. Hoffman, whose munificent counsel and steadfast presence helped me overcome obstacles and stay on course while navigating through this process.

In am immensely appreciative of the support, guidance, and mentorship I received from my advisor, Dr. James T. Randerson. Jim's excitement in the face of scientific curiosity is infectious, and I consider myself extremely fortunate to have this as my inspiration. I have found many doors that were opened to me thanks to his stature and reputation. His preeminence as a scientist is matched only by the grace of his compassion and his commitment to humanity and justice.

I would like to thank all of the people who have made these years in Irvine enjoyable ones. I would

especially like to thank Sungduk Yu and Chris McGuire for accompanying me on many adventures in local and distant hills and for helping me keep the rubber side down (most of the time).

I also owe a tremendous debt of gratitude to all of the teachers who have inspired and educated me in various subjects throughout the years, in particular (but in no particular order), Brian Litwak, Marcia Weiss, Barry Smolin, Mackie Burnette, Sandra Tsing Loh, Howard Cowan, Todd Winkler, Sue Polep, Betty Sazer, Sylvia Pepper, Josh Bender, David Sears, Sid Lasaine, Molly O'Brien, Jan Tullis, Thompson Webb III, and Stanley Trimble.

Finally, I must share the credit for all I have accomplished with my family and friends, whose love and support has been the cornerstone of my experience. The unmitigated inspiration provided by my nieces, Sonia and Lily, along with my sister Maya, was as important to me as the unconditional endorsement from my parents, Natasha and Joe. I also want to thank my uncle Peter, who started me down this path of scientific discovery with a trip to Mono Lake, for helping me break the "Vorster curse" and finish this dissertation. And I want to extend special thanks to those whose frendship and partnership carried me through this process.

CURRICULUM VITAE

Paul Alexander Levine

EDUCATION

Doctor of Philosophy in Earth System Science **2019**
University of California, Irvine *Irvine, California*

Master of Science in Earth System Science **2016**
University of California, Irvine *Irvine, California*

Master of Arts in Geography **2007**
University of California, Los Angeles *Los Angeles, California*

Bachelor of Arts in Geology/Biology **2000**
Brown University *Providence, Rhode Island*

AWARDS

Received NASA Earth and Space Science Fellowship 16-EARTH16F-0196 (2016)

FIRST AUTHOR JOURNAL PUBLICATIONS

Levine, P. A., J. T. Randerson, Y. Chen, M. S. Pritchard, M. Xu, and F. M. Hoffman (2019), Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to El Niño–Southern Oscillation, *Journal of Climate*, *32*(4), 1273–1292. http://dx.doi.org/10.1175/JCLI-D-18-0150.1

Levine, P. A., J. T. Randerson, D. M. Lawrence, and S. C. Swenson (2016), Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations, *Hydrology and Earth System Sciences*, *20*, 4837–4856. http://dx.doi.org/10.5194/hess-20-4837-2016

FIRST AUTHOR CONFERENCE PRESENTATIONS

Levine, P. A., J. T. Randerson, Q. Zhu, W. M. Riley, A. M. Hoyt, S. E. Trumbore, Z. Shi, S. D. Allison, and F. M. Hoffman (2019), Global radiocarbon observations suggest a reduced soil carbon sink in the Energy Exascale Earth System Model. Presented at the 2019 CESM Land Model and Biogeochemistry Working Group Meetings, NCAR, Boulder, Colo., 11–13 Feb. (oral)

Levine, P. A., J. T. Randerson, Q. Zhu, W. M. Riley, A. M. Hoyt, S. E. Trumbore, Z. Shi, S. D. Allison, and F. M. Hoffman (2018), Global carbon-14 observations constrain rates of

soil organic matter decomposition in the Energy Exascale Earth System Model, Abstract B33B-03 presented at 2018 Fall Meeting, AGU, Washington, D.C., 10–14 Dec. (oral)

Levine, P. A., M. Xu, Y. Chen, J. T. Randerson, and F. M. Hoffman (2018), Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to ENSO in CLM4.5. Presented at the 23rd Annual CESM Workshop, Biogeochemistry Working Group Session, NCAR, Boulder, Colo., 18–20 Jun. (oral)

Levine, P. A., M. Xu, Y. Chen, J. T. Randerson, and F. M. Hoffman (2017), Remote SST Forcing and Local Land-Atmosphere Moisture Coupling as Drivers of Amazon Temperature and Carbon Cycle Variability, Abstract H42B-08 presented at 2017 Fall Meeting, AGU, New Orleans, Louisiana, 11–15 Dec. (oral)

Levine, P. A., J. T. Randerson, D. M. Lawrence, and S. C. Swenson (2016), Evaluating Land-Atmosphere Moisture Feedbacks in Earth System Models With Spaceborne Observations, Abstract H51I-1635 presented at 2016 Fall Meeting, AGU, San Francisco, Calif., 11–15 Dec. (poster)

Levine, P. A., C. de Linage, I. Veligocna, and J. T. Randerson (2014), Global spaceborne assessment of the relationship between terrestrial water storage and evaporative demand, Abstract H24F-02 presented at 2014 Fall Meeting, AGU, San Francisco, Calif., 15–19 Dec. (oral)

SYNERGISTIC ACTIVITIES

- Contributor to the International Soil Radiocarbon Database (ISRaD), an international consortium of researchers using observations of carbon-14 to study soil processes
- Honorary fellow in the Machine Learning and Physical Sciences (MAPS) US National Science Foundation Research Traineeship program at UC Irvine
- Former writer and managing editor for *The Loh Down on Science*, a science radio program hosted by Sandra Tsing Loh and syndicated by National Public Radio
- Selected to attend the 2017 ComSciCon National Workshop at Harvard University for graduate student leaders in science communication
- Participant in Climate Literacy Empowerment and Inquiry (CLEAN) Education, a 501(c)(3) non-profit dedicated to teaching the science of climate change to local primary- and secondary-school students
- Co-founder of UC Irvine School of Physical Sciences school-based council for Diverse Educational Community and Doctoral Experience (DECADE), an initiative dedicated to fostering an environment of diversity and inclusivity for graduate students

ABSTRACT OF THE DISSERTATION

Land-atmosphere feedbacks in the energy, water, and carbon cycles of Earth system models

By

Paul Alexander Levine

Doctor of Philosophy in Earth System Science

University of California, Irvine, 2019

Professor James T. Randerson, Chair

Feedbacks between the land surface and the atmosphere are expected to play a role in the response of the Earth system to climate change. Earth system models (ESMs) are designed to simulate these processes, but their implementation is only as good as scientific knowledge and technological capabilities allow. This puts ESMs in a dual role with respect to land-atmosphere feedbacks: On the one hand, they represent the most important tools researchers have for understanding land-atmosphere feedbacks and how they will affect the Earth system. On the other hand, they acquire substantial uncertainty as a result of limitations in their representation of feedback processes.

In this dissertation, I used ESMs within the context of both roles. I used them as tools for understanding how land-atmosphere feedbacks affect the Earth system and our understanding of it. I also used observational data to constrain land-atmosphere feedbacks in ESMs with the goal of evaluating and improving their ability to project future responses to climatic forcing.

The goal of my first science chapter was to evaluate how well ESMs represent seasonal-scale feedbacks between terrestrial water availability and atmospheric conditions. I developed a metric that was designed to use terrestrial water storage data from the Gravity Recovery and Climate Experiment (GRACE) satellites in conjunction with other remote sensing data sources for atmospheric temperature and humidity, precipitation, and downwelling shortwave radiation. First, I used GRACE to identify the months of the year when the land surface loses more water to evapotranspiration

and runoff than it gains from precipitation. I then related the interannual water storage anomaly at the onset of this interval with subsequent atmospheric conditions as a measure of the land surface forcing on the atmosphere, as well as relating the atmospheric conditions with the subsequent water storage anomaly as a measure of the land surface response to atmospheric forcing. By calculating the equivalent quantities in an ensemble of ESMs, I demonstrated that the models tended to overestimate the strength of these relationships. These results are consistent with complementary efforts using other data sources and spatiotemporal scales, and suggest that overly simplistic representations of the heterogeneity within vegetation cover and bare soil could cause models to overestimate soil moisture feedbacks with the atmosphere.

My next chapter was partially motivated by a remaining question from the previous research: How much of the measured relationships was driven directly by interactions between soil moisture and the atmosphere, and how much is due to co-variability due to external forcing such as sea surface temperature (SST) anomalies. An additional motivation stemmed from the well-known relationship between the El Niño-Southern Oscillation (ENSO) and interannual variability in the growth rate of atmospheric CO_2 concentrations. This relationship is attributed primarily to the response of tropical terrestrial ecosystems, and temperature has been widely implicated as the primary driver. However, there is an ongoing debate in the literature suggesting that hydrology is also an important driver, which, as I demonstrated in the previous chapter, is not independent of temperature.

To address these questions, I performed a set of experiments using the Energy Exascale Earth System Model (E3SM). I modified E3SM in order to decouple the interannual variability of soil moisture and SST, which allowed me to isolate the influence of each of these factors in the response of the Amazon rainforest to ENSO. I found that in E3SM, soil moisture served to amplify and extend the land surface response to ENSO in the Amazon. SST anomalies coupled with atmospheric circulation drove an immediate response, which coincides with the Amazon wet season. As the ecosystem was generally not water limited at this time, temperature played the dominant role in carbon cycle variability. However, soil moisture anomalies persisted into the dry season, intensifying

and extending the response of both temperature and the terrestrial CO_2 flux. This highlights the importance of considering the feedback between soil moisture and temperature when considering their relative importance as drivers of CO_2 variability in the Amazon.

For my final science chapter, I used a global dataset of vertically resolved soil radiocarbon observations to evaluate the representation of soil carbon processes in the Community Land Model (CLM) and the E3SM land model (ELM). I found that while ELM slightly overestimated radiocarbon ages in temperate latitudes, it underestimated them in boreal and tropical latitudes, particularly at depth, enough to lead to a young bias globally. CLM, on the other hand, underestimated radiocarbon ages at all depths, latitudes, and vegetation types. This suggests that carbon was cycling through the soil too quickly in the models, leading them to overestimate the rate at which soils could sequester carbon in response to increasing atmospheric CO_2 concentrations.

I used the observed radiocarbon profiles to constrain kinetic rate constants and transfer coefficients in ELM, which improved the young age bias and improved a low bias in the tropical soil carbon stock. Constraining the model increased the global soil carbon stock as a result of the improvements in the tropics, and reduced the contribution of soil to the terrestrial carbon sink. This suggests that soil is likely to respond very slowly to increasing atmospheric CO_2 concentrations, and is unlikely to serve as a short-term carbon sink.

These three chapters illustrate the range applications for ESMs in the context of land-atmosphere feedbacks. ESMs can be evaluated for how well they reproduce observations of relationships between multiple climatic variables, as in the first chapter. ESMs can be used to understand land-atmosphere feedbacks in the Earth system, by isolating factors that are difficult or impossible to disentangle in nature, as in the second chapter. ESMs can be constrained by observations in order to reduce uncertainty in their simulation of land-atmosphere feedbacks, as in the final chapter. This dissertation represents steps toward understanding how energy, water, and carbon flow through terrestrial ecosystems, and improving the representation of these processes in ESMs.

Chapter 1

Introduction

1.1 Feedbacks in the Earth system

One of the major challenges to making accurate predictions of future climatic responses to current anthropogenic forcing is uncertainty in the quantification of feedbacks [*Bony et al.*, 2006; *Dirmeyer et al.*, 2006a; *Soden and Held*, 2006; *Booth et al.*, 2012; *Arora et al.*, 2013; *Bodman et al.*, 2013; *Friedlingstein et al.*, 2014; *Berg et al.*, 2016a]. The concept of feedbacks was adapted from control theory [*Bates*, 2007], and has long been recognized as a feature of the Earth's climate [*Hansen et al.*, 1984; *Woodwell et al.*, 1998], but the terminology can be somewhat ambiguous in the literature [*Sherwood et al.*, 2015]. In this dissertation, I use the term "feedbacks" to encompass all relationships in which perturbations to a given variable cause a chain reaction of events that ultimately affect this variable beyond its initial perturbation. In the case of positive feedbacks, the chain reaction pushes the variable in the same direction as the initial perturbation, and the feedback loop will continue to intensify until some external factor affects one or more elements in the relationship. In the case of negative feedbacks, the chain reaction pushes the variable in the opposite direction as the initial perturbation, yielding a dampening effect.

In the Earth's climate system, multiple interacting relationships can lead to overall feedback loops with both positive and negative components. In these situations, quantification of the total feedback is hindered by the inability to examine individual relationships in isolation. For example, anomalously dry soil conditions would tend to reduce plant growth, which would reduce the withdrawal of water from the soil in a negative feedback loop [*Seneviratne et al.*, 2010]. At the same time, those drier soils may increase temperature, and thereby evaporative demand, which would increase the amount of water that is transferred from the soil to the atmosphere in a positive feedback loop [*Dirmeyer*, 2006, 2011]. In nature, it is extremely difficult to investigate these two relationships independently of one another [*Koster et al.*, 2006].

In addition, some feedback loops occur on time scales for which we lack sufficient observational data to properly quantify [*Manzoni and Porporato*, 2009; *Sun and Wang*, 2012]. Nonlinearities in relationships often means that the strength or even the sign of relationships in the present day may not reflect those quantities in the future [*Wang et al.*, 2014]. Furthermore, feedback relationships may respond to thresholds that have not yet been realized in the historical record [*Seneviratne et al.*, 2010].

1.2 Feedbacks between the land surface and the atmosphere

The research presented here focuses on climate system feedbacks between the land surface and the atmosphere. Many processes in the land surface are driven by atmospheric forcing from air temperature, downwelling radiation, precipitation, humidity, and wind. At the same time, the land surface acts as a lower boundary to atmospheric circulation, and various characteristics (i.e., albedo, roughness length, and stomatal resistance) and processes (i.e., infiltration, evapotranspiration, primary production, and respiration) influence the atmosphere by regulating exchanges of energy, water, and carbon.

In this dissertation, I focus on two specific types of land-atmosphere feedbacks:

- Soil moisture feedback, in which the amount of water in the land surface exerts a limitation on evapotranspiration, thereby influencing the amount of energy and water that are returned to the atmosphere.
- The carbon-concentration feedback, in which atmospheric carbon dioxide (CO_2) concentrations influence the exchange of carbon between the land surface and the atmosphere.

The first limb of the soil moisture feedback, in which soil moisture influences the atmosphere, is frequently referred to in the literature as land–atmosphere coupling [*Koster et al.*, 2006; *Seneviratne et al.*, 2013]. This describes conditions in which soil moisture exerts a limitation on evapotranspiration, thereby controlling how net radiation is partitioned between turbulent fluxes in the planetary boundary layer (PBL). For example, when evapotranspiration is limited by insufficient soil moisture availability, the latent heat flux is decreased and the sensible heat flux is increased, limiting moisture and increasing temperature in the PBL [*Findell and Eltahir*, 2003a]. The sign of the overall feedback loop then depends on the response of the atmosphere to the warmer and drier PBL. Reduced precipitation and/or increased evaporative demand would cause even greater reductions in soil moisture, thus constituting an overall positive feedback loop [*Eltahir*, 1998; *Findell et al.*, 2011]. However, it is also possible that the warmer PBL could enhance convective precipitation [*Findell and Eltahir*, 2003b; *Taylor et al.*, 2012; *Guillod et al.*, 2015], which would lead to soil moisture increases, thus constituting an overall negative feedback loop.

The terrestrial carbon-concentration feedback is frequently referred to in the literature as the CO_2 fertilization effect [*Cox et al.*, 2013; *Keenan et al.*, 2016]. Evidence that enhanced atmospheric CO_2 concentrations stimulate increased vegetation growth has been observed in both field experiments [*Norby and Zak*, 2011] and satellite remote sensing [*Schimel et al.*, 2015]. This negative feedback serves to mitigate rising atmospheric CO_2 [*Keenan et al.*, 2016], but the long-term efficacy of this feedback depends on how carbon is transferred from more vulnerable forms in vegetation to more

stable forms in the soil. Additionally, it is possible that enhanced carbon input to the soil could actually stimulate heterotrophic respiration via priming [*Heimann and Reichstein*, 2008; *Stockmann et al.*, 2013], which would reduce the strength of the negative feedback loop.

1.3 Earth system models as tools to understand feedbacks

Earth system models (ESMs) are important tools for understanding land-atmosphere feedbacks and their implications for a changing climate. When used in a coupled mode, i.e., with a land surface model interacting with an atmospheric circulation model, ESMs simulate the feedbacks between these two systems. ESMs overcome the previously mentioned limitations in two ways: first, they allow the impact of feedbacks to be considered over longer periods of time than the observational record and under conditions with no present-day analog, and second, they allow for mechanisms to be excluded and included in ways that are not possible in the real world [*Koster et al.*, 2006; *Seneviratne et al.*, 2013; *Swann et al.*, 2016; *Berg et al.*, 2016a; *Kooperman et al.*, 2018]. However, uncertainty in the representation of land-atmosphere feedbacks is a major source of divergence among ESM projections of future climate change [*Dirmeyer et al.*, 2006a; *Friedlingstein et al.*, 2014; *Berg et al.*, 2016a].

Given the importance of ESMs as tools in combination with their uncertainty representing land-atmosphere feedbacks, I consider there to be three broad categories of research in this area:

- ESMs can be *evaluated* based on observations in order to determine how well they can represent land-atmosphere feedbacks.
- ESMs can be *manipulated* in order to simulate conditions that are not possible to observe in the real Earth system.
- ESMs can be *improved* in order to better represent important processes.

Each of these three categories is represented by a chapter in this dissertation, with Chapter 2 focusing on *evaluating* feedbacks in multiple ESMs, Chapter 3 focusing on *manipulating* an ESM to tease apart mechanisms that cannot be separated in nature, and Chapter 4 focusing on attempts to *improve* an ESM.

Chapter 2 describes and discusses research I conducted to assess how well ESMs represent the relationships that are part of soil moisture feedbacks. A substantial body of research has been devoted to assessing the strength of land–atmosphere coupling in models [*Koster et al.*, 2006; *Dirmeyer et al.*, 2013], but interpretation of the outcomes has been hindered by a lack of large-scale observational data that allows the modeled relationships to be evaluated at a global scale. To assess the strenth of soil moisture feedbacks at larger spatial scales, I developed a metric that was specifically designed to take advantage of the Gravity Recovery and Climate Experiment (GRACE), a gravimetric satellite remote sensing platform that measures terrestrial water storage anomalies [*Swenson and Wahr*, 2006; *Landerer and Swenson*, 2012].

I combined GRACE data with global observations of temperature, humidity, precipitation, and radiation from multiple additional satellite remote sensing platforms. I used this combination of data to quantify the degree of coherence between terrestrial water storage anomalies, as a proxy for soil moisture, and the above-mentioned atmospheric variables at multiple points in time. I then calculated the same quantities from ESMs for comparison with the observational metrics, and was able to demonstrate that ESMs tend to overestimate the strength of the relationships in this feedback loop compared with the satellite observations, in a manner consistent with previous studies that have relied on different data sources across shorter time scales and/or smaller spatial scales. This chapter is a slightly modified version of the following publication:

Levine, P. A., J. T. Randerson, D. M. Lawrence, and S. C. Swenson (2016), Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations, *Hydrology and Earth System Sciences*, *20*, 4837–4856. http://dx.doi.org/10.5194/hess-20-4837-2016

One of the remaining questions after the research for Chapter 2 was how much the relationships quantified by the metrics resulted directly from coupling between the land surface and the atmosphere relative to the amount they co-varied due to external forcing. In particular, the El Niño-Southern Oscillation (ENSO) is known to exert an influence on both the land surface and the atmosphere, particularly in the tropics, over the seasonal time scales I had focused on. ESMs should ideally be able to reproduce the correct response to both direct land-atmosphere interactions and external forcing from sea surface temperature (SST) anomalies associated with ENSO. The metrics I developed serve as useful model benchmarks by measuring the combined effect of this combination of factors. However, the question of how important each of these factors is relative to the other part of the initial motivation for the research that I describe in Chapter 3.

The primary motivation for the research in Chapter 3 is related specifically to the forcing mechanisms from soil moisture to temperature, and the impact of these processes on the terrestrial carbon cycle. Interannual variability in the growth rate of atmospheric CO_2 is highly correlated with ENSO, and this relationship is attributed primarily to the tropical land surface [*Cox et al.*, 2013]. However, the relative importance of temperature versus hydrological anomalies is an area of ongoing debate in the literature, with some studies arguing that temperature alone is a sufficient explanatory variable [*Cox et al.*, 2013; *Piao et al.*, 2013; *Wang et al.*, 2013, 2014], and others highlighting the importance of precipitation and water storage variability [*Foley et al.*, 2002; *Zeng et al.*, 2005; *Qian et al.*, 2008; *Keppel-Aleks et al.*, 2014; *Wang et al.*, 2016a; *Humphrey et al.*, 2018]. While the temperature anomalies associated with ENSO are well-understood features of atmospheric circulation, it is possible that some portion of those anomalies may be driven by contemporaneous soil moisture anomalies via land–atmosphere coupling. In this case, some portion of the hydrologic influence on the carbon cycle may be obscured by correlated temperature variability.

To assess the importance of this interaction, I conducted an experiment by manipulating an ESM to decouple soil moisture variability from SST variability. I focused on the watershed of the Amazon River, which has a strong response to ENSO that was reproduced well by the ESM. The ESM

manipulation experiment allowed me to quantify how much of the temperature anomalies associated with ENSO was caused directly by remote SST forcing, and how much was caused indirectly by SST-driven soil moisture anomalies via land-atmosphere coupling. I found that the relative importance of each factor depended on the season and location, with the eastern Amazon showing a strong influence of land–atmosphere coupling on temperature, and a strong carbon cycle response associated with soil moisture variability. This chapter is a slightly modified version of the following publication:

> Levine, P. A., J. T. Randerson, Y. Chen, M. S. Pritchard, M. Xu, and F. M. Hoffman (2019), Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to El Niño–Southern Oscillation, *Journal of Climate*, *32*(4), 1273–1292. http://dx.doi.org/10.1175/JCLI-D-18-0150.1

The research conducted for Chapter 4 focused on the carbon-concentration feedback, and specifically, improving how part of that feedback is represented in an ESM. While most work on the carbon-concentration feedback in ESMs primarily considers the response of vegetation to elevated atmospheric CO_2 concentrations [*Hajima et al.*, 2014; *Sun et al.*, 2014], there has been less attention paid to its consequences for long-term carbon storage in the soil [*Todd-Brown et al.*, 2013]. This may be partially due to the large uncertainties in the highly simplistic representations of soil carbon in most ESMs. However, as ESMs adopt more sophisticated approaches toward modeling carbon, it is essential to evaluate their simulations relative to observations, and improve them where possible.

In Chapter 4, I describe and discuss my attempt to improve soil carbon representation in an ESM with a relatively advanced soil carbon parameterization [*Koven et al.*, 2013]. I used a large database of soil carbon observations that include measurements of the radioactive isotope carbon-14 (radiocarbon). Radiocarbon gives information about the rates at which carbon cycles from the atmosphere through the land surface [*Trumbore*, 2009], which is an important characteristic in ESMs to which they are highly sensitive [*Friedlingstein et al.*, 2014]. I developed a method for calibrating an ESM using the radiocarbon observations, in order to improve its representation of the soil carbon cycle and,

ultimately, the carbon-climate feedback.

The calibrated ESM, by virtue of being observationally constrained with both radiocarbon ages and total carbon stocks, reduces uncertainty in the simulated response to elevated atmospheric CO_2 concentrations and global temperatures. The particular ESM I used for the calibration was able to reproduce observed radiocarbon profiles comparatively well in its default parameterization, so the calibration yielded relatively small changes. Nevertheless, the small simulated 20^{th}-century soil carbon sink in the default parameterization was reduced even further with the calibrated parameters. This suggests that ESMs are likely to be overestimating global soil carbon accumulation, and that soils are unlikely to serve as an effective sink for anthropogenic CO_2 emissions.

Finally, in Chapter 5, I discuss the implications of my dissertation research, and potential directions for future research.

Chapter 2

Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations

2.1 Introduction

Land–atmosphere feedbacks can result from the coupling of the terrestrial moisture state with temperature, precipitation, or radiation [*Findell and Eltahir*, 1997; *Koster et al.*, 2004; *Betts et al.*, 2014; *Guillod et al.*, 2015]. Land–atmosphere coupling occurs when terrestrial moisture anomalies influence the partitioning of surface energy between latent and sensible heat fluxes that, in turn, influence the development of the planetary boundary layer (PBL) [*Seneviratne et al.*, 2010]. Temperature coupling generally leads to a positive feedback, with wetter soil contributing to a higher evaporative fraction (EF; the ratio of the latent heat flux to the sum of the sensible and latent

heat fluxes), a lower surface temperature, and decreased evaporative demand [*Hirschi et al.*, 2011; *Miralles et al.*, 2012]. Precipitation coupling can lead to both positive and negative feedbacks, as the influence of EF on the development of the PBL can serve to either enhance or suppress cloud formation and precipitation [*Findell and Eltahir*, 2003b; *Guillod et al.*, 2015]. Cloud radiative coupling can likewise lead to positive or negative feedbacks as insolation and evaporative demand, as a function of cloud cover, are either enhanced or suppressed [*Betts*, 2009; *Cheruy et al.*, 2014]. Temperature, precipitation, and radiation feedbacks each stem from coupling between terrestrial moisture and evapotranspiration (ET), which occurs most strongly in conditions of intermediate moisture availability [*Seneviratne et al.*, 2010].

Evidence of these feedbacks has been observed in both in situ and remotely sensed data [*Findell and Eltahir*, 1997; *Eltahir*, 1998; *Guillod et al.*, 2014, 2015]. Some observational analyses have found land–atmosphere feedback strength to be relatively weak compared to the influence of large-scale atmospheric forcing [*Alfieri et al.*, 2008; *Phillips and Klein*, 2014]. Other observational studies have highlighted the role of these feedback mechanisms in the initiation and exacerbation of climatic extremes such as droughts and heat waves [*Hirschi et al.*, 2011; *Miralles et al.*, 2012; *Whan et al.*, 2015].

2.1.1 Land–atmosphere coupling in climate models

Large-scale land–atmosphere coupling in general circulation models has been demonstrated by a series of experiments from the Global Land–Atmosphere Coupling Experiment (GLACE) project [*Koster et al.*, 2004, 2006; *Guo et al.*, 2006]. The GLACE efforts found that coupled climate models differed greatly in the extent to which soil moisture variations affect precipitation and surface air temperature, but models generally agreed on the spatial distribution of relative coupling strength, with "hotspots" of strong coupling during boreal summer found in the central United States, northern Amazonia, the Sahel, western Eurasia, and northern India. These hotspots were found in regions of

intermediate soil wetness, which is consistent with the understanding that strong land–atmosphere coupling occurs under conditions in which terrestrial moisture availability limits ET [*Seneviratne et al.*, 2010]. GLACE efforts also showed that correct soil moisture initialization improves seasonal forecast skill of temperature and, to a lesser extent, precipitation, particularly in cases with a large initial soil moisture anomaly [*Koster et al.*, 2010, 2011].

Additional studies have considered land–atmosphere feedbacks in the coupled Earth system models (ESMs) used by the Intergovernmental Panel on Climate Change (IPCC) [*Seneviratne et al.*, 2006, 2013; *Notaro*, 2008; *Dirmeyer et al.*, 2013]. *Notaro* [2008] was able to confirm the boreal summer GLACE hotspots, as well as identify several additional austral summer hotspots, in the models used for the Fourth Assessment Report of the IPCC. Analysis of long-term projections from the phase 5 of the Coupled Model Intercomparison Project (CMIP5) indicated an increased control of land surface moisture on boundary layer conditions with climate change [*Dirmeyer et al.*, 2013]. The GLACE-CMIP5 experiment found that modeled coupling strength plays an important role in simulated response to global warming, with greater warming evident in more strongly coupled models due to interactions between soil moisture, temperature, and precipitation [*Seneviratne et al.*, 2013; *Berg et al.*, 2015; *May et al.*, 2015].

Despite the importance of land–atmosphere coupling in both short-term predictability of climatic extremes and long-term uncertainty in climate change, validation efforts have suggested that climate models may not be correctly representing the strength, and in some cases even the sign, of these feedbacks [*Ferguson et al.*, 2012; *Hirschi et al.*, 2014]. The metrics developed for the GLACE are based on model experiments with no direct observational equivalents. However, correlation-based metrics that do enable direct comparison with observations suggest that models may overestimate land–atmosphere coupling strength [*Dirmeyer et al.*, 2006a]. *Zeng et al.* [2010] found that version 3 of the Community Climate System Model (CCSM3) showed a higher coupling strength than reanalysis or observational data. *Mei and Wang* [2012] found that coupling strength was reduced when the Community Atmosphere Model (the land surface component of CCSM3) was updated

from version 3 (CAM3) to version 4 (CAM4), though the coupling strength of the updated version was still stronger than observations and reanalysis.

The Local Land–Atmosphere Coupling (LoCo) project has focused on developing a suite of metrics for diagnosing land–atmosphere coupling strength in observations and models. LoCo metrics consider both the influence of soil moisture on EF, and the influence of EF on diurnal-scale boundary layer development [*Santanello et al.*, 2009]. *Ferguson et al.* [2012] used the LoCo approach to compare global remote sensing data sets of soil moisture, EF, and the lifting condensation level with several land surface models and reanalyses. They found that even though the models were able to simulate the correct spatial pattern of stronger coupling in moist–arid transitional regions, the models tended to simulate a stronger influence of soil moisture on surface turbulent fluxes than what was observed in the satellite data. *Guillod et al.* [2014] used a combination of flux tower, remote sensing, and reanalysis data sets to demonstrate that the measured strength of coupling between EF and precipitation depends greatly on the data source and scale, and that a strong coupling apparent in a previous analysis [*Findell et al.*, 2011] was not consistent with the observations.

While many of the previously mentioned studies have confirmed the long-standing suspicion that models may overestimate coupling strength relative to observations, more recent work has indicated that observations and models may not even agree on the sign of the precipitation feedback. *Taylor et al.* [2012] performed a spatial analysis of the relationship between soil moisture and afternoon precipitation using data from remote sensing, reanalysis, and coupled models. They found evidence of a negative feedback in the remote sensing observations, with afternoon rain being more likely over regions of drier soil, as opposed to the positive feedback that was apparent in the models. *Guillod et al.* [2015] addressed these findings by replicating the spatial analysis and complementing it with a temporal analysis. They found a negative spatial feedback, consistent with the one found by *Taylor et al.* [2012], but a positive temporal feedback, with afternoon precipitation at a given location being more likely after mornings of relatively moist soil.

These studies highlight the need for continued efforts toward evaluating the coupling strength of

models relative to observations using a wide array of data sources at a range of spatiotemporal scales. Apparent coupling strength depends greatly on the spatial scales of analysis [*Hohenegger et al.*, 2009], indicating that observations at the scale of flux towers should not be expected to yield the same coupling strength as those at the scale of global climate models [*Guillod et al.*, 2014]. Consistency between the spatial scale of observations and models is greatly assisted by Earth observation satellites that have been continuously monitoring several relevant land surface and atmospheric variables over multiple years [*Teixeira et al.*, 2014]. Measured or modeled coupling strength will also depend on the timescales in question [*Guillod et al.*, 2015], and while the LoCo efforts have improved the understanding of synoptic and diurnal-scale mechanisms, there is an additional need to examine these processes on seasonal to interannual time periods.

2.1.2 New metrics to benchmark land–atmosphere coupling in models

Here we introduce a set of metrics for measuring the strength of land–atmosphere interactions on seasonal timescales by combining satellite remote sensing data sets of terrestrial water storage, precipitation, shortwave radiation, and surface atmospheric temperature and water vapor during 2002–2014. These new metrics complement previous studies and are unique in several ways. In particular, we designed our metrics to consider interannual variability of entire seasons in order to complement the temporal resolution of LoCo metrics, which focus on day-to-day variability within one or more seasons. Land–atmosphere coupling on seasonal timescales has been shown to be essential in enabling tropical forests to survive during the dry season in the Amazon [*Lee et al.*, 2005] and as a mechanism enabling seasonal forecasts of fire risk [*Chen et al.*, 2013, 2016].

Until recently, studies using remote sensing data to look for evidence of land–atmosphere coupling relied on products that provide information about surface soil moisture [*Ferguson et al.*, 2012; *Taylor et al.*, 2012]. Consideration of root-zone soil moisture has recently been accomplished only indirectly via data-assimilated estimates [*Guillod et al.*, 2015]. The inability to directly consider root-zone

soil moisture has been suggested as an explanation for the relatively weak coupling observed using remote sensing data [*Hirschi et al.*, 2014]. In order to include root-zone soil moisture, as well as other sources of moisture available across entire seasons, here we analyzed remote sensing data of the entire terrestrial water storage (TWS) column.

The metrics introduced here were specifically designed to use the monthly TWS anomaly (TWSA) product from the Gravity Recovery and Climate Experiment (GRACE) mission [*Wahr et al.*, 2004; *Landerer and Swenson*, 2012]. The GRACE TWSA product integrates soil moisture at all layers along with surface, canopy, snow/ice, and aquifer storage, as each of these components represents a potential source of moisture for fulfilling evaporative demand. For example, in areas where agricultural ecosystems are important, diversion of lake and river water resources and withdrawal from aquifers may contribute to irrigation fluxes and thus ET. Furthermore, surface storage of liquid water and snow represents sources of water that are available for and potentially limiting to ET. Under these conditions, month-to-month TWS anomalies capture portions of the terrestrial water cycle that soil moisture alone may not.

Previous studies have largely focused on land surface moisture availability as a forcing mechanism on the atmosphere, as this relationship has important implications for seasonal predictability as well as the projection of the frequency and severity of climatic extremes. However, the land surface response to the atmosphere is governed by many of the same processes through which terrestrial moisture availability forces atmospheric conditions, and it determines the conditions that drive subsequent land surface forcing. It is therefore critical to assess the response of land surface moisture to atmospheric conditions, as an accurate representation of these processes is essential for generating the correct terrestrial moisture variability that will go on to influence the atmosphere. As far as we can tell, this response limb of the land surface feedback loop has not been systematically integrated with existing analyses of land–atmosphere coupling strength.

Our globally applicable approach used the annual cycle of TWS drawdown and recharge to isolate the months of the year during which the land surface loses moisture, which we refer to as the drawdown

Table 2.1: Remote sensing products used for analysis.

Variable	Abbr.	Data product	Spatial	Temporal	Reference
Terrestrial water storage	TWS	GRACE Tellus RL05.1	1°	monthly	*Landerer and Swenson* [2012]
Vapor pressure deficit	VPD	AIRS AIRX3STM v6	1°	monthly	*Susskind et al.* [2014]
Precipitation	PPT	GPCP 1DD v1.2	1°	daily	*Huffman et al.* [2009]
Downwelling shortwave radiation	SW↓	CERES EBAF Ed2.8	1°	monthly	*Loeb et al.* [2009]

interval (Figure 2.1a). We selected this interval because past work has shown that the land surface's influence on the atmosphere is most prevalent during summer in the Northern Hemisphere [*Cheruy et al.*, 2014; *Phillips and Klein*, 2014] and during the dry season in tropical forests [*Harper et al.*, 2013; *Lorenz and Pitman*, 2014]. This approach allowed us to investigate land surface coupling at a global scale, and to extend metrics developed in previous work for pre-defined monthly intervals corresponding to boreal summer [e.g., *Koster et al.*, 2006; *Guo and Dirmeyer*, 2013] to be applicable to any seasonality.

In our analysis, separate metrics were calculated to consider the influence of TWS at the onset of the drawdown interval on atmospheric conditions in subsequent months, and simultaneously, the influence of atmospheric conditions during the drawdown interval on terrestrial water storage at the end of the season. We refer to these two relationships as the forcing and response limbs, respectively, of the fully coupled feedback loop between the land surface and the atmosphere (Figure 2.1). We estimated the strength of these feedbacks during 2002–2014 using GRACE and other satellite remote sensing data (Table 2.1). We then used the satellite observations to evaluate the strength of these feedbacks in the Community Earth System Model (CESM) Large Ensemble (LENS) [*Kay et al.*, 2014] and in several models that contributed simulations to CMIP5 (Table 2.2).

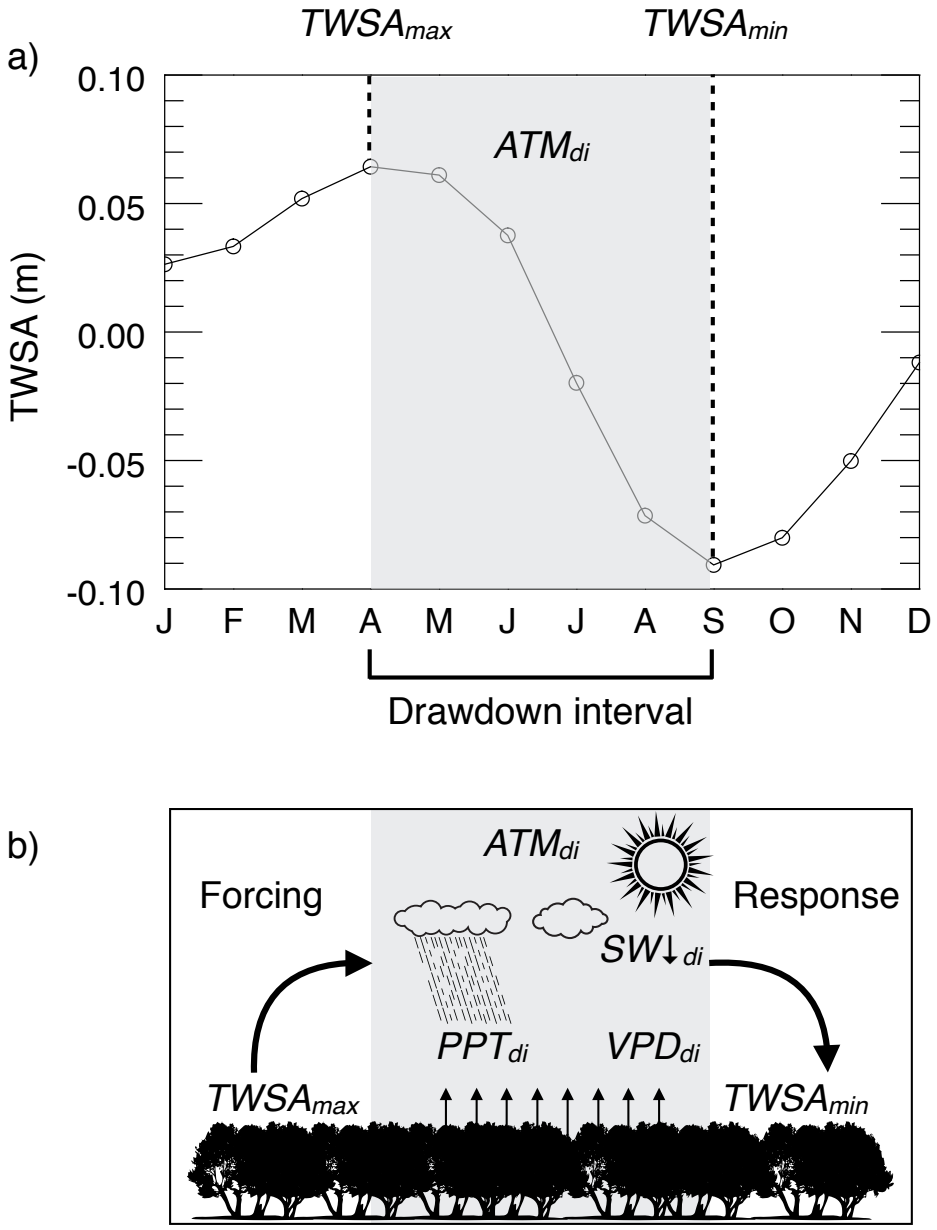

Figure 2.1: Conceptual description of coupling metrics: **(a)** example TWSA climatology from a typical midlatitude location in central North America (38° N, 92° W) illustrating the definition of the drawdown interval as the months from the maximum TWSA through the minimum TWSA. $TWSA_{max}$ and $TWSA_{min}$ are the TWSA values (in units of water height) during the maximum and minimum months, respectively, and ATM_{di} is the atmospheric variable of interest averaged across the months of the drawdown interval. **(b)** Representation of the interactions between TWS and atmospheric component, demonstrating the forcing limb of the feedback loop, in which $TWSA_{max}$ forces subsequent atmospheric conditions, as well as the response limb, in which $TWSA_{min}$ responds to the atmospheric state during the drawdown interval.

2.2 Methods

2.2.1 Remote sensing data

We obtained level 3 TWSA data from GRACE using the University of Texas at Austin Center for Space Research (CSR) spherical harmonic solutions [*Swenson*, 2012]. Global land data at a 1° resolution were scaled using the coefficients provided by *Landerer and Swenson* [2012]. The study period was limited to September, 2002 through November, 2014, in order to minimize temporal gaps. GRACE data during the study period included eight non-consecutive and two consecutive missing months, which were filled using linear interpolation. At each grid cell, the TWSA time series was decomposed into linear trend, seasonal cycle, and interannual variability components using ordinary least squares regression. This decomposition allowed us to estimate a mean annual cycle at each grid cell with minimal influence of any long-term trend.

Level 3 near-surface temperature and relative humidity were obtained globally at a monthly, 1° resolution from the ascending (daytime) orbit of the Atmospheric Infrared Sounder (AIRS) platform [*Susskind et al.*, 2014]. Vapor pressure deficit (VPD) was calculated from the AIRS data using the August–Roche–Magnus approximation to the Clausius–Clapeyron relation [*Lawrence*, 2005]. Precipitation (PPT) data were obtained from the Global Precipitation Climatology Project (GPCP), a merged satellite and gauge-based data set [*Huffman et al.*, 2009], at a daily, 1° resolution and then integrated monthly. Downwelling shortwave radiation (SW↓) was obtained globally at a monthly, 1° resolution from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) surface product [*Loeb et al.*, 2009]. More information describing the remote sensing and reanalysis data products used in our analysis is summarized in Table 2.1.

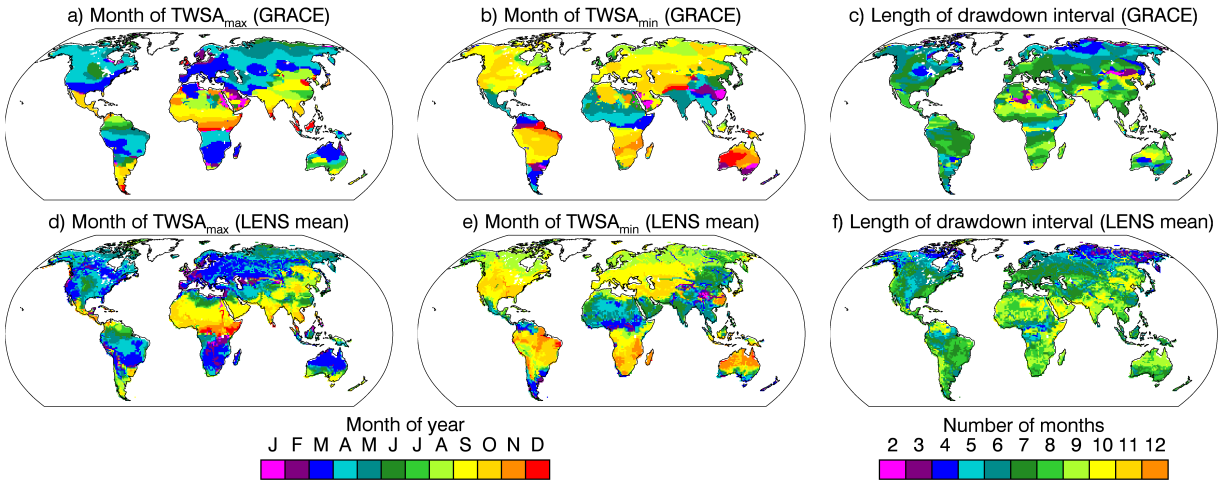

Figure 2.2: Month of maximum and minimum TWSA and the length of the drawdown interval from GRACE **(a–c)** and the LENS ensemble mean **(d–f)**. Months of maximum and minimum were based on the climatology of detrended TWSA over the 146 months in the GRACE record.

2.2.2 Drawdown interval

As a first step, we used the mean annual cycle from GRACE to determine the months of the maximum and minimum TWS anomalies in order to define the drawdown interval at each 1° land grid cell (Figure 2.2). Northern Hemisphere middle and high latitudes exhibited a drawdown interval beginning in the spring (MAM, March–April–May) and ending in the late summer or fall (ASO, August–September–October), reflecting the timing of the boreal summer growing season. At lower latitudes, the North American, African, and Asian monsoons were evident, with Mexico, India, and the Sahel showing a drawdown interval beginning in September, after the monsoonal precipitation has peaked, and ending the following spring after the winter dry season. The onset of the drawdown interval reversed abruptly at the Equator in Africa and Asia, with the drawdown interval reflecting a winter dry season in the austral low latitudes transitioning to a summer growing season in the austral midlatitudes. Within the months of our study period, the portion of land grid cells that experience 11, 12, and 13 complete drawdown intervals are 9.4, 90.5, and 0.1%, respectively.

Existing literature generally defines “land–atmosphere coupling” as the extent to which atmospheric

conditions are forced by the land surface state, and would use the term "atmosphere–land coupling" to refer to the land surface response to atmospheric drivers [*Seneviratne et al.*, 2010]. In this study, we develop what we refer to as "coupling metrics" to indicate the strength of both limbs of the fully coupled land–atmosphere feedback loop. We use the terms "forcing" and "response" to indicate whether we are considering the forcing of the atmosphere by the land surface, or the response of the land surface to the atmosphere (Figure 2.1b).

We defined our forcing metric as the Pearson product-moment correlation coefficient between the TWS anomaly at the onset of the drawdown interval ($TWSA_{max}$) and the surface atmospheric conditions during the drawdown interval (abbreviated here as ATM_{di}). In our analysis, we selected three variables to represent the atmospheric state: VPD, SW↓, and PPT. These atmospheric variables were averaged during the drawdown interval, including during the months of climatological maximum and minimum TWSA. We chose these variables because they represent various aspects of evaporative supply (PPT) and demand (VPD and SW↓).

Similarly, we defined our response metric as the correlation coefficient between ATM_{di} and the land surface state at the end of the drawdown interval ($TWSA_{min}$). Although most previous diagnoses of land–atmosphere coupling has focused on the forcing limb, we argue the response limb is equally important as a metric for model evaluation. Specifically, if variability in the balance between evaporative supply and demand does not lead to the correct TWS variability, then the incorrect TWS response will feed back into subsequent forcing on the atmosphere.

We note that these metrics do not provide distinctive information for measuring the strength of land–atmosphere coupling or the land surface response. While the metrics include the influence of direct land–atmosphere interactions, they are also potentially influenced by atmospheric and soil moisture persistence, as well as remote forcing from sea surface temperatures (SSTs) [*Orlowsky and Seneviratne*, 2010; *Mei and Wang*, 2011]. Nevertheless, these metrics may still serve as useful benchmarks against which to evaluate the ability of ESMs to reproduce the proper relationships based on the combination of these factors.

Here we note that our evaluation of both the forcing and response metrics will follow a nomenclature that considers strong coupling as acting in the direction of an overall positive feedback loop. In regions with a strong positive feedback, higher than average TWS would be followed by lower than average VPD, as more available water is able to fulfill evaporative demand. Therefore, strong TWS forcing on VPD would be associated with a negative correlation coefficient. Higher VPD during the drawdown interval would increase evaporative demand, potentially leading to a more negative TWS anomaly; therefore, a strong response of the land surface to VPD would also be associated with a negative correlation coefficient.

Because the partitioning of surface fluxes can, depending on the spatiotemporal scale, cause a change of either sign to cloudiness and precipitation [*Taylor et al.*, 2012; *Guillod et al.*, 2015], correlation coefficients of either sign could indicate strong land surface forcing on PPT and SW↓. However, the response metrics would be expected to show greater consistency. Higher PPT during the drawdown interval would be expected to increase TWS (positive correlation), while higher SW↓ would increase evaporative demand, thereby decreasing TWS (negative correlation). Therefore, to maintain consistent nomenclature based on evaluating the strength of a positive moisture feedback, we consider strong coupling in both the forcing and response metrics to be associated with a positive correlation in the case of PPT and a negative correlation in the case of SW↓.

2.2.3 Community Earth System Model Large Ensemble

We used the metrics described above to evaluate feedback strength in the CESM LENS. LENS comprises an ensemble of 38 fully coupled runs in which air temperature initial conditions are perturbed slightly (by an amount less than the round-off error) to reveal the internal variability inherent within the coupled climate model. LENS has demonstrated that the uncertainty in climate projections due to internal climate variability inherent in CESM is comparable to the ranges of output within the entire CMIP5 experiment [*Kay et al.*, 2014]. LENS uses version 1 of CESM (CESM1)

with version 5 of the Community Atmosphere Model (CAM5) and version 4 of the Community Land Model (CLM4) at a horizontal resolution of 1°. The ensemble run follows protocols from the CMIP5 experiment, with historical radiative forcing for the 20th century and representative concentration pathway 8.5 (RCP8.5) forcing for the 21st century.

The LENS data were chosen as a starting point for feedback evaluation for two reasons. First, the availability of a TWS variable in these simulations enabled a direct comparison with metrics derived using data from GRACE. The TWS field in CLM4 included water from surface and canopy storage, snow and ice, soil moisture, and a dynamic aquifer, in addition to river water storage terms from the coupled River Transport Module (RTM). The coupling of CLM4 with RTM has been shown to be important for simulating both the annual cycle and interannual variability of TWS in comparison with GRACE [*Kim et al.*, 2009].

Second, the ensemble allowed us to test the importance of internal model variability for the diagnosis of feedback strength. Because the complete satellite record was relatively short (containing no more than 12 drawdown intervals at any location), comparison with an equivalent single time series of model output could be influenced by a model's internal decadal-scale variability [*Kay et al.*, 2014]. Analyzing the full ensemble from LENS enabled us to assess the sensitivity of our forcing and response metrics to this variability. We extracted from each ensemble member the equivalent months of the satellite record, with data prior to December 2005 coming from the historical runs, and data from January 2006 onward coming from the RCP8.5 simulations.

2.2.4 Assessment of uncertainty

To assess the sensitivity of our metrics to observational uncertainty, we used a Monte Carlo sampling approach. For each of the 38 members of LENS, we calculated coupling metrics 10 times with random noise added to both TWSA and atmospheric variable time series at each grid cell. The noise was randomly generated from a Gaussian distribution with a mean of zero and a standard deviation

equal to 25% of the standard deviation of the original data. Comparing these results with those from the unaltered data provided some indication of how much our coupling metrics are degraded by random noise as an approximation of observational uncertainty.

In addition, to assess how our analysis may be influenced by uncertainty due to the selection of satellite data, we substituted data from the European Centre For Medium-range Weather Forecasting (ECMWF) Interim Reanalysis (ERA-Interim) [*Dee et al.*, 2011] in place of AIRS-, GPCP-, and CERES-derived variables. We only used atmospheric reanalysis data for this sensitivity analysis, as these data benefit from assimilation of observations, while we continued to use GRACE for TWSA. Comparing results from this GRACE-reanalysis hybrid to those using only satellite data provided a general ndication of how sensitive our coupling metrics were to the data source.

2.2.5 CMIP5 analysis

To extend our analysis to models that did not output an explicit TWS field, we compared accumulated residuals of precipitation, evapotranspiration, and total runoff (surface and subsurface) with the explicit TWS variable in the LENS simulations. We also compared coupling metrics calculated from LENS using accumulated residuals with those calculated from the explicit TWS field. After we determined that the accumulated residuals of the water balance represented much of the variability in the explicit TWS variable and yielded coupling metrics with similar distributions within LENS, we calculated equivalent metrics for several model simulations in the CMIP5 archive (Table 2.2). We selected the CMIP5 models that were similar to LENS (CESM1-CAM5 and CESM1-BGC) as well as the models that participated in the GLACE-CMIP5 experiment [*Seneviratne et al.*, 2013] for which each necessary output field was available (CCSM4, GFDL-ESM2M, GFDL-ESM2G, and IPSL-CM5A-LR).

Table 2.2: CMIP5 models used for analysis.

Model acronym	Atmospheric model	Land surface model	Horizontal grid	Ensemble size	Reference(s)
CCSM4	National Center for Atmospheric Research (NCAR) Community Atmosphere Model version 4 (CAM4)	Community Land Model version 4 (CLM4)	288×192	6	*Lawrence et al.* [2011]; *Neale et al.* [2013]
CESM1-CAM5	NCAR Community Atmosphere Model version 5 (CAM5)	CLM4	288×192	3	*Lawrence et al.* [2011]; *Meehl et al.* [2013]
CESM1-BGC	NCAR CAM4 with biogeochemistry	CLM4	288×192	1	*Lawrence et al.* [2011]; *Neale et al.* [2013]; *Lindsay et al.* [2014]
IPSL-CM5A-LR	Laboratoire de Météorologie Dynamique atmospheric model (LMDZ5A)	Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE)	96×96	3	*Cheruy et al.* [2013]; *Dufresne et al.* [2013]; *Hourdin et al.* [2013]
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory (GFDL) Earth System Model 2 (ESM2)	Land Model 3.0 (LM3.0)	144×90	1	*Shevliakova et al.* [2009]; *Dunne et al.* [2012]
GFDL-ESM2M	GFDL ESM2	LM3.0	144×90	1	*Shevliakova et al.* [2009]; *Dunne et al.* [2012]

2.3 Results

2.3.1 Drawdown interval and interannual variability

A comparison of the months of maximum and minimum terrestrial water storage as determined by climatologies of GRACE and the LENS ensemble mean indicated that the model largely reproduces the timing of TWSA seasonality evident in the satellite observations (Figure 2.2). Geographic patterns of seasonality were consistent between the model and observations, though a phase shift in the drawdown interval is apparent in eastern Canada and central Eurasia where LENS had a 1-month early bias for both the maximum and minimum TWSA, in southeast North America where the onset of the modeled drawdown interval was slightly later than the observations, and in parts of east Asia and Australia where the modeled drawdown interval ended earlier than in the observations. However,

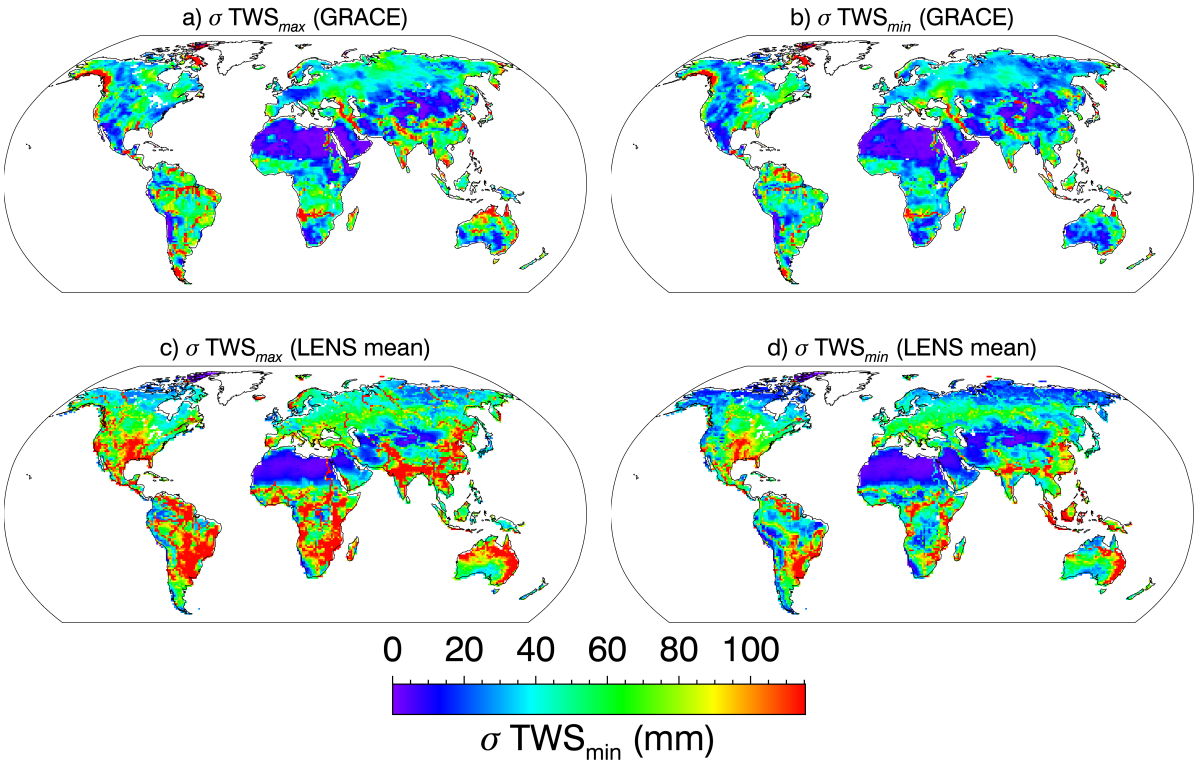

Figure 2.3: Interannual variability (standard deviation) of $TWSA_{max}$ and $TWSA_{min}$ from GRACE **(a, b)** and the LENS ensemble mean **(c, d)**.

despite capturing generally correct timing, the model exhibited higher interannual variability of $TWSA_{max}$ and $TWSA_{min}$ across the 11–12 drawdown intervals compared with the satellite data (Figure 2.3) particularly in the southern United States, southern South America, central and eastern Africa, southern Asia, and eastern Australia. One possible explanation for this is the presence of multi-year trends in aquifer storage in CLM4 that are not consistent with GRACE [*Swenson and Lawrence*, 2015].

A comparison of the interannual variability of atmospheric variables across multiple drawdown intervals between the model and satellite data showed various degrees of consistency (Figure 2.4). The magnitude and geographic pattern of VPD_{di} was generally consistent, though LENS showed greater interannual variability than AIRS in central and western North America, South America, northern and southern Africa, and southern Asia. In the case of PPT_{di}, LENS showed less interannual variability than GPCP in southeast North America and much of South America, but the two were

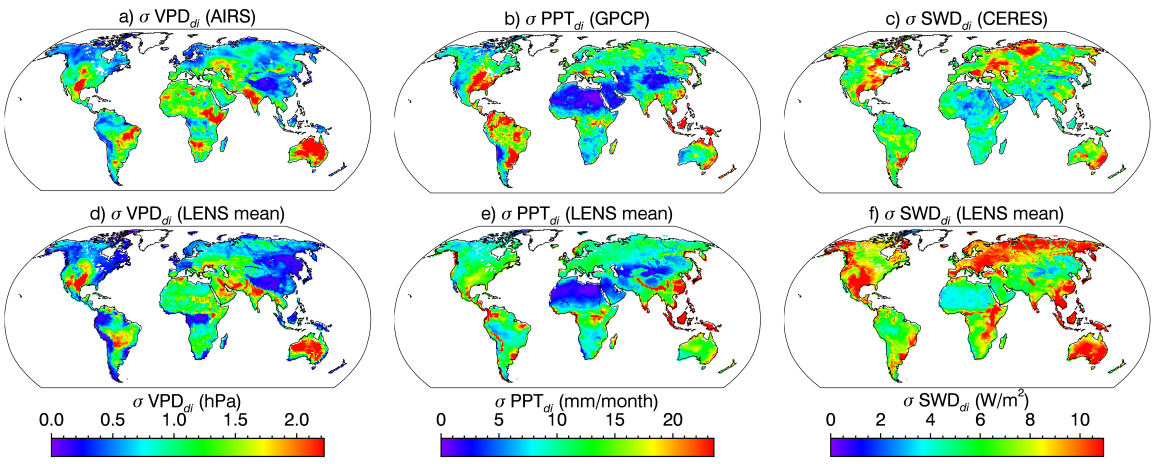

Figure 2.4: Interannual variability (standard deviation) of atmospheric variables: VPD_{di} from AIRS **(a)**, PPT_{di} from GPCP **(b)**, $SW{\downarrow}_{di}$ from CERES **(c)** and the equivalent quantities from the LENS ensemble mean **(d–f)**.

largely consistent elsewhere. $SW{\downarrow}_{di}$ was the least consistent between the model and satellite data, as LENS showed greater interannual variability than CERES in southern North America, northern Eurasia, most of Africa, and most of Australasia.

Comparing both the timing of TWS dynamics and the interannual variability of TWS and the atmospheric variables between the observations and model output provides a context for interpreting the correlation-based metrics we present next. Although there are some inconsistencies, as noted above, the model largely reproduced the same patterns evident in the remote sensing data. In many regions, the interannual variability in model output was similar to the observed variability, indicating that CESM was able to simulate reasonably well the baseline properties (timing and variability) that influence feedback dynamics.

2.3.2 Evaluating feedbacks for a single model simulation

The forcing metric for VPD derived from GRACE and AIRS showed regions of strong coupling, in which $TWSA_{max}$ was negatively correlated with VPD_{di}, in the northern Great Plains, northern

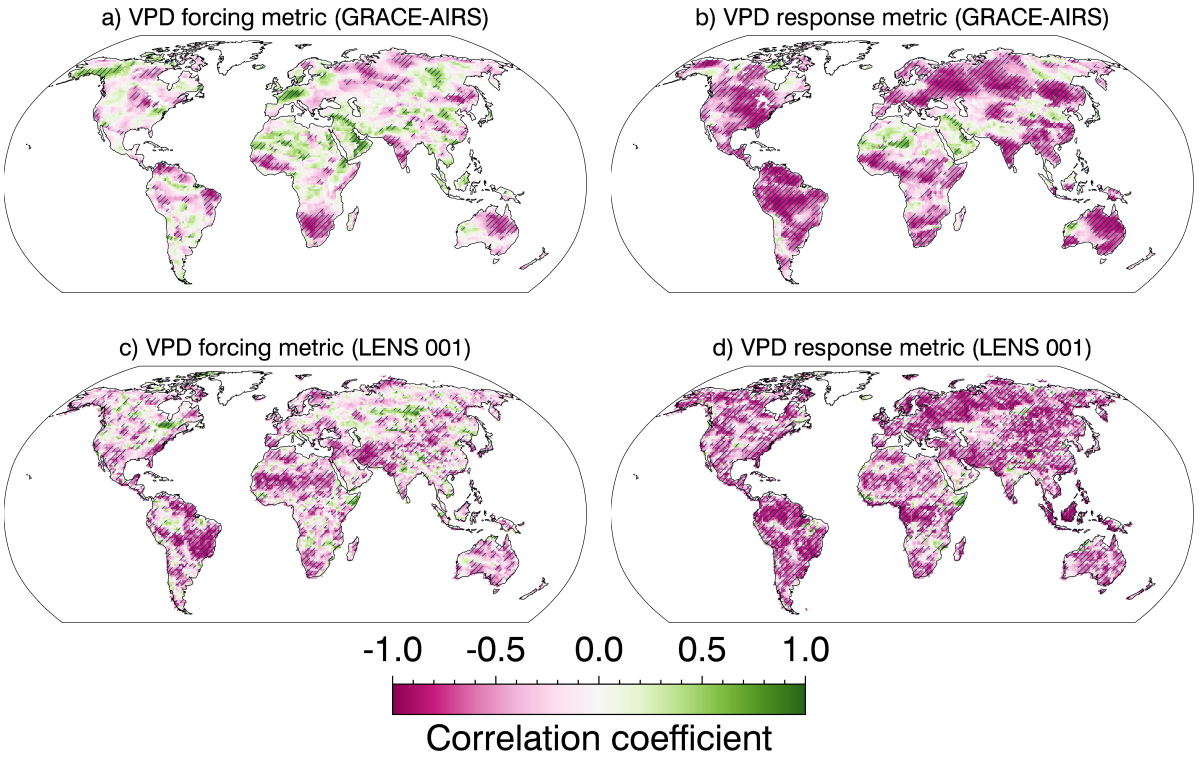

Figure 2.5: Forcing and response metrics for VPD from GRACE/AIRS **(a, b)** and LENS ensemble member 001 **(c, d)**. Cross-hatching indicates a correlation coefficient that is statistically significant at $p \leq 0.05$.

South America, southern Africa, southern and western India, north central Eurasia, and northern Australia (Figure 2.5a). Regions with strong positive correlation were much less common, and were largely confined to areas of very low GRACE-derived TWSA_{max} variability (Figure 2.2a). Positive correlations are unlikely to reflect direct land–atmosphere coupling. Instead, they demonstrate how remote SST forcing, depending on persistence and time delays with atmospheric responses, can lead to apparent negative relationships such as those demonstrated by [*Wei et al.*, 2008]. In comparison with the satellite data, the VPD forcing metrics from the first ensemble member of LENS (Figure 2.5c) showed much stronger coupling in the southern and eastern Amazon, and marginally stronger coupling strength across many regions in temperate Asia.

The response metrics for VPD showed much stronger coupling than the forcing metrics in both the satellite data and the model (Figure 2.5b and d). Satellite data yielded negative correlation

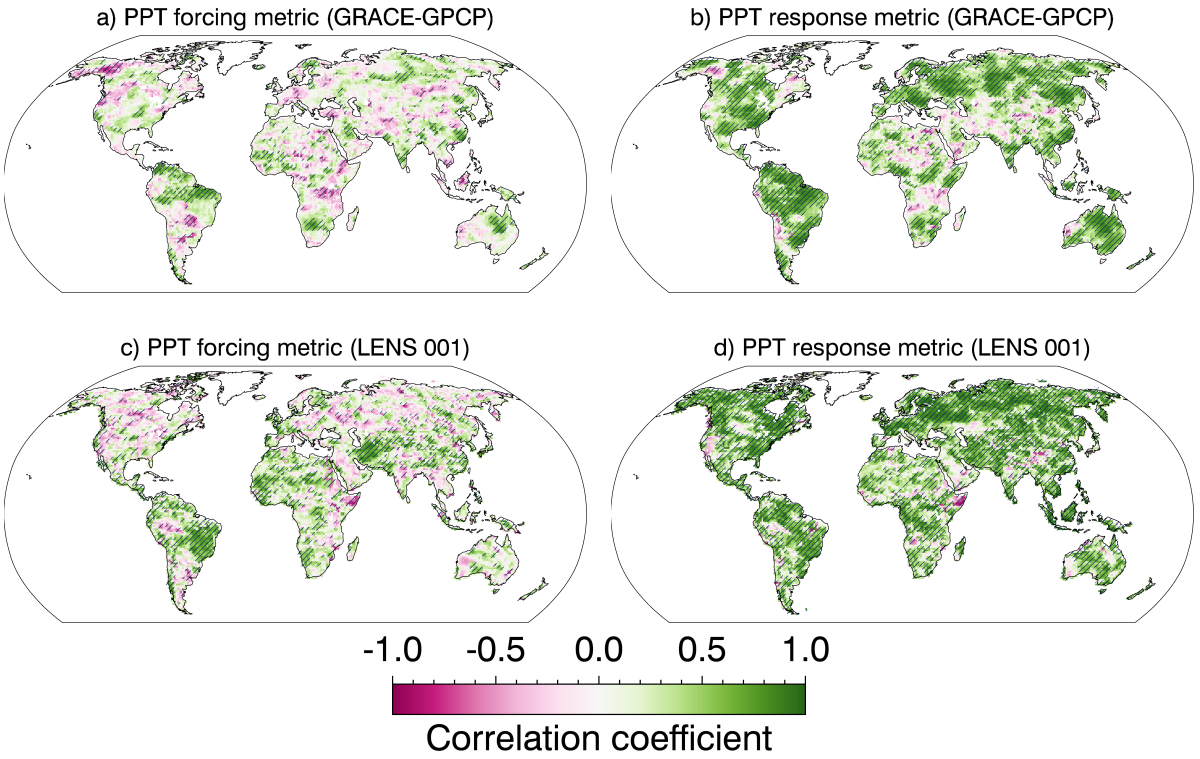

Figure 2.6: Forcing and response metrics for PPT from GRACE/GPCP **(a, b)** and LENS ensemble member 001 **(c, d)**. Cross-hatching indicates a correlation coefficient that is statistically significant at $p \leq 0.05$.

coefficients nearly everywhere, with positive correlations found only in arid regions of low TWS variability. Particularly strong response metrics were found in eastern North America, northern South America, western Eurasia, the Sahel, India, and eastern Australia. The first ensemble member from LENS showed widespread negative correlations, and did not show the positive correlations found in the satellite data. Response coupling in LENS was much more spatially homogeneous than in the satellite data, though northern South America and western Eurasia still showed stronger coupling than elsewhere.

Many of the areas that showed a strong forcing metric for VPD also showed a relatively strong forcing metric for PPT, though the PPT forcing metric was overall weaker than that for VPD (Figure 2.6a). The response metric for PPT was generally positive, indicating that for much of the globe, a more positive TWSA_{min} was associated with higher precipitation rates (Figure 2.6b). Both the forcing

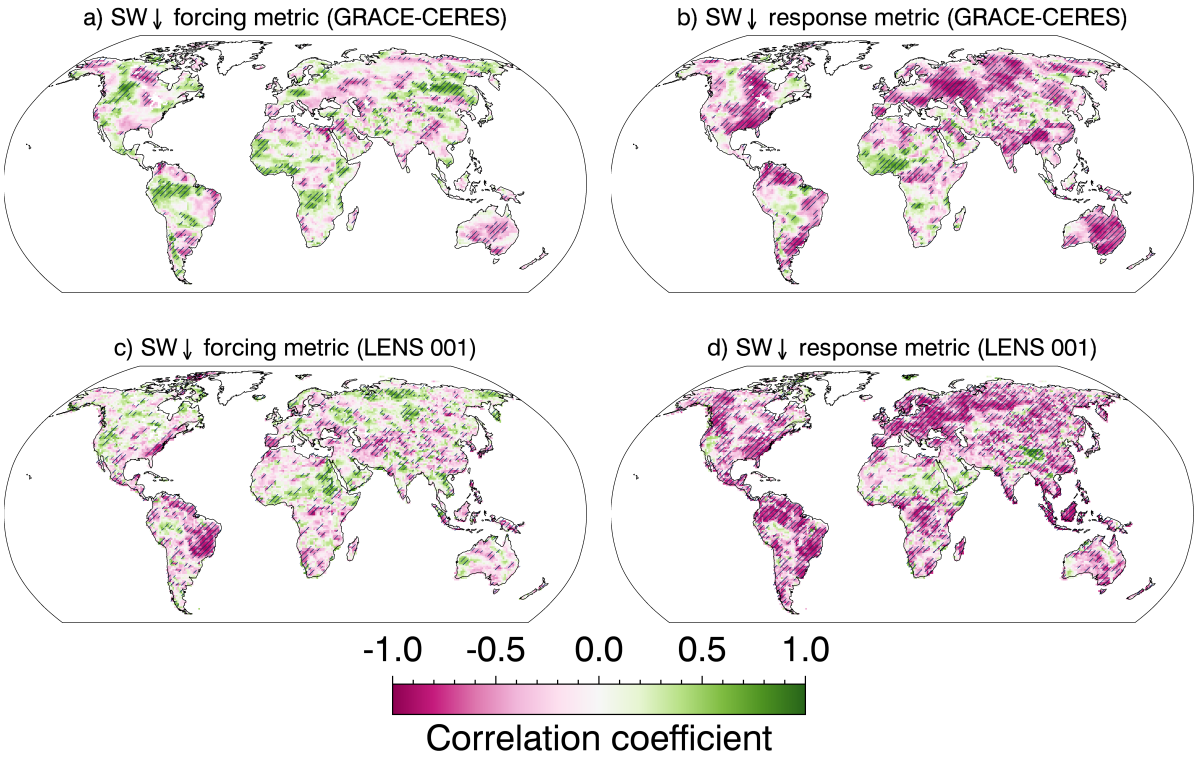

Figure 2.7: Forcing and response metrics for SW↓ from GRACE/CERES **(a, b)** and LENS ensemble member 001 **(c, d)**. Cross-hatching indicates a correlation coefficient that is statistically significant at $p \leq 0.05$.

and response metrics were somewhat stronger in the LENS member relative to those evident in the satellite data (Figure 2.6c and d).

The forcing metrics for SW↓ showed a mixture of positive and negative correlations, indicating that higher TWSA_{max} was either positively or negatively coupled with shortwave radiation (Figure 2.7a). This finding is consistent with both positive and negative coupling between cloud cover and terrestrial moisture observed over shorter timescales [*Taylor et al.*, 2012; *Guillod et al.*, 2015]. The response metrics for SW↓ were generally negative, indicating that greater seasonal shortwave radiation was associated with more negative TWSA_{min} (stronger coupling), with western Africa being a notable exception (Figure 2.7b). The LENS member showed generally stronger coupling in both the forcing and response metrics for SW↓ (Figure 2.7c and d).

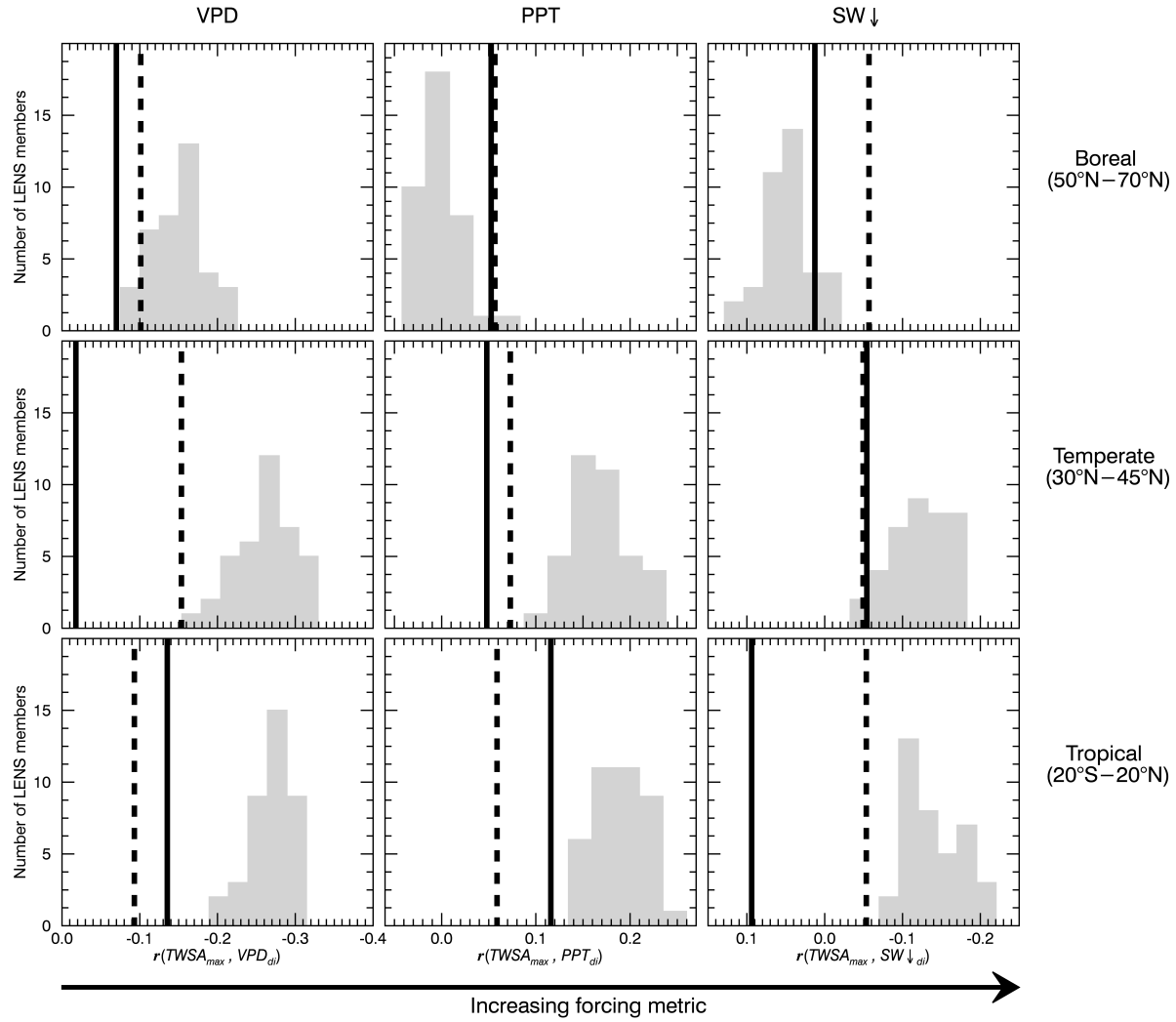

Figure 2.8: Ensemble histogram of forcing metrics from the 38 simulations in LENS (gray bars) compared to satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate set of observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

2.3.3 Evaluating the CESM Large Ensemble

In temperate and tropical regions, forcing metrics were generally stronger in LENS (more positive correlations for PPT, more negative for VPD and SW↓) than in the satellite and reanalysis data, indicating a stronger land surface forcing of the surface atmospheric state in the model than in the observations (Figure 2.8). In boreal regions, forcing metrics were much weaker (closer to zero) than

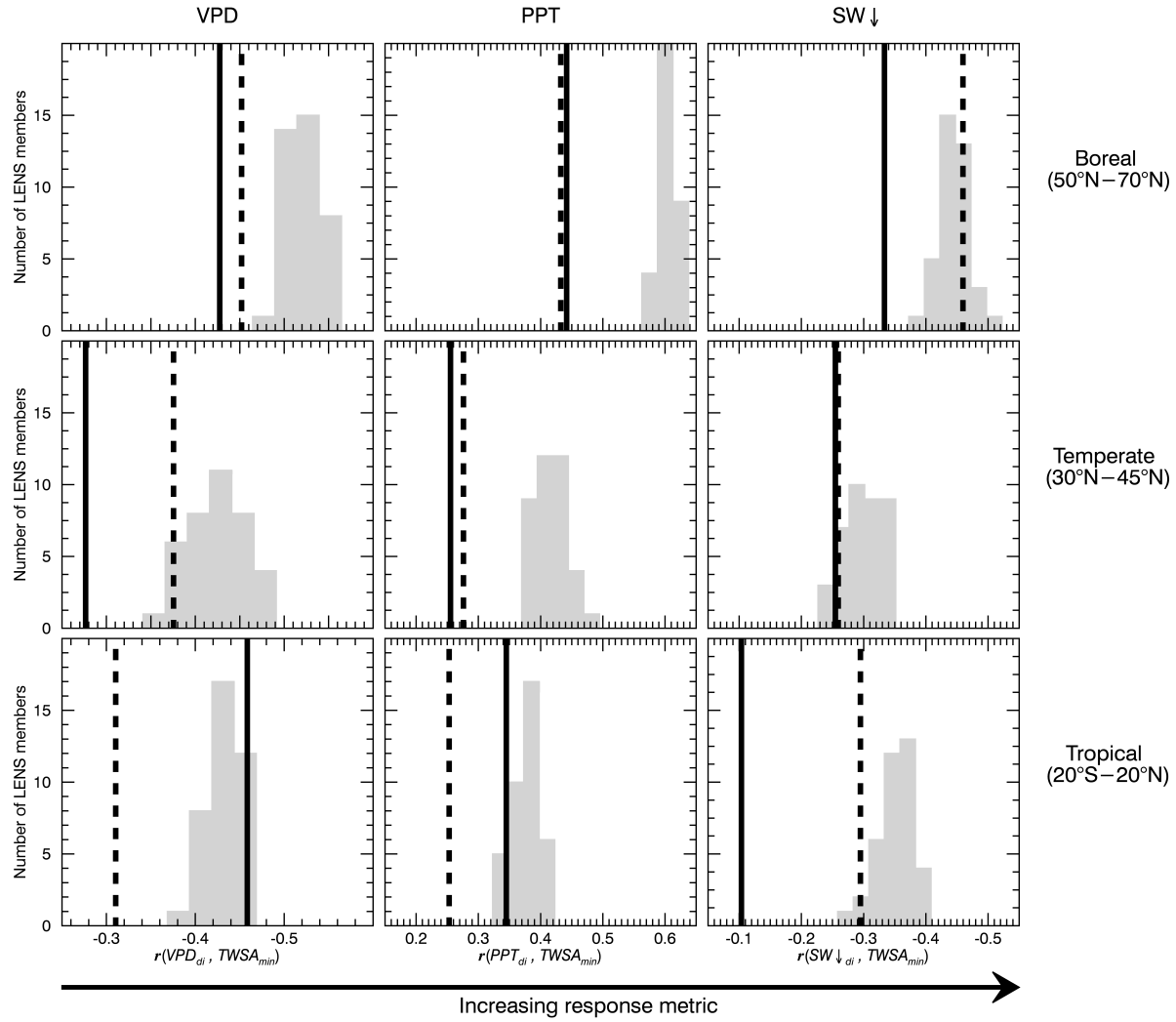

Figure 2.9: Ensemble histogram of response metrics from the 38 simulation in LENS (gray bars) compared to satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate set of observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

at lower latitudes in both the satellite data and in LENS, indicating very little relationship between $TWSA_{max}$ and ATM_{di}. This is consistent with high levels of climate variability in many high-latitude regions driven by the Arctic Oscillation, the North Atlantic Oscillation, and other dynamical modes [*Cohen and Barlow*, 2005]. Furthermore, at high latitudes, ET is generally energy limited rather than moisture limited, which would lead to weak forcing metrics as moisture availability would not strongly influence atmospheric conditions.

Response metrics were also generally higher in LENS than in both the satellite and reanalysis data (Figure 2.9). Noticeable exceptions were the VPD and PPT response metrics in the tropics, which were close to the satellite observations, and the boreal SW↓ and tropical PPT response metrics, which were close to the reanalysis estimates. Despite the internal variability evident within the model ensemble, and the difference between metrics as measured by the satellite data compared with the reanalysis data, the general pattern indicated that modeled response metrics were higher than those from observations and reanalysis.

2.3.4 Analysis of uncertainty

The internal variability across the ensemble of simulations in LENS yielded a distribution of forcing and response metrics with a spread on the same order of magnitude as the difference between modeled and satellite-derived zonal averages. The distribution of coupling metrics from LENS revealed the sensitivity of the relationships to decadal climate variability given the relatively short TWS time series. Comparing this distribution with the spread between the purely satellite-derived metrics and GRACE-reanalysis hybrid indicated the sensitivity of our metrics to the choice of data source. The differences between satellite and reanalysis metrics were generally greater in the tropics, particularly for VPD and SW↓, and in midlatitude VPD for both forcing and response variables. Elsewhere, the differences were generally similar to or less than the differences between the observationally constrained zonal averages and the LENS distributions.

Comparing the original LENS forcing and response metrics with those calculated after adding random noise to LENS (Figures 2.10 and 2.11) provided an estimate of the metrics' sensitivity to observational uncertainty. Adding random noise with 25% of the standard deviation of the original data to the model time series of TWSA and atmospheric variables at each grid cell does degrade the metrics slightly, causing areal averages to be closer to zero, but the differences are relatively small compared to the differences between observed and modeled averages as well as the spread of the

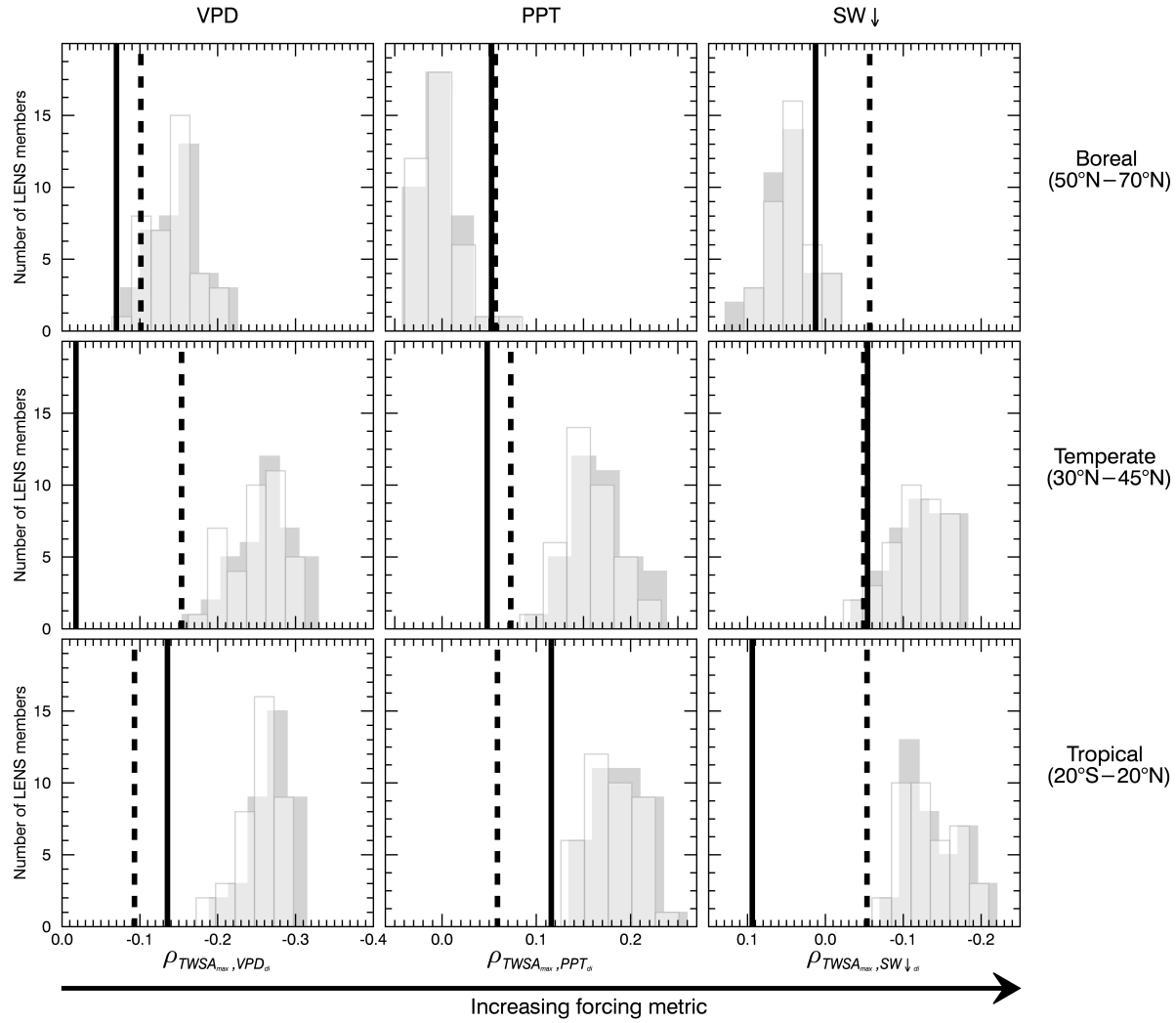

Figure 2.10: Ensemble histogram of forcing metrics from LENS (grey bars) and LENS plus random noise (white bars), with the satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

ensemble itself. This sensitivity analysis provided evidence that observational errors likely have a relatively small impact on the quality of our satellite-derived metrics.

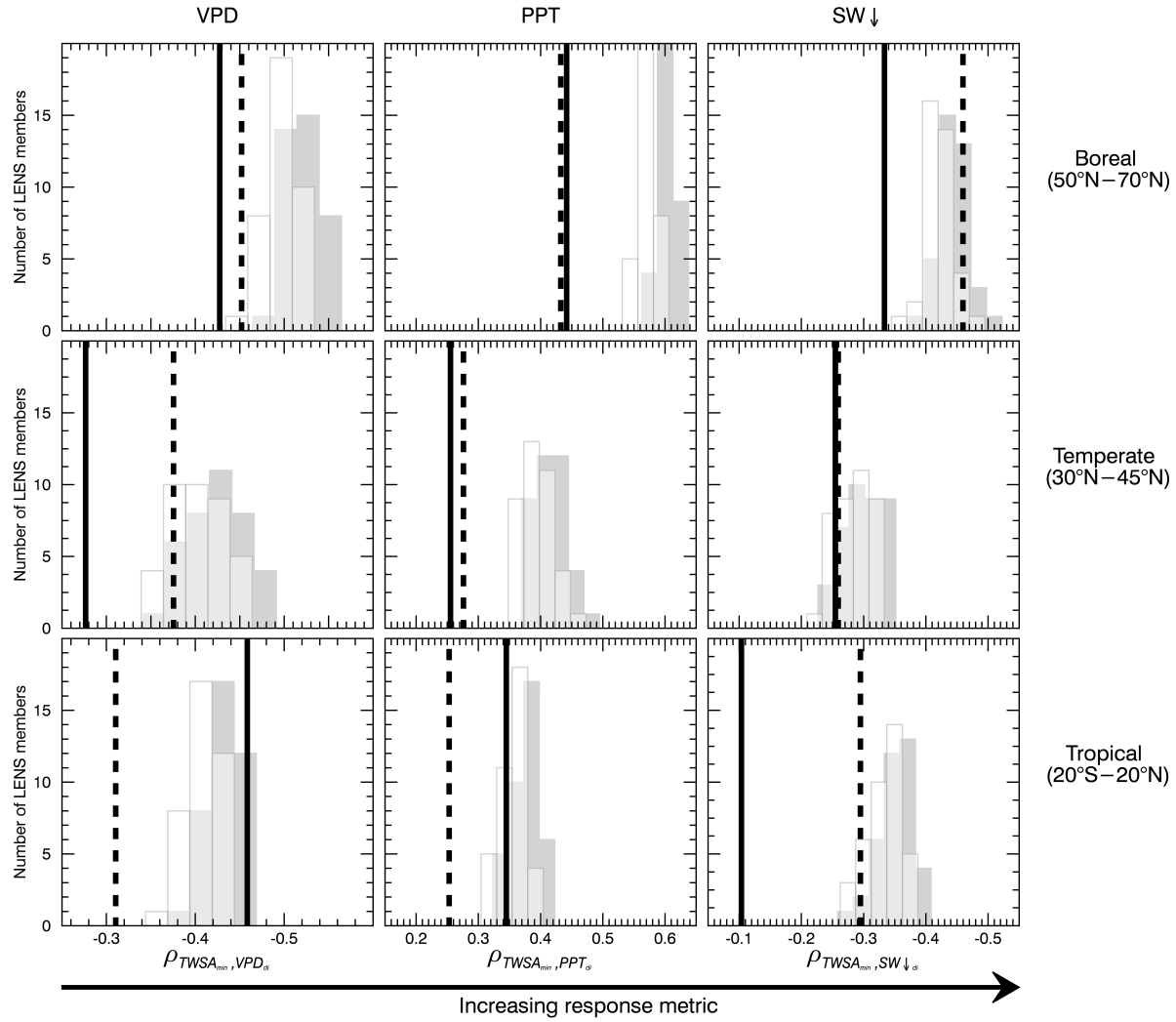

Figure 2.11: Ensemble histogram of response metrics from LENS (grey bars) and LENS plus random noise (white bars), with the satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

2.3.5 Evaluating CMIP5 models

Comparison of the explicit TWS field from LENS with the accumulated residuals of the surface water budget, as well as the forcing and response metrics calculated using both (Figures 2.12 and 2.13), indicated that the alternative formulation provides an acceptable substitute when an explicit TWS field is not available from an ESM. More specifically, it suggests that water storage

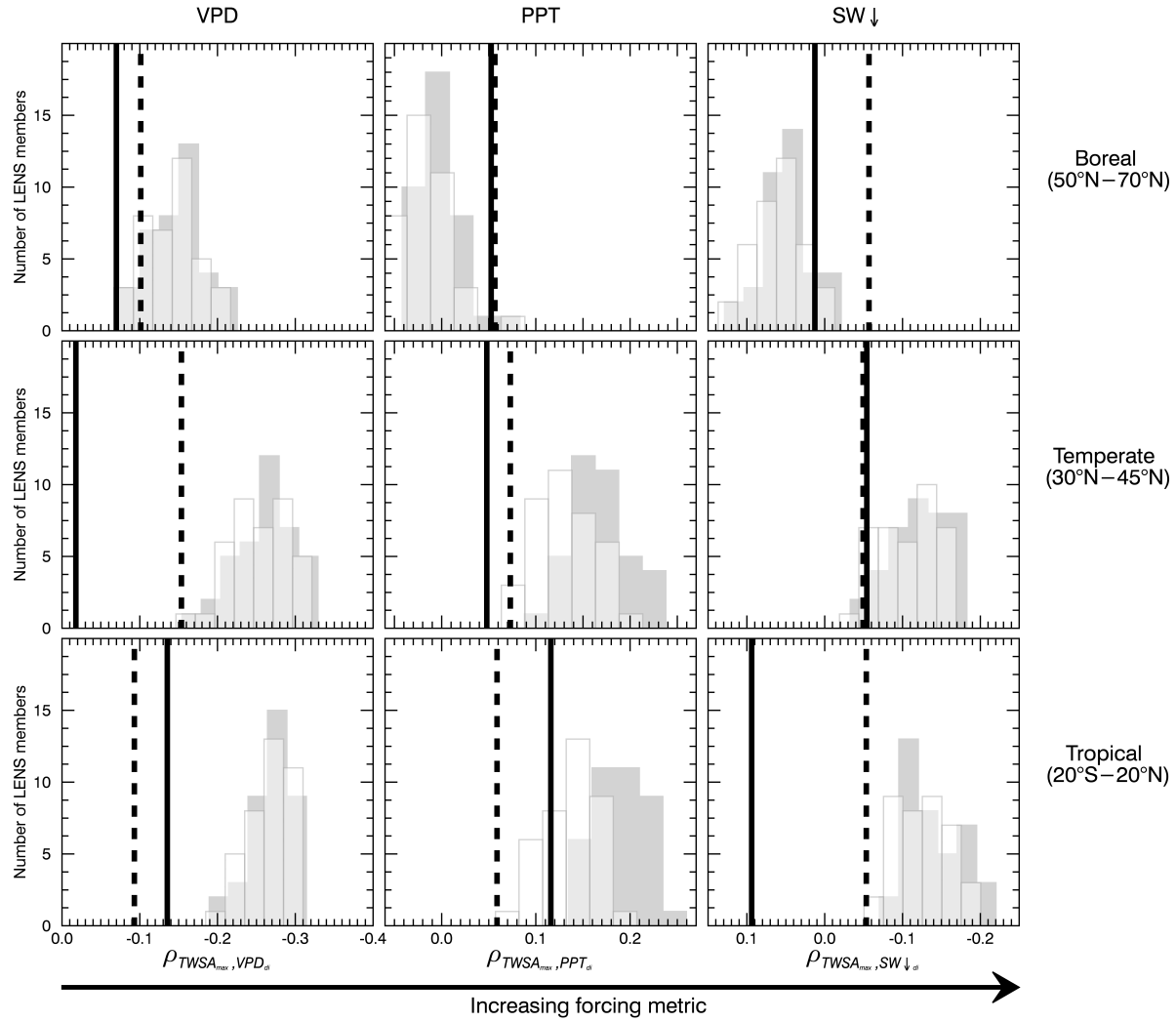

Figure 2.12: Ensemble histogram of forcing metrics from LENS with explicit TWS (grey bars) and accumulated residuals of the water balance (white bars) with the satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

in rivers, lakes, and other parts of the terrestrial hydrologic system that are downstream from grid cell-level runoff did not significantly degrade the set of metrics evaluated here. Some degradation of the forcing metrics for PPT was apparent in the middle and low latitudes, but the remaining metrics are not highly sensitive to TWS formulation. This suggests that metrics calculated for CMIP5 output using accumulated residuals could be reasonably and effectively compared with the metrics derived from LENS and the observations (Figure 2.14).

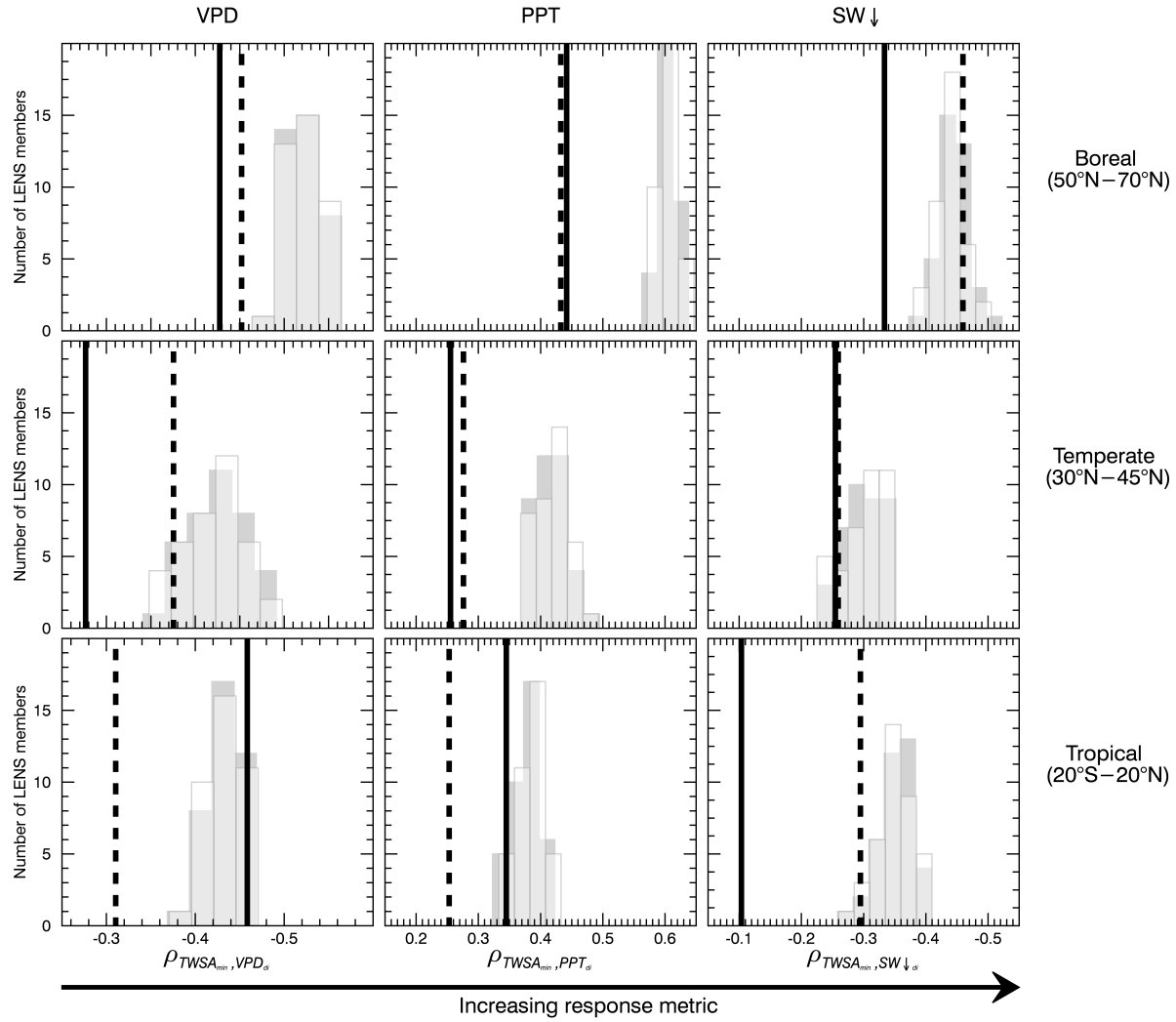

Figure 2.13: Ensemble histogram of response metrics from LENS with explicit TWS (grey bars) and accumulated residuals of the water balance (white bars) with the satellite observations from GRACE/AIRS/GPCP/CERES (solid black line) and the alternate observations from GRACE and ERA-Interim (dashed black line), averaged across land regions within different latitude bands.

As with LENS, the metrics derived from CMIP5 output indicated generally stronger coupling metrics than the observations for both the forcing and response limbs. Exceptions include the VPD response metric in the tropics, the boreal PPT and SW↓ forcing metrics, and the midlatitude SW↓ response metrics. The spread between various models was generally greater than the spread within any single model with a multi-member ensemble. Of the four models that use CLM4 for the land surface, the two that use CAM5 for the atmosphere (LENS and CESM1-CAM5) were clustered

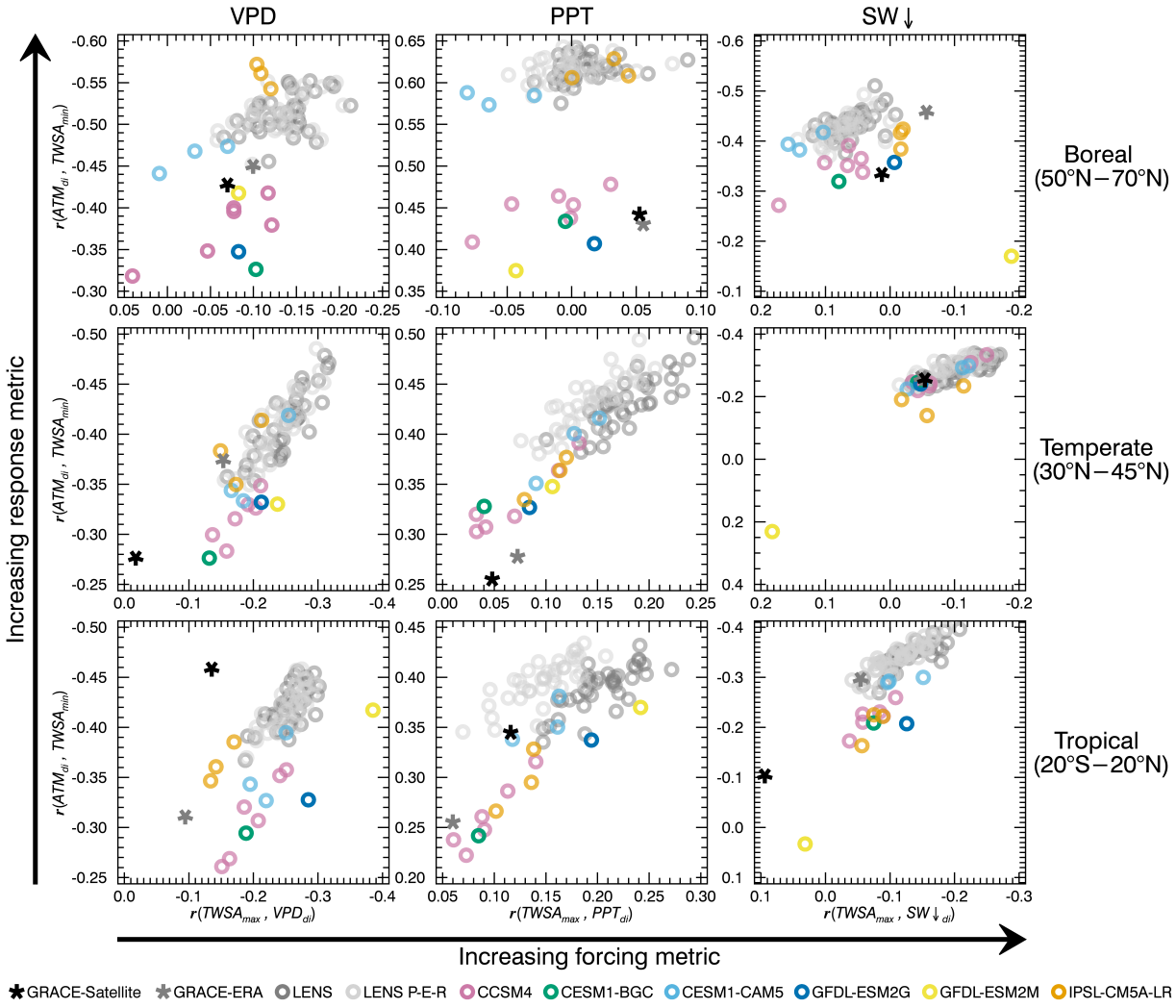

Figure 2.14: Scatter plots of forcing and response metrics for LENS and CMIP5 models with observations, averaged across land regions within different latitude bands. For LENS, we show metrics calculated using the explicit TWS output (darker gray) and TWSA estimates from the accumulated residuals of the surface water balance (lighter gray). For CMIP5 models, we calculated metrics using TWSA estimates from the accumulated residuals of the surface water balance.

close together, and exhibited generally the strongest forcing and response metrics. The two that use CAM4 (CCSM4 and CESM1-BGC) were close to each other, but with lower metrics in both forcing and response than the CAM5 models. The two GFDL models were both within the general ensemble range in the metrics for both VPD and PPT, but GFDL-ESM2M was an extreme outlier in both forcing and response metrics for SW↓.

Comparison of CMIP5 and LENS models indicated a mostly positive relationship between forcing and response metrics in temperate and tropical latitude bands. In boreal latitudes, there was little distinction between the forcing metrics of the different models, all of which were close to zero, though there were some clear differences within the response metrics. In temperate and tropical latitudes, models that showed the strongest forcing metrics generally also showed the strongest response metrics for a given variable. This relationship suggests that analysis of the response limb of the feedback loop is important for understanding how conditions are set up for subsequent forcing via land–atmosphere coupling.

2.4 Discussion

2.4.1 Benchmarking models with observed coupling metrics

The metrics developed here from satellite observations provide a means for evaluating land–atmosphere feedback strength on seasonal to interannual timescales in coupled ESMs. The use of correlation coefficients in this study does not enable a direct assessment of whether the relationships are directly causal, as correlation between atmospheric and terrestrial conditions could result from atmospheric persistence and remote forcing from SSTs [*Orlowsky and Seneviratne*, 2010; *Mei and Wang*, 2011]. Nevertheless, the satellite-derived metrics provide a meaningful constraint against which coupled models can be benchmarked, as these models need to correctly represent the combined effects of persistence, remote SST forcing, and land–atmosphere coupling.

The forcing metrics, by indicating the relationship between antecedent TWS and subsequent atmospheric characteristics, provide observational constraints to complement previous research in large-scale land–atmosphere coupling in global models [e.g., *Koster et al.*, 2006; *Guo and Dirmeyer*, 2013; *Seneviratne et al.*, 2013]. Observed forcing metrics were found to be strong in some of the regions of intermediate wetness in which ET is limited by terrestrial moisture availability, in addition

to some regions in the moist tropics in which ET is generally considered to be energy limited. Recent observational analyses by *Hilker et al.* [2014] demonstrated that at least in the Amazon, deep rooting-zone water supplies can become seasonally depleted, leading to a stronger land–atmosphere coupling. This is consistent with findings that deep rooted plants vertically redistribute soil water to shallower layers, allowing higher levels of evapotranspiration to be sustained during the dry season [*Lee et al.*, 2005] . It is also consistent with recent work demonstrating that TWSAs can be used as predictors for fire season severity in the Amazon [*Chen et al.*, 2013].

The inclusion of response metrics in our analysis allows the full feedback loop to be considered by recognizing the two-way dependence between the land surface and the atmosphere. The generally higher correlation coefficients in observed response metrics indicates the importance of the land surface response in priming the system for subsequent forcing on the atmosphere. For example, if the TWS response is too strongly coupled to the atmosphere, a small change in atmospheric conditions could yield an unrealistically large change in TWS. The unrealistically large TWS anomaly, in turn, would have the potential to impart a larger land surface forcing of the atmosphere in subsequent time steps. That models and ensemble members with high forcing metrics were also generally found to have high response metrics (Figure 2.14) highlights the need to consider this.

Both the forcing and response metrics as calculated from the output of the ESMs analyzed in the current study indicated generally stronger coupling compared with those derived from the satellite observations. There are exceptions to this pattern, but it holds generally true, particularly across middle and lower latitudes, and particularly in the LENS data. This is consistent with previous studies conducted at finer temporal resolutions [*Ferguson et al.*, 2012] and across more limited spatial domains [*Hirschi et al.*, 2011]. As described below, there are several possible explanations as to why models may simulate a stronger feedback than is observed in the satellite record.

2.4.2 Possible explanations for enhanced feedback strength in models

One set of possible explanations for the stronger coupling metrics in models relative to observations involves models overestimating the amount of water available for ET during the drawdown interval. The land surface influence on the atmosphere requires water to be a limiting factor to ET but not limiting enough to prevent it altogether. Under more moisture-limited conditions, a drawdown interval may experience multiple shorter time periods during which ET is inhibited due to insufficient water, and the terrestrial moisture state exerts no control over flux partitioning. These periods of insufficient moisture would tend to reduce the overall feedback strength integrated across the duration of the drawdown interval. Model shortcomings that make water too readily available for ET could reduce the amount of time spent in a periods of insufficient moisture during the drawdown interval, thereby unrealistically strengthening the longer-term feedback. We note that the opposite could take place under near-saturated conditions if a model overestimates the amount of time in which ET is energy limited, but we would not expect these conditions to be as prevalent during the drawdown interval that was the time period of focus in our analysis.

ESMs are known to simulate unrealistically homogeneous rainfall intensity, with overestimates of drizzle and underestimates of large infrequent events [*Dai*, 2006]. Infrequent high-intensity rainfall events would yield much more runoff from saturated soil, which would lead to a weaker connection between the land and atmosphere than frequent low-intensity drizzle. If a model simulates too much drizzle, precipitation could lead to too much storage, which would cause a model to overestimate the response metrics. Too much storage also could allow water to be too readily available for ET, causing an overestimate of the forcing metrics. Contributions from drizzle could be offset if insufficient rainfall intensity does not allow for high enough throughfall or soil moisture recharge. The issue of rainfall intensity is related to issues of convective parameterization (described below), and may be addressed in future versions of ESMs through atmospheric superparameterization, in which a model's convective parameterization is replaced with embedded cloud-resolving models [*Kooperman et al.*, 2016].

A misrepresentation of either the amount of bare soil or of bare soil processes also could lead to overestimates of the amount of water available for ET and thereby coupling strength. Current land surface schemes of ESMs are based on the "big leaf" model paradigm, which could lead to overestimates of ET if runoff and groundwater recharge are underestimated as a consequence of an unrealistically small bare soil fraction. In addition, even if bare soil fraction were correct, overestimates of ET due to an incomplete representation of surface resistance of bare soil, as found in CLM4 by *Swenson and Lawrence* [2014], would amplify positive feedbacks.

Additional explanations for why models may overestimate feedback strength include the parameterization of convection in the PBL or stomatal conductance responses to soil moisture. Previous work using a regional climate model (RCM) with a higher spatial resolution have determined that convective parameterizations are as important as spatial resolution in the simulation of precipitation coupling [*Hohenegger et al.*, 2009]. *Taylor et al.* [2013] similarly found parameterized convection in an RCM yielding a positive coupling in contrast to the negative coupling found in both observations and model runs with explicitly simulated convection. If negative coupling mechanisms are present in reality but absent from models, this could contribute to an overestimate of coupling metrics and underrepresentation of negative feedbacks in models. Similarly, the diversity of stomatal conductance parameterizations in CMIP5 ESMs is relatively low [*Medlyn et al.*, 2011; *Swann et al.*, 2016], and if stomatal apertures close too rapidly in response to an initial deficit in terrestrial water storage, transpiration–humidity feedbacks may be intensified in an unrealistic manner.

One factor that could contribute toward stronger coupling metrics in models relative to observations is the effect of observational uncertainty combined with a relatively short time series. Adding random error to one or more variables in a correlation analysis will reduce the correlation coefficient, and this degradation has been shown to be sensitive to the length of data sets used to establish metrics of land–atmosphere coupling [*Findell et al.*, 2015]. Given the relatively short time series available for the current analysis, the correlation coefficients derived from remote sensing data may be reduced due to observational uncertainty, unlike those derived from internally consistent models.

We found that adding random noise to LENS at 25% of the standard deviation of the original data caused some degradation of our area-averaged coupling metrics, but only by a small amount relative to the difference between LENS and the observations (Figures 2.10 and 2.11). We chose 25% as a qualitative upper bound on likely uncertainties introduced from random observational error within the TWSA and atmospheric variable time series. This highlights the need for developing more quantitative error estimates in remote sensing and reanalysis products. More generally, this sensitivity analysis suggests that our coupling metrics, when averaged across large areas, may be useful in identifying robust data–model differences.

Another possible explanation stems from the fact that our coupling metrics include co-variability due to atmospheric persistence and remote forcing by SST [*Orlowsky and Seneviratne*, 2010; *Mei and Wang*, 2011] alongside the direct influence of land–atmosphere interactions. For this reason, we caution that overestimates of coupling metrics do not imply that the land–atmosphere feedback is necessarily stronger, but could be due to an overestimate of SST-driven correlations between the land surface and the atmosphere. *Wei et al.* [2008] demonstrated that negative correlations between soil moisture and subsequent precipitation can be explained by precipitation persistence combined with negative temporal autocorrelation of precipitation associated with intra-seasonal modes such as the Madden–Julian Oscillation (MJO). Poor representation of the MJO period in CMIP5 models leads to unrealistic patterns of precipitation persistence [*Hung et al.*, 2013]. If models are failing to capture MJO-driven negative correlations, this could lead to overly strong positive correlations relative to observations. However, this would depend on the length of the drawdown interval relative to persistence time and the period of intra-seasonal modes.

2.4.3 Uncertainties and future applications

The current study demonstrates the utility of the coupling metrics presented here, but conclusions are limited by the time span of the satellite record. While LENS enables the internal variability of these

relationships to be investigated within the model, it is unclear how much natural climate variability affects these relationships in reality on timescales longer than the satellite record. Furthermore, we acknowledge that observational error over an insufficiently long time series could reduce the apparent strength of correlations [*Findell et al.*, 2015]. Therefore, the utility of the coupling metrics we present will increase alongside the length of the time series available from remote sensing platforms. This emphasizes the importance of the GRACE follow-on mission [*Flechtner et al.*, 2014] and the need for continuity in the record between missions.

Furthermore, incorporating additional remote sensing products can reduce uncertainties inherent in the satellite-derived data sets. We presented metrics derived using ERA-Interim in place of AIRS, GPCP, and CERES in order to qualitatively illustrate this uncertainty. We found a non-negligible amount of uncertainty in both forcing and response metrics due to inconsistencies between the remote sensing and reanalysis products. Future work will address these uncertainties by incorporating additional observations and observationally constrained data sets such as those from the Global Soil Wetness Project [*Dirmeyer et al.*, 2006b] and the Global Land Data Assimilation System [*Rodell et al.*, 2004]. In addition, as increasingly long time series of data become available from the Soil Moisture Ocean Salinity [*Mecklenburg et al.*, 2012] and Soil Moisture Active Passive [*Panciera et al.*, 2014] missions, the metrics developed here can be applied to those data sets as well, which will elucidate the importance of surface soil moisture relative to the total TWS column in these interactions.

Finally, the issue of causality and the possibility that correlations result primarily from atmospheric persistence and remote forcing from SST rather than land–atmosphere interactions may be addressed using sensitivity experiments similar to those of the GLACE and GLACE-CMIP experiments. While the previous experiments have tested the importance of soil moisture interaction with the atmosphere, additional experiments could expand upon these methods by treating SST variability similar to terrestrial soil moisture availability. Such experiments could determine the relative importance of remote SST forcing, including the effect of atmospheric persistence, and local land–atmosphere

coupling in explaining correlations between TWS and atmospheric conditions.

As these sources of uncertainty are diminished, the coupling metrics introduced here may be used to assess whether improvements to model biogeophysics and parameterizations yield relationships that are more consistent with observations. CMIP5 models are known to have a high ET bias [*Mueller and Seneviratne*, 2014], which could be due in part to the explanations proposed as possible reasons for overestimated coupling metrics in models. As data become available from phase 6 of the Coupled Model Intercomparison Project (CMIP6), these metrics could provide an assessment of whether improvements to ET processes in models also improves the relationship between the land surface and the atmosphere.

2.5 Conclusion

We have developed a new approach for measuring the strength of the two-way feedback relationships between TWS and the atmosphere. This approach was designed specifically to take advantage of TWSA data from the GRACE mission, along with concurrently collected remote sensing and reanalysis data sets of atmospheric variables, in a manner that could then be applied to Earth system models. The coupling metrics described here quantify the relationships between both antecedent TWS and subsequent atmospheric conditions, as well as antecedent atmospheric conditions and subsequent TWS.

Regions of strong forcing, in which the TWSA at the beginning of the drawdown interval was related to the subsequent atmospheric state, coincided with the semi-arid zones previously found to be hot spots of land–atmosphere coupling, as well as some new tropical zones that may have moisture-limited ET regimes. Regions of strong response metrics, in which the TWSA at the end of the drawdown interval is related to the atmosphere, are much more widespread. Modeled coupling metrics are generally found to be stronger than those observed in the satellite record. If

this discrepancy is due to models overestimating the two-way feedback between the land surface and the atmosphere, this could lead to models incorrectly projecting future warming trends and climatic extremes [e.g., *Hirschi et al.*, 2011; *Seneviratne et al.*, 2013; *Cheruy et al.*, 2014; *Miralles et al.*, 2014].

The results of this study are consistent with previous studies at smaller temporal scales indicating land–atmosphere coupling strength may be stronger in models than in observations. There are several possible mechanisms that may contribute to the overestimation of land–atmosphere coupling in models, and future studies may incorporate the metrics introduced here to assess the role of these mechanisms. These metrics will become increasingly useful as the temporal coverage of the remote sensing record grows longer and additional missions come online.

Acknowledgements

This chapter was adapted from the following publication:

Levine, P. A., J. T. Randerson, D. M. Lawrence, and S. C. Swenson (2016), Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations, *Hydrology and Earth System Sciences*, *20*, 4837–4856. http://dx.doi.org/10.5194/hess-20-4837-2016

 We received funding support from the United States Department of Energy Office of Science Biogeochemistry Feedbacks Scientific Focus Area and the Climate Change Prediction Program, cooperative agreement (DE-FC03-97ER62402/A010). The CESM Large Ensemble Community Project was supported by the National Science Foundation (NSF) with supercomputing resources provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory (CISL). We also acknowledge the World Climate Research

Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 2 of this paper) for producing and making available their model output. For CMIP, the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led the development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Chapter 3

Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to El Niño–Southern Oscillation

3.1 Introduction

3.1.1 El Niño-Southern Oscillation and Amazon climate variability

El Niño-Southern Oscillation (ENSO) is the dominant mode of interannual climate variability in the Earth's tropics. During the positive phase of the ENSO cycle (El Niño), high sea surface temperature (SST) anomalies in the eastern and central equatorial Pacific persist for several months during boreal fall and winter. Positive SST anomalies are associated with a weakened Walker circulation, which leads to climatic teleconnections globally, and particularly over the tropical land surface

[*Trenberth et al.*, 2002]. Under El Niño conditions, the tropical land surface as a whole experiences anomalously high surface air temperatures, and anomalously low precipitation and terrestrial water storage, while the opposite occurs during the negative phase of the ENSO cycle (La Niña) [*Llovel et al.*, 2011; *Tyrrell et al.*, 2014; *Reager et al.*, 2016].

ENSO-driven land surface responses are particularly prevalent in the Amazon River Basin [*Marengo*, 1992; *Foley et al.*, 2002; *Chen et al.*, 2011; *de Linage et al.*, 2013]. Recent El Niño events have exacerbated the effects of global warming, with the 2015–2016 El Niño leading to record-breaking temperatures and extreme drought in the Amazon [*Jiménez-Muñoz et al.*, 2016]. The Amazon is known as a "hot spot" of land-atmosphere coupling, in which variation in soil moisture leads to variation in the partitioning of turbulent surface fluxes between latent and sensible heat, subsequently influencing development of the planetary boundary layer and local atmospheric conditions [*Fu et al.*, 2001; *Lee et al.*, 2011; *Ma et al.*, 2011; *Sun and Wang*, 2013]. An outstanding challenge is to determine how much the Amazonian temperature anomalies during different phases of the ENSO cycle are the direct result of atmospheric circulation changes due to SST variability, and how much they are indirectly driven by the local land surface response to moisture redistribution [*Sun and Wang*, 2013; *Spennemann and Saulo*, 2015; *Levine et al.*, 2016].

The ENSO cycle is also associated with interannual variability in the atmospheric CO_2 growth rate (CGR), which increases during El Niño and decreases during La Niña [*Bacastow*, 1976; *Keeling and Revelle*, 1985; *Jones et al.*, 2001]. While interannual variability in the carbon sink of the tropical land surface has been strongly implicated in CGR variability, there is ongoing debate in the literature as to the exact mechanisms that lead to this relationship. Some studies attribute CGR variability primarily to the effects of temperature on terrestrial ecosystems [*Cox et al.*, 2013; *Piao et al.*, 2013; *Wang et al.*, 2013, 2014]. Other studies indicate that CGR is driven at least in part by the terrestrial response to precipitation and water storage variability [*Foley et al.*, 2002; *Zeng et al.*, 2005; *Qian et al.*, 2008; *Keppel-Aleks et al.*, 2014; *Wang et al.*, 2016a; *Humphrey et al.*, 2018]. While the relative importance of these drivers varies across tropical continents, both high temperature and

low precipitation in the Amazon during a recent El Niño event were associated with an observed positive CO_2 flux anomaly [*Liu et al.*, 2017].

3.1.2 Land-atmosphere coupling in the Amazon's response to ENSO

Of course, the effects of temperature and hydrology on the terrestrial carbon cycle should not be viewed as mutually exclusive. The role of land-atmosphere coupling in the Amazon suggests that when considering the impact of SST forcing on terrestrial ecosystems, temperature anomalies and moisture redistribution may not be independent. However, studies exploring the relative importance of precipitation and temperature controls on terrestrial carbon cycle variability may obscure this interdependence by externally forcing land surface models with time series of these variables from observations or reanalysis [*Zeng et al.*, 2005; *Qian et al.*, 2008; *Piao et al.*, 2013; *Wang et al.*, 2013, 2016a].

The Amazon's ecological response to ENSO is delayed by a series of cascading effects. ENSO teleconnections with the Amazon are strongest during boreal winter, with El Niño events decreasing precipitation between November and April [*Chen et al.*, 2017]. These months coincide with the dominant wet season for most of the Amazon. Therefore, even though years that begin under El Niño conditions may receive anomalously low precipitation during these months, the terrestrial ecosystem may not respond immediately, as water availability is not a limiting factor during the wet season. The Amazon rainforest is known to sustain plant growth during the dry season by utilizing water from deep soil, which can be hydraulically redistributed to the surface by deep roots [*Nepstad et al.*, 1994; *Lee et al.*, 2005], and transferred from groundwater into the root zone through capillary rise [*Miguez-Macho and Fan*, 2012]. Therefore, wet season precipitation deficits and concurrent radiation surpluses may reduce the soil moisture store for the subsequent dry season, leading to decreased evapotranspiration and primary production [*Juárez et al.*, 2007; *Chen et al.*, 2013; *Hilker et al.*, 2014; *Bowman et al.*, 2017; *Liu et al.*, 2017; *Swann and Koven*, 2017].

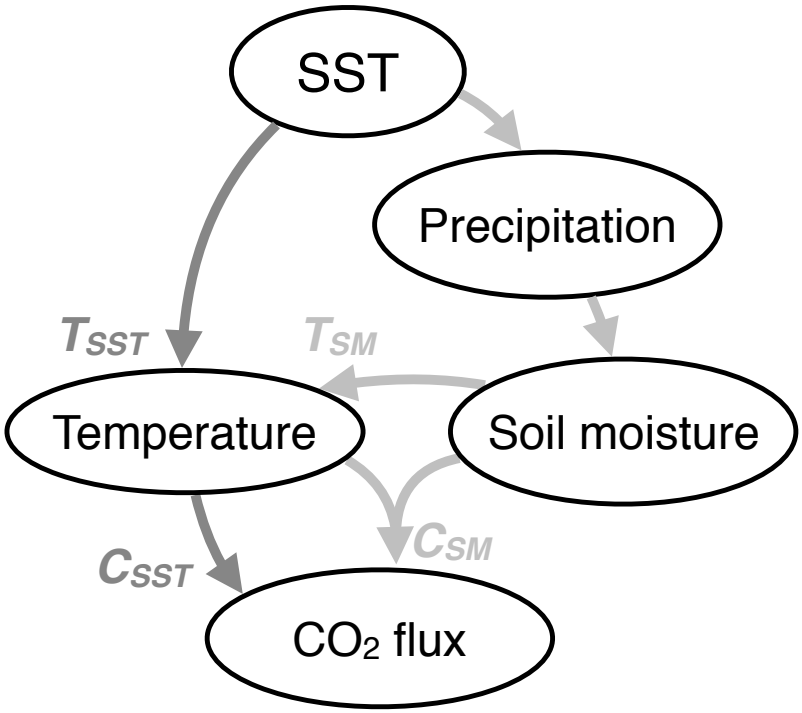

Figure 3.1: Conceptual diagram of the hypothesized pathways through which SST and soil moisture variability impact temperature and the carbon cycle. T_{SST} is the portion of surface air temperature anomalies driven directly by remote SST forcing. T_{SM} is the portion of temperature anomalies driven indirectly by SST-driven soil moisture anomalies via land-atmosphere coupling. C_{SST} is the portion of the carbon cycle anomalies driven by T_{SST}. C_{SM} is the portion of the carbon cycle anomalies driven directly by soil moisture anomalies along with the portion of the temperature signal driven by soil moisture anomalies (T_{SM}).

Here, we were motivated by the hypothesis that some portion of the Amazon temperature anomalies associated with ENSO was due to land-atmosphere coupling resulting from the cascading effects of precipitation and subsequent soil moisture anomalies. We posited that following El Niño events, drier soils in the Amazon would inhibit evaporative cooling, indirectly contributing to warmer temperatures above and beyond direct SST forcing. Furthermore, we hypothesized that the temperature anomalies driven by land-atmosphere coupling may have contributed to variability in CO_2 fluxes between the terrestrial ecosystem and the atmosphere. This hypothesis is illustrated conceptually in Figure 3.1.

To test this hypothesis, we performed a series of mechanism-denial experiments in an Earth system model. As described in Section 3.2.3.2.2, we first ran a control simulation driven by observed SSTs

and with fully interactive land-atmosphere feedbacks. We then modified the land surface model to assimilate soil moisture in order to isolate the atmospherically mediated effects of remote SST variability on land surface temperature from those occurring indirectly as a result of SST-induced soil moisture anomalies. Our goal was to determine how much the cascading effects of SST anomalies on soil moisture anomalies contribute to interannual variability in surface air temperature and the net CO_2 flux over the Amazon basin.

3.2 Methods

3.2.1 Model description

We used version 0.3 of the Energy Exascale Earth System Model (E3SM), a branch of the Community Earth System Model (CESM) version 1.3 beta that is now under development by the United States Department of Energy Office of Science [*Terai et al.*, 2017]. Global E3SM simulations were run in a configuration similar to the Atmospheric Model Intercomparison Project (AMIP) [*Gates et al.*, 1999], with an interactive atmosphere and land surface, and with prescribed SSTs and sea ice fractions. The atmosphere was simulated by version 5.3 of the Community Atmosphere Model [*Neale et al.*, 2012] with the spectral element dynamical core (CAM-SE) [*Dennis et al.*, 2012], with the land surface simulated by version 4.5 of the Community Land Model with prognostic biogeochemistry (CLM4.5-BGC) [*Oleson et al.*, 2013].

The prescribed SST and sea ice time series were from version 2 of the NOAA 1/4° daily Optimum Interpolation Sea Surface Temperature (OISSTv2) [*Banzon et al.*, 2016]. We chose this observationally-based data set in order to simulate the climatic response to the ENSO time series that took place during the observational record. Radiative forcing of greenhouse gas and aerosols was fixed at levels from the year 2000, so that the forced component of interannual variability would be due entirely to SST anomalies. We ran E3SM simulations on a cubed-sphere grid of NE30

(~ 100-km grid spacing), with 30 vertical layers in the atmosphere and 10 hydrologically active soil layers with exponentially increasing thicknesses to a depth of ~3.8 m. History fields were conservatively regridded to a 1° uniform equirectangular grid.

3.2.2 Experimental setup

The control run (hereafter referred to as the AMIP simulation) was forced with SSTs and sea ice fractions from daily OISSTv2 data from 1982 through 2016. Soil moisture was simulated interactively by CLM4.5-BGC, and recorded in each layer at every time step. We used this interannually varying time series of soil moisture from the AMIP simulation, as well as its climatology, to prescribe soil moisture in our subsequent experimental simulations (described below, and summarized in Table 3.1).

Table 3.1: Description of E3SM experiments and applicability to the conceptual pathways from Figure 3.1.

Simulation	SST source	Soil moisture	Pathways
AMIP	Time varying observations	Interactive from CLM4.5	All
SST_{var}	Time varying observations	Prescribed from AMIP climatology	T_{SST} & C_{SST}
SM_{var}	Observations climatology	Prescribed from AMIP (time varying)	T_{SM} & C_{SM}
NO_{var}	Observations climatology	Prescribed from AMIP climatology	None

In the first experimental simulation (SST_{var}), we maintained interannual variability in the SST forcing, but we prescribed soil moisture in each layer to the annual climatology of the data from the AMIP simulation. This served to isolate the direct influence of SSTs on interannual variability in the Amazonian temperature (T_{SST} in Figure 3.1) and carbon cycle (C_{SST}), while excluding the influence of soil moisture (T_{SM} and C_{SM}). Our approach was similar to the Global Land-Atmosphere Coupling Experiment of the Coupled Model Intercomparison Project, Phase 5 (GLACE-CMIP5) [*Seneviratne et al.*, 2013], but with prescribed SSTs that capture the observed ENSO time series.

In an additional experimental simulation (SM_{var}), we used a climatology of OISSTv2, but retained

the interannual variability of soil moisture by prescribing the full time series of data from the AMIP simulation. This approach, which enabled us to isolate the indirect effects of SST-driven soil moisture variability from the direct effects of SST on atmospheric circulation, is similar to part of the experiment in *Orth and Seneviratne* [2017], but with a soil moisture time series that was created by forcing the climate model with observed SSTs. Finally, we performed a simulation that we refer to as NO_{var}, in which both SSTs and soil moisture were prescribed to a climatology. Comparing results from this simulation to the AMIP simulation enabled us to quantify the unforced internal variability of the atmosphere apart from any contributions from either remote SST forcing or local soil moisture response.

3.2.3 Observations and reanalysis

The Amazon is a region where interannual variability of climate is known to be strongly related to interannual SST variability [*Marengo*, 1992; *Foley et al.*, 2002; *Chen et al.*, 2011; *de Linage et al.*, 2013]. The AMIP simulation, driven by observed SSTs, was expected to reproduce the portion of the actual interannual variability of the Amazon climate that is controlled by SST-driven teleconnections, to the extent the associated atmospheric bridge mechanisms are represented in the model. We benchmarked temperature, precipitation, and total water storage anomalies from the AMIP run with observational and reanalysis data sets in order to determine how well E3SM was able to reproduce ENSO teleconnections impacting the land surface.

For benchmark temperature data, we used version 4.01 of the Climatic Research Unit Time Series (CRU TS4.01) [*Harris et al.*, 2014], the European Centre for Medium Range Forecasts Interim Reanalysis (ERA-Interim) [*Dee et al.*, 2011], and version 2 of the Modern-Era Retrospective Analysis for Research and Applications (MERRA-2) [*Gelaro et al.*, 2017]. For benchmark precipitation data, we used version 2.3 of the Global Precipitation Climatology Project (GPCP v2.3) [*Huffman et al.*, 2009] and the CPC Merged Analysis of Precipitation (CMAP) [*Xie and Arkin*, 1997]. Temperature

and precipitation benchmark data sets were conservatively regridded to the same 1° equirectangular grid as the model. For total water storage benchmarks, we used version 2 of the Gravity Recovery and Climate Experiment (GRACE) Jet Propulsion Laboratory (JPL) RLO5M mascon solution [*Watkins et al.*, 2015; *Wiese et al.*, 2015]. GRACE data were preserved in their original resolution of 3° mascons on a 0.5° grid, while modeled total water storage data were regridded to the same 0.5° grid and subsequently averaged over each of the 3° GRACE mascons.

3.2.4 Temporal classification and aggregation

We calculated a Niño3.4 index that represented the mean SST anomaly averaged over the eastern tropical Pacific (5°S–5°N and 170°–120°W). The monthly OISSTv2 data were smoothed with a 3-month center mean moving window, and the anomalies were estimated using a monthly climatology constructed from the entire time series (1982–2016). El Niño and La Niña years were defined as years that begin during an interval in which the Niño3.4 index exceeds a threshold of ±0.5°C (positive for El Niño, negative for La Niña) for five or more consecutive months. According to these criteria, El Niño years in our analysis were 1983, 1987, 1988, 1992, 1995, 1998, 2003, 2005, 2007, 2010, 2015, and 2016, and La Niña years were 1985, 1989, 1996, 1999, 2000, 2001, 2008, 2011, and 2012.

To focus on interannual variability, we calculated monthly anomalies by subtracting linear trends and long-term means. Several of the analyses described below involved aggregating monthly anomalies into wet and dry seasons. We defined wet and dry seasons from a terrestrial perspective, based on the mean annual cycle of total water storage in the Amazon basin measured by GRACE. As such, the wet season comprised the three months with the largest increase in total water storage (January—March), while the dry season included the three months with the largest decrease (July—September).

The SST anomalies used to define El Niño and La Niña years generally exhibited their largest values (positive or negative) during or soon before the wet season. The initial onset of El Niño and La Niña

conditions generally occurred around the end of the dry season. In consideration of the cascading response of the Amazon to ENSO [*Chen et al.*, 2017], we associated the dry season that follows the peak of the SST anomaly with the El Niño or La Niña years. For example, the SST anomalies that defined the 1983 El Niño began at the end of the 1982 dry season, but we considered only the dry season in 1983 to be associated with this particular El Niño event.

3.2.5 Data analysis

Temperature variability decomposition

We assessed the relative contributions of various drivers of temperature variability by comparing the variance of temperature anomalies from the AMIP simulation with those from the experimental simulations. The magnitude of temperature variability resulting from remote SST forcing and local land-atmosphere coupling may not be additive, because feedbacks with the atmosphere are prevented in the simulations with prescribed soil moisture. However, subtracting the variance of temperature anomalies in the NO_{var} from those of SST_{var} and SM_{var} provides a useful indication of how much additional variability is due solely to the direct SST forcing and subsequent soil moisture response, respectively. Normalizing these quantities by the variance of the AMIP simulation indicates the relative importance of each of these mechanisms in explaining the overall variability of temperature.

Model benchmarking

To determine mean state biases, we compared wet- and dry-season multi-year means of temperature and accumulated precipitation in both the AMIP simulation and the benchmark data. We computed summary statistics [*Taylor*, 2001] comparing monthly anomalies from benchmark data sets with those of the equivalent quantity in the AMIP simulation to evaluate how well E3SM v0.3 was able to simulate observed interannual variability. In order to assess how realistically the model responds

to ENSO, we compared correlations of the Niño3.4 index with both model and benchmark data during the wet and dry season. Finally, to gauge whether the model exhibited the correct sensitivity of temperature to soil moisture, we replicated the analyses described below with the benchmark temperature data sets.

El Niño–La Niña contrast

We calculated mean anomalies of surface air temperature, terrestrial ecosystem fluxes, and relevant biogeophysical variables using all of the El Niño and La Niña years, respectively. The resulting mean El Niño and La Niña composites do not represent the temporal structure of individual years within the ENSO cycle. Nevertheless, they serve as useful indicators of the timing of the different drivers of variability across the experimental simulations.

We defined the ENSO-driven contrast as the difference between the mean El Niño and La Niña composites during the Amazonian wet and dry seasons. Calculating this metric for temperature and the net CO_2 flux provided the basis upon which we quantitatively estimated the conceptual pathways in our hypothesis (Figure 3.1). We did not expect ENSO-driven contrasts to necessarily sum linearly between simulations, due to internal variability in addition to soil moisture feedbacks in the coupled AMIP simulation that were not captured by the mechanism-denial simulations. Therefore, we estimated each pathway as the range bounded on one side by the fraction of the contrast produced by including that mechanism, and on the other side by the fraction that denying that mechanism failed to produce. The width of the ranges provided an estimate of the importance of nonlinearities and soil moisture feedbacks (as well as any internal variability that may be present) in explaining the full contrast of the coupled AMIP simulation.

For example, to consider the conceptual pathways for temperature (T_{SST} and T_{SM} in Figure 3.1), we first calculated the ENSO-driven contrast (EC) in temperature for the AMIP, SST_{var}, and SM_{var} simulations, notated as EC_{AMIP}, EC_{SST}, and EC_{SM} respectively, and then estimated each pathway

as

$$T_{SST} \in \left\{ x \middle| \frac{EC_{SST}}{EC_{AMIP}} \leq x \leq \left(1 - \frac{EC_{SM}}{EC_{AMIP}}\right) \right\} \tag{3.1}$$

and

$$T_{SM} \in \left\{ x \middle| \frac{EC_{SM}}{EC_{AMIP}} \leq x \leq \left(1 - \frac{EC_{SST}}{EC_{AMIP}}\right) \right\}, \tag{3.2}$$

with an equivalent set of estimates for carbon fluxes (C_{SST} and C_{SM} in Figure 3.1) based on the ENSO-driven contrast of NEE.

3.3 Results

3.3.1 Drivers of interannual variability in temperature

The overall variability in detrended monthly surface air temperature as simulated by E3SM had a variance of about 0.15–0.45 $^{\circ}C^2$ throughout most of the Amazon (Figure 3.2a), with a mean variance of 0.32 $^{\circ}C^2$ across the entire watershed. The fractions of the variance attributable to each driver were approximately additive, and indicated the relative importance of each driver across the watershed. The variability was relatively high in the south, as a consequence of contributions from internal atmospheric variability (Figure 3.2b). In central and western portions of the basin, where the total variability was low, the variability was mostly driven by SST (Figure 3.2c). The contribution from soil moisture variability was mainly observed in the eastern part of the Amazon basin (Figure 3.2d).

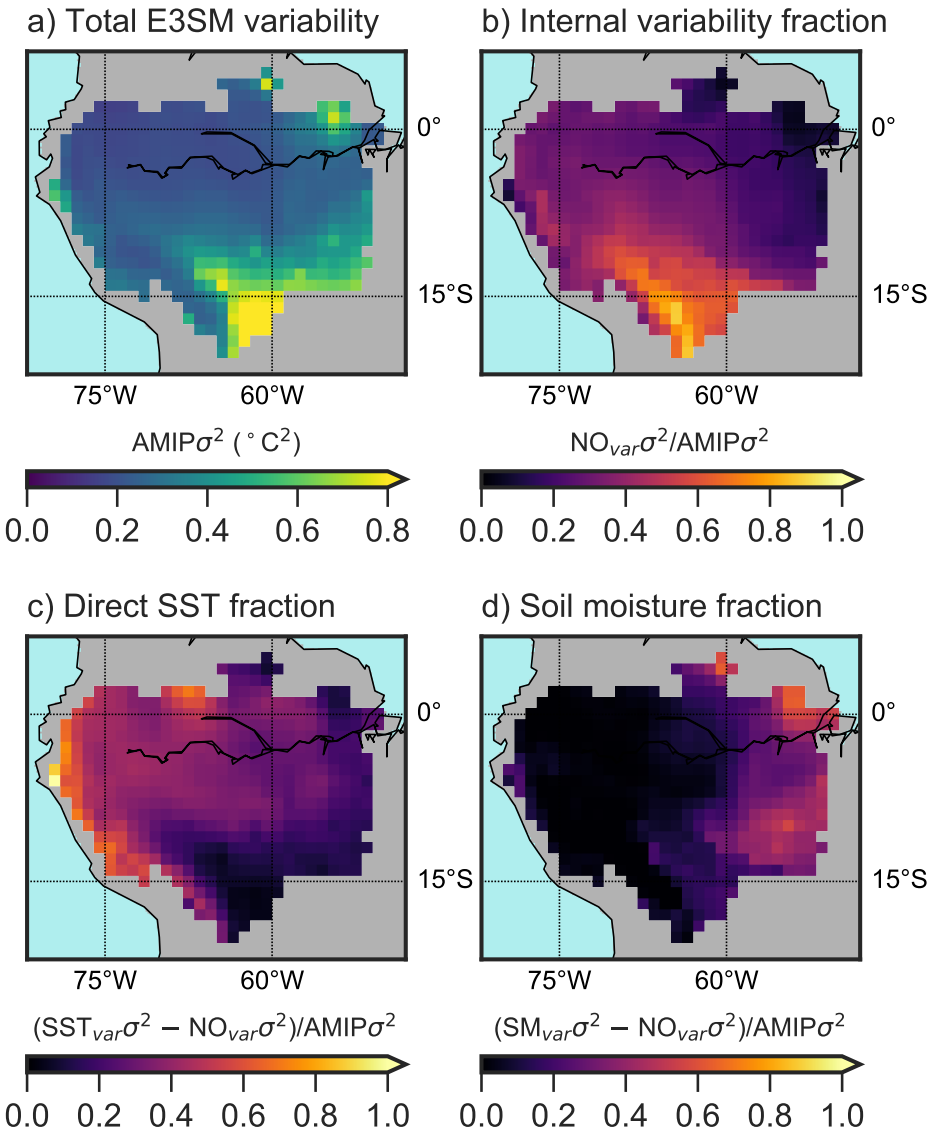

Figure 3.2: Decomposition of interannual variability of surface air temperature in E3SM for 1982–2016. Variance of detrended monthly temperature anomalies from the AMIP simulation, in units of $°C^2$ (a). Ratio of the variance of temperature anomalies from the NO_{var} simulation to that of the AMIP simulation (b). Ratio of the difference between the variances of temperature anomalies from the SST_{var} and NO_{var} simulations to that of the AMIP simulation (c). Ratio of the difference between the variances of temperature anomalies from the SM_{var} and NO_{var} simulations to that of the AMIP simulation (d).

3.3.2 Benchmarking E3SM with observations and reanalysis

Model validation

E3SM produced mean annual temperature within the range of the benchmark data sets over most of the Amazon, with the exception of the Andes (Figure 3.3). The AMIP run was cooler than the CRU TS4.01 data by approximately 2–3 °C, warmer than the ERA-Interim data by approximately 1–2 °C, and within about 1 °C of the MERRA-2 data over the non-montane portion of the basin. The disagreement across benchmark data meant the bias between the model and any one dataset is less than the bias between the gridded observations from CRU and the reanalysis data. E3SM represented the magnitude of interannual variability within the range of the benchmark data (Figure 3.4), and did a reasonable job simulating its timing (Figure 3.5).

E3SM simulated a mean precipitation that was considerably lower than the observations from GPCP throughout most of the Amazon, producing excess rainfall only at the Andean margin (Figure 3.6), with a bias pattern in this region consistent with other CMIP5 models [*Joetzjer et al.*, 2013; *Yin et al.*, 2013]. A similar spatial pattern was evident in the magnitude of variability, as the ratio of modeled to observed standard deviations (of monthly anomalies) was low throughout most of the Amazon while high in the Andes (Figure 3.7). Despite relatively low correlation of modeled precipitation anomalies with those of both precipitation data sets (Figure 3.8), modeled terrestrial water storage anomalies during the dry season were correlated relatively highly with GRACE in the eastern Amazon (Figure 3.9). This apparent contradiction likely results from the greater spatial homogeneity of water storage relative to precipitation; even if the model could not capture the precise spatial structure of precipitation anomalies, runoff distributed the resulting water storage anomalies more evenly.

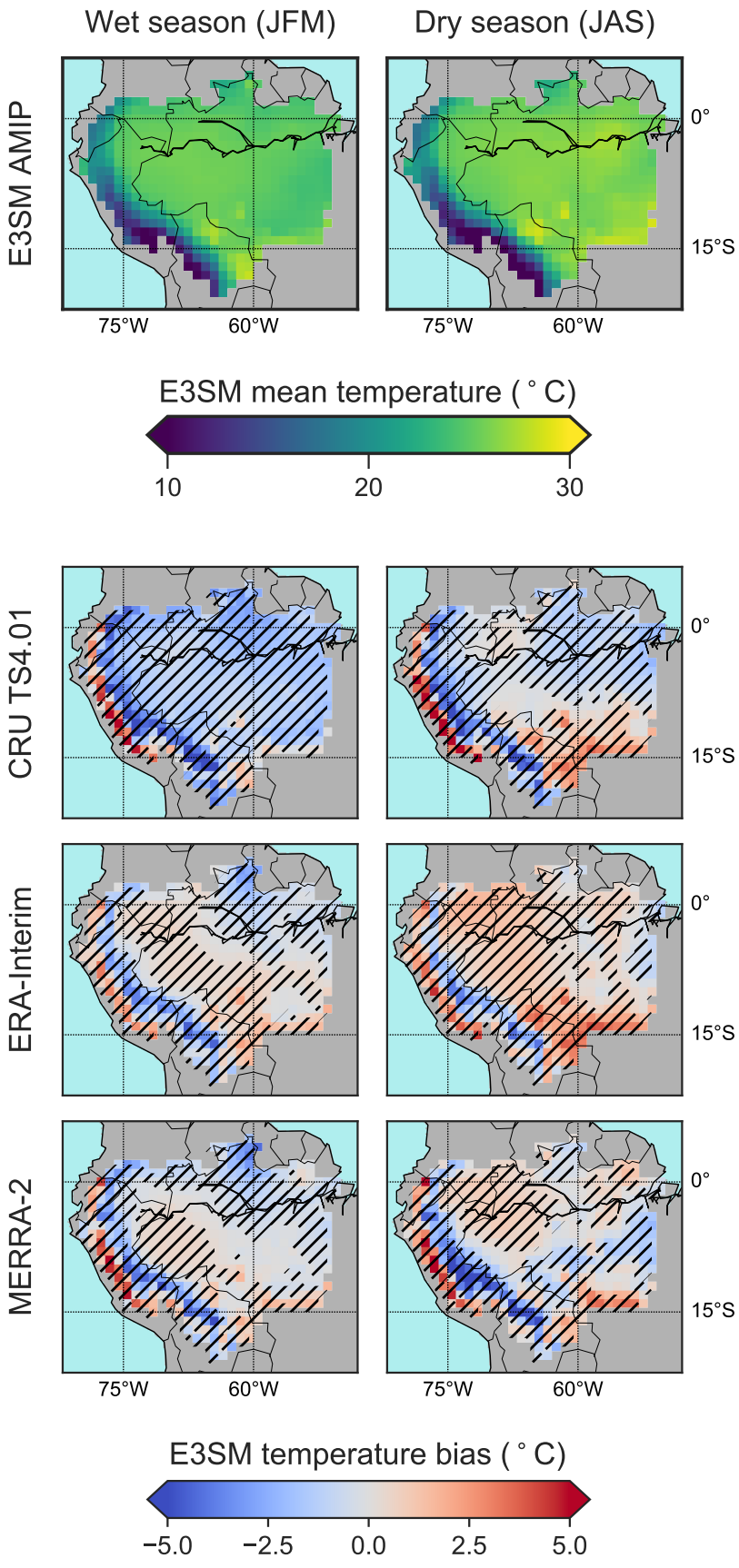

Figure 3.3: The first row is the mean temperature (in °C) in the E3SM AMIP simulation during the wet season (JFM) and dry season (JAS) for the years 1982—2016. The remaining rows are the temperature biases (model minus benchmark data) for the CRU TS4, ERA-Interim, and MERRA-2 data sets for the wet season (left column) and dry season (right column) of the same years. Crosshatching indicates that the difference between means (n =35) is statistically significant at $p \leq 0.05$.

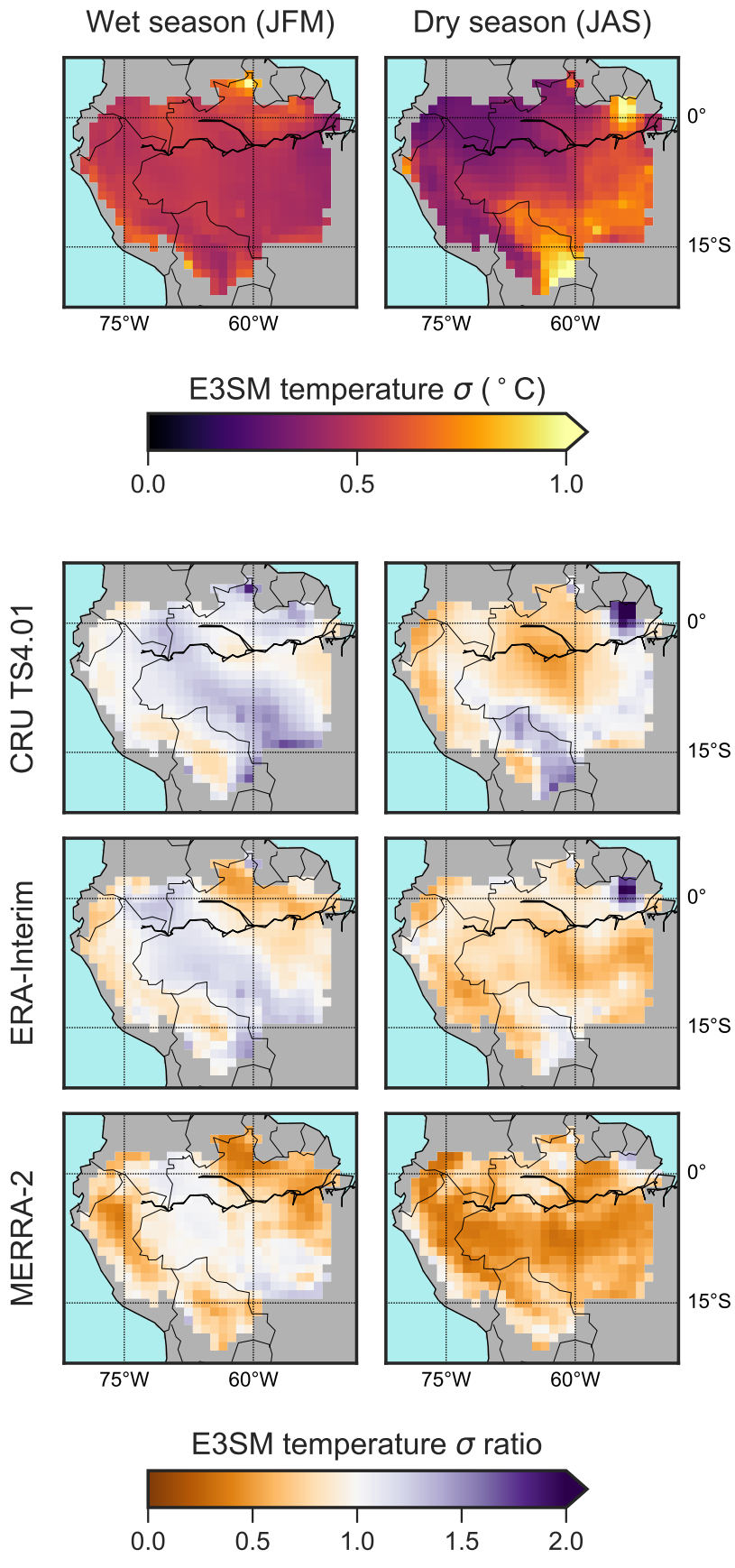

Figure 3.4: The first row is the standard deviation of temperature (in °C) in the E3SM AMIP simulation during the wet season (JFM) and dry season (JAS) across the years 1982—2016. The remaining rows are the ratios of the modeled standard deviation to those from the CRU TS4, ERA-Interim, and MERRA-2 data sets for the wet season (left column) and dry season (right column) across the same years.

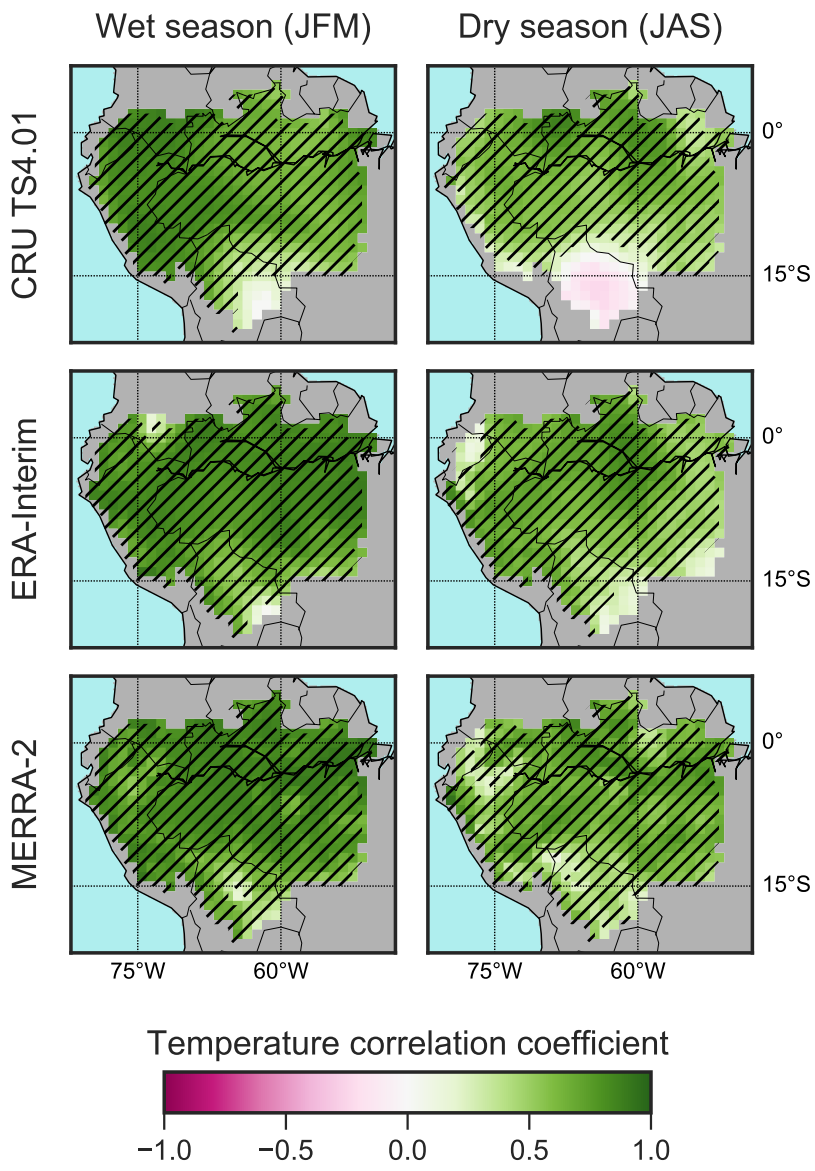

Figure 3.5: Pearson's correlation coefficient for the temperature in the E3SM AMIP simulation versus those from the CRU TS4, ERA-Interim, and MERRA-2 data sets during the wet season (JFM, left column) and dry season (JAS, right column) for the years 1982-–2016. Correlations were calculated between time series of the seasonal means across the 35 years in the study period. Cross-hatching indicates that the correlation coefficient was statistically significant at $p \leq 0.05$.

Benchmarking the ENSO–Amazon teleconnection

E3SM simulated reasonably well the relationship between Niño3.4 SSTs during December–February and precipitation anomalies during subsequent wet (January–March) and dry (July–September) seasons (Figure 3.10). SST anomalies were negatively correlated with wet-season precipitation in most of the northern Amazon basin, particularly towards the east, which was consistent with both benchmark data sets. Statistically significant positive correlations in the southern Amazon were present in E3SM but not in the benchmark data. During the dry season, E3SM showed a

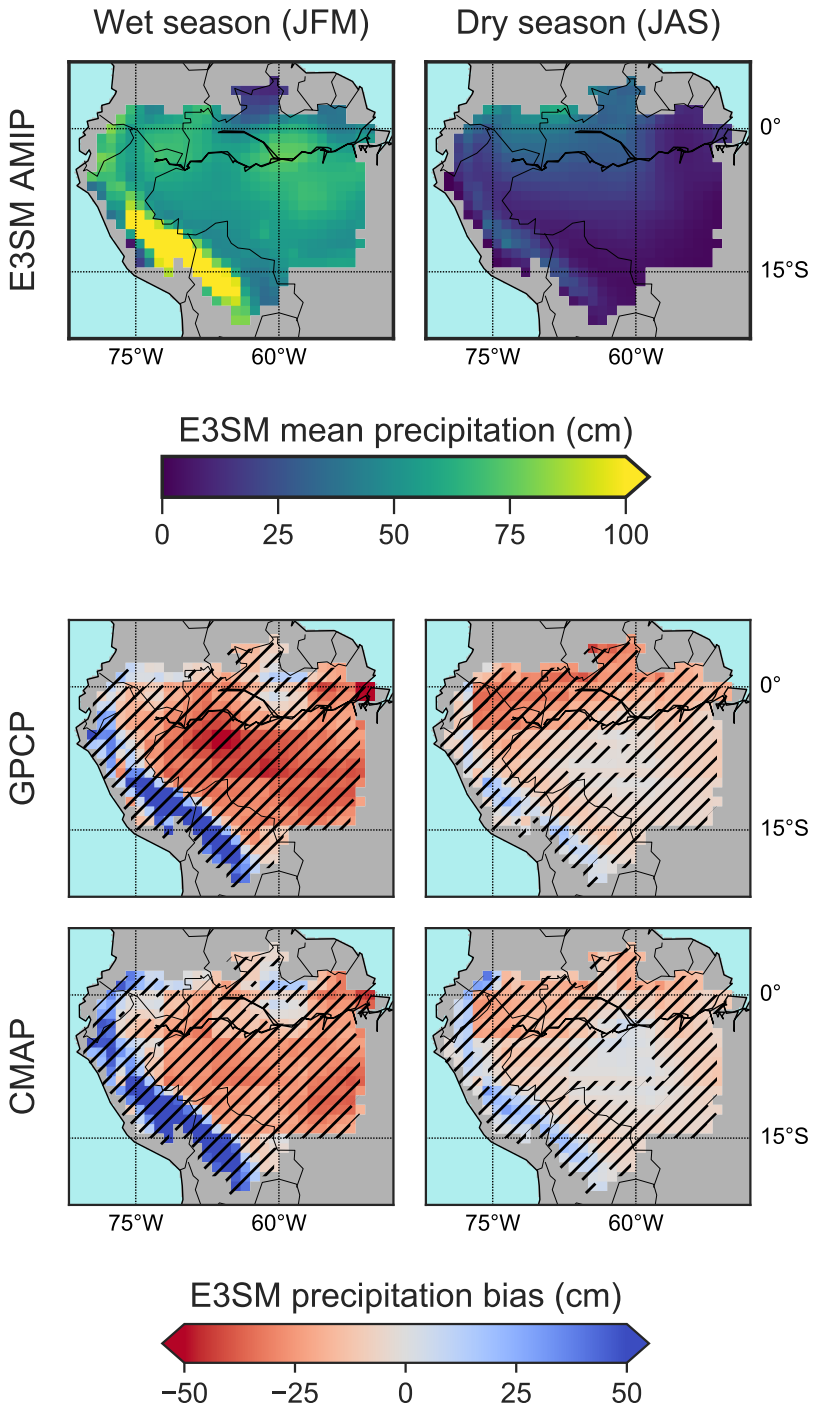

Figure 3.6: The first row is the mean precipitation (in cm) in the E3SM AMIP simulation accumulated during the wet season (JFM) and dry season (JAS) for the years 1982—2016. The remaining rows are the precipitation biases (model minus benchmark data) for the GPCP v2.3 and CMAP data sets for the wet season (left column) and dry season (right column) of the same years. Crosshatching indicates that the difference between means (n =35) is statistically significant at $p \leq 0.05$.

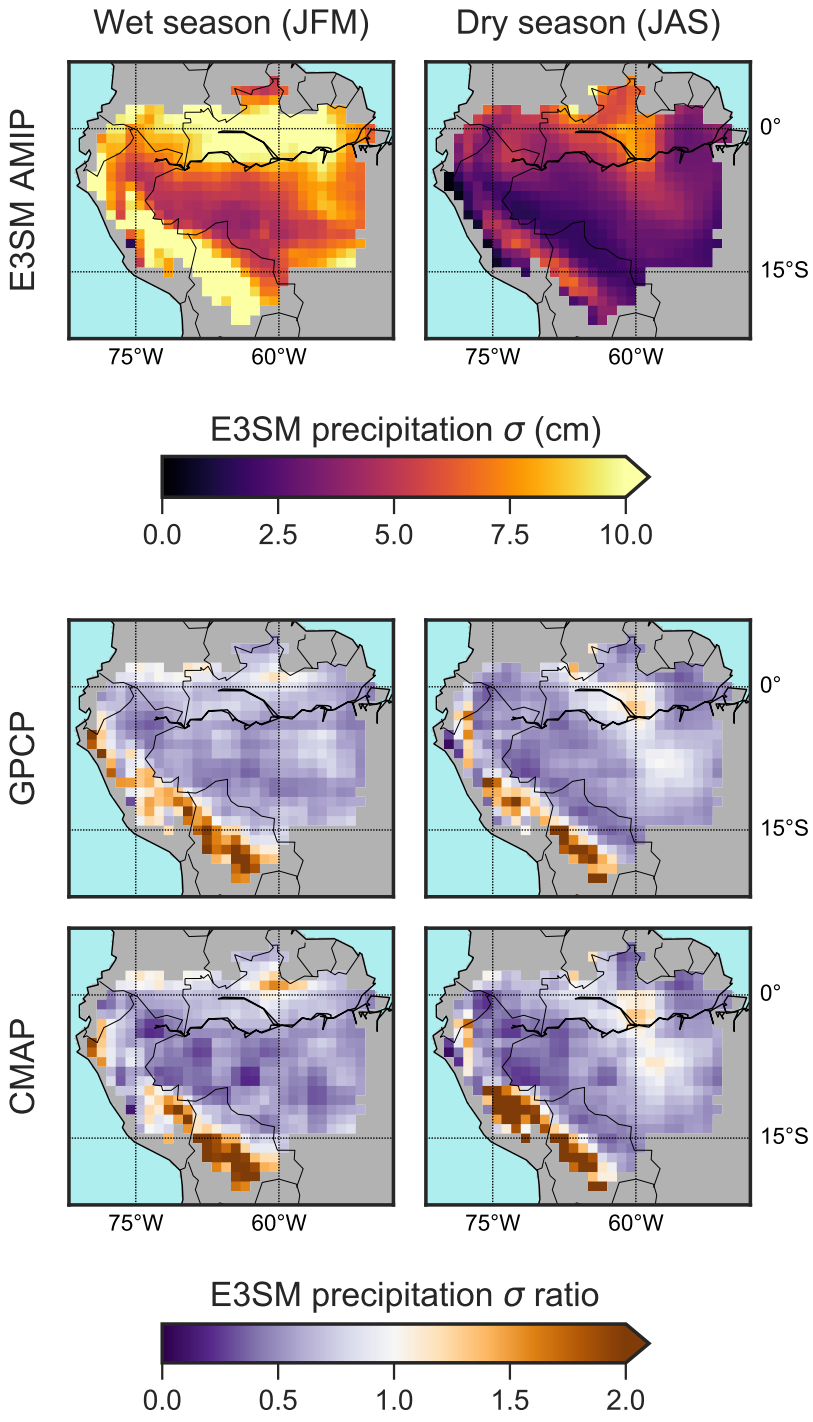

Figure 3.7: The first row is the standard deviation of precipitation (in cm) in the E3SM AMIP simulation accumulated during the wet season (JFM) and dry season (JAS) across the years 1982—2016. The remaining rows are the ratios of the modeled standard deviation to those from the GPCP v2.3 and CMAP data sets during the wet season (left column) and dry season (right column) across the same years.

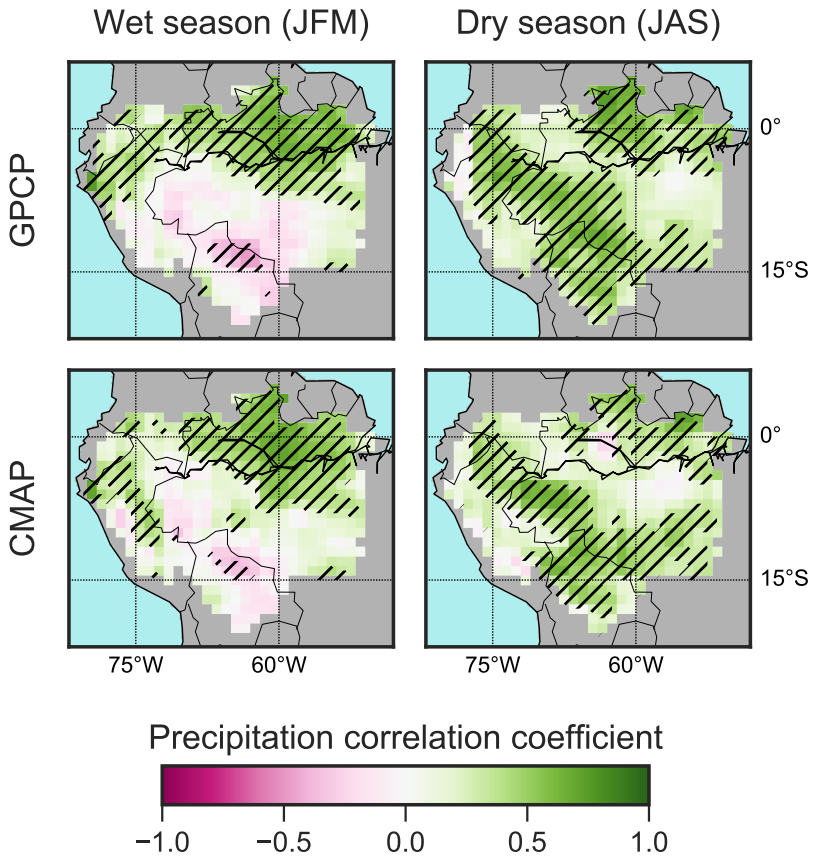

Figure 3.8: Pearson's correlation coefficient for the precipitation in the E3SM AMIP simulation versus those from the GPCP v2.3 and CMAP data sets during the wet season (JFM, left column) and dry season (JAS, right column) for the years 1982—2016. Correlations were calculated between time series of the seasonal means across the 35 years in the study period. Cross-hatching indicates that the correlation coefficient was statistically significant at $p \leq 0.05$.

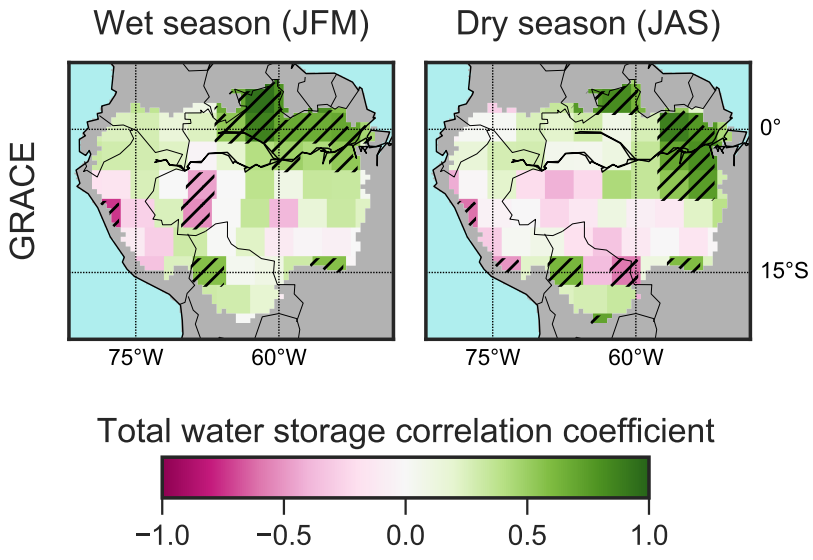

Figure 3.9: Pearson's correlation coefficient for the total water storage in the E3SM AMIP simulation versus that from the GRACE data set during the wet season (JFM, left column) and dry season (JAS, right column) for the years 2003—2016. Correlations were calculated between time series of the seasonal means across the 14 years in the GRACE record. Cross-hatching indicates that the correlation coefficient was statistically significant at $p \leq 0.05$.

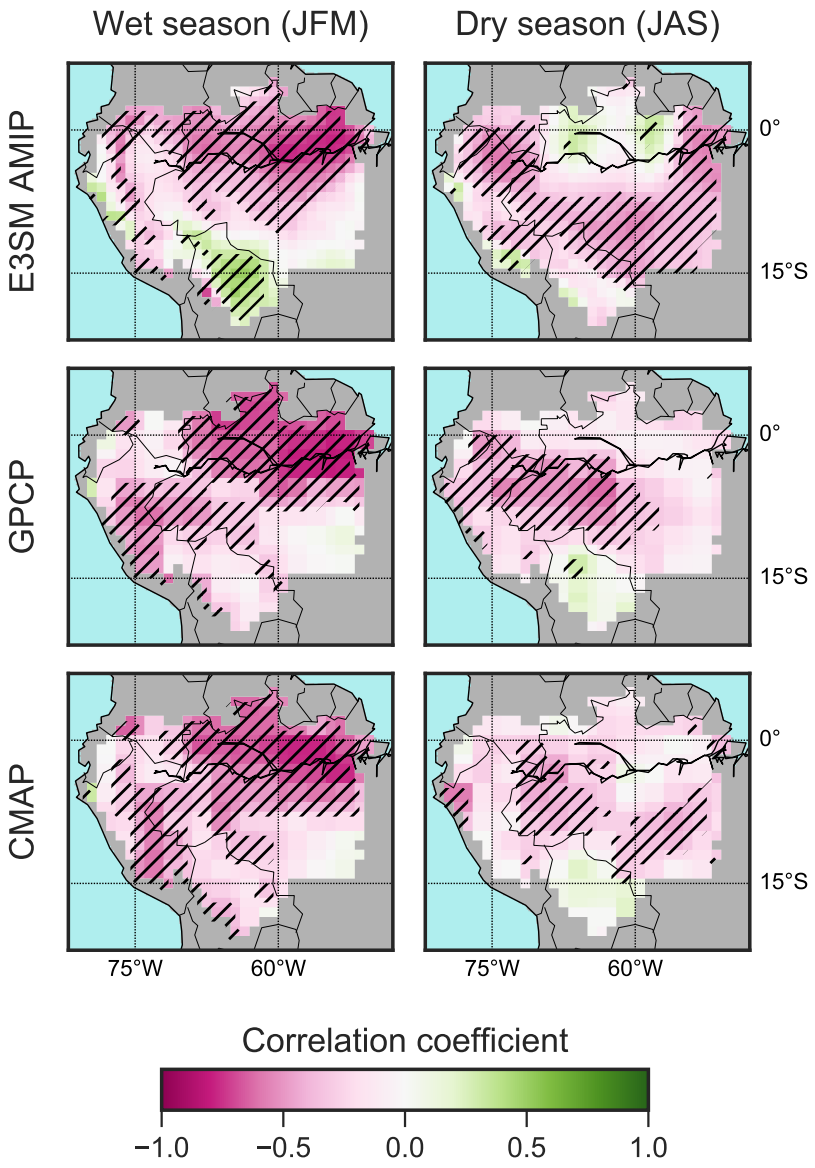

Figure 3.10: Pearson's correlation coefficient comparing the Niño3.4 index averaged over December–February with precipitation anomalies from E3SM (AMIP simulation), GPCP v2.3, and CMAP averaged over the wet season (January–March, left column) and dry season (July–September, right column) for 1982–2016. Cross-hatching indicates the correlation coefficient was statistically significant at $p \leq 0.05$.

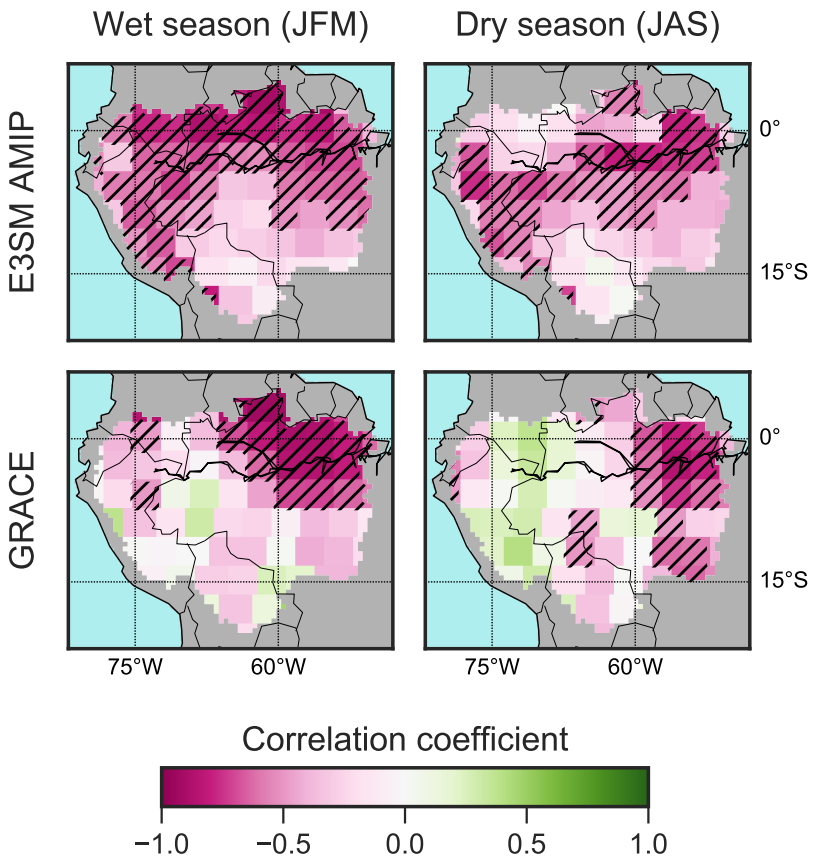

Figure 3.11: Pearson's correlation coefficient comparing the Niño3.4 index averaged over December–February with total water storage anomalies from E3SM (AMIP simulation) and GRACE averaged over the wet season (January–March, left column) and dry season (July–September, right column) for 2003–2016. Cross-hatching indicates the correlation coefficient was statistically significant at $p \leq 0.05$.

stronger negative correlation in the eastern Amazon than the GPCP precipitation data and, to a lesser extent, the CMAP data. However, the impact of this correlation bias only weakly affected soil moisture variability because of the smaller overall amount of precipitation during this time of year and carry-over in moisture storage from the previous wet season.

E3SM simulated ENSO teleconnections with total water storage in the Amazon qualitatively better than precipitation, compared with the available data from GRACE (Figure 3.11). In the wet season, E3SM and GRACE showed a similar pattern, suggesting that evapotranspiration and runoff must have had compensating biases. In the dry season, there was less agreement between E3SM and GRACE in the west and central Amazon, but both the model and observations showed significant correlations

in the east, despite less significant dry-season correlations in the benchmark precipitation data.

Although there were large differences among benchmark data sets, both model and benchmark data showed a positive correlation between the Niño3.4 index and the surface air temperature anomalies in the Amazon during both the wet and dry seasons (Figure 3.12). The correlations were higher in the wet season than in the dry season across all data sets. The dry season exhibited both the greatest disparity between benchmark data sets as well as the greatest disagreement between the model and benchmark data. E3SM and MERRAv2 both showed significant correlations across most of the Amazon during the dry season, but the ERA-Interim and CRU TS4.01 data sets each showed respectively less.

3.3.3 Amazon temperature response to ENSO forcing

The ENSO-driven contrast

The ENSO-driven contrast in temperature from the different model experiments revealed the season when and location where temperature anomalies originated from direct SST forcing and land-atmosphere moisture coupling (Figure 3.13). During the wet season, El Niño years exhibited higher temperatures than La Niña years throughout the Amazon, with most of the difference in the AMIP simulation attributable to forcing from remote SSTs (the SST_{var} simulation) and very little from land-atmosphere moisture coupling (the SM_{var} simulation). During the dry season, however, the ENSO-driven contrast was stronger in the SM_{var} simulation, particularly in the eastern part of the basin. Thus, land-atmosphere moisture coupling played a major role in determining the ENSO–temperature teleconnection during the dry season, and, more generally, in extending the duration of ENSO-induced temperature anomalies within the Amazon.

We further examined temperature anomalies in the eastern Amazon (east of 60°W) by exploring their temporal evolution during a full ENSO cycle (Figure 3.14). Table 3.2 summarizes the seasonal

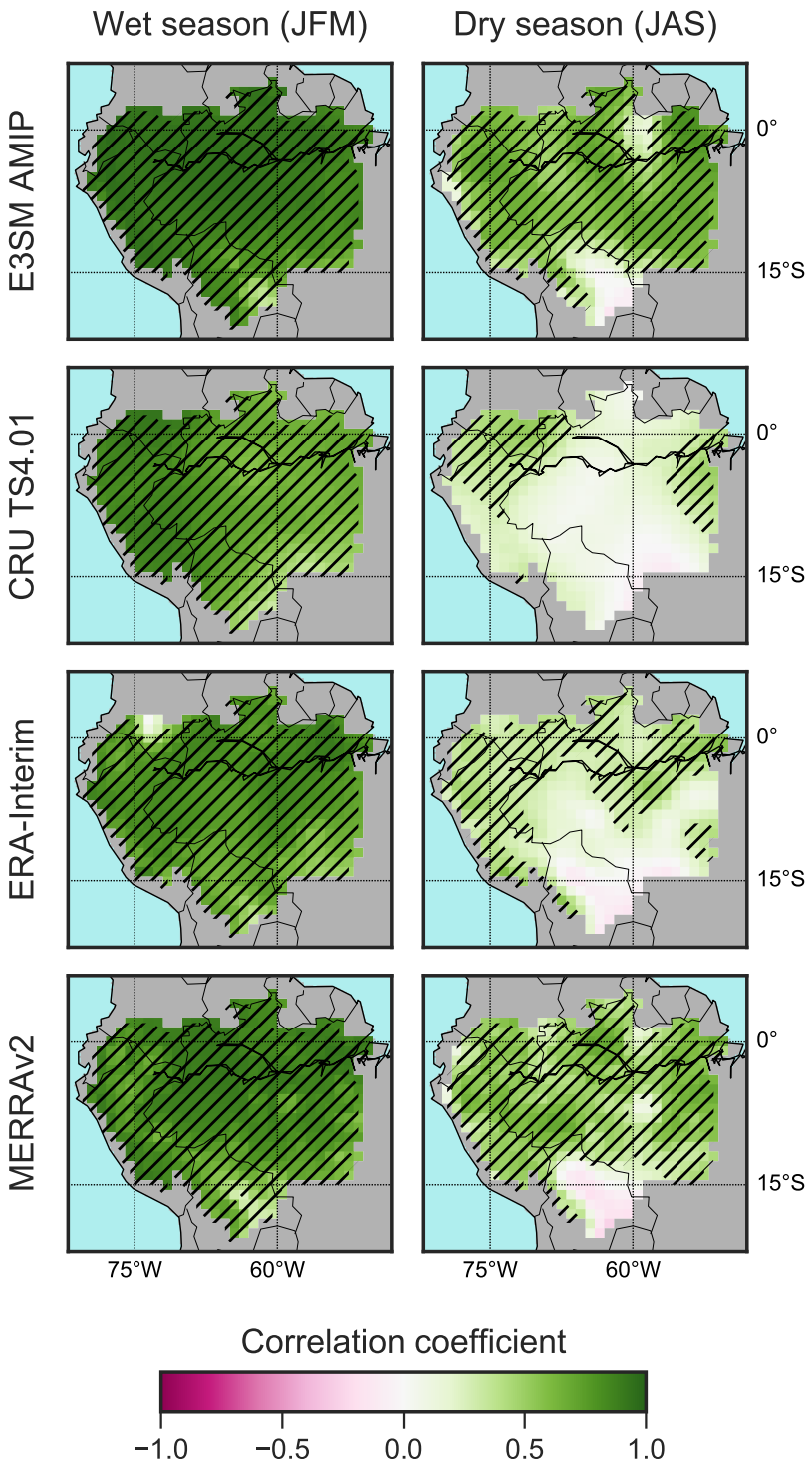

Figure 3.12: Pearson's correlation coefficient comparing the Niño3.4 index averaged over December–February with temperature anomalies from E3SM (AMIP simulation), CRU TS4.01, ERA-Interim, and MERRA-2 averaged over the wet season (January–March, left column) and dry season (July–September, right column) for 1982–2016. Cross-hatching indicates the correlation coefficient was statistically significant at $p \leq 0.05$.

means of wet- and dry-season contrasts in this region, which were used to quantitatively estimate our hypothesized conceptual pathways reported in Table 3.3. During the wet season in the eastern Amazon, the fully coupled (AMIP) temperature anomaly was 0.81 °C higher for El Niño years than it was for La Niña years. Direct SST forcing reproduced 81% of that anomaly when isolated in SST_{var}, and failed to reproduce 82% when denied in SM_{var}. Thus, wet-season T_{SST} was estimated as 81–82%, and, similarly, T_{SM} as 18–19%. During the dry season, the AMIP temperature anomaly showed a second peak of nearly the same magnitude (0.71 °C). However, during this season, land-atmosphere moisture coupling was the dominant driver (T_{SM} = 67–82%), while direct SST forcing had only a secondary effect (T_{SST} = 18–33%).

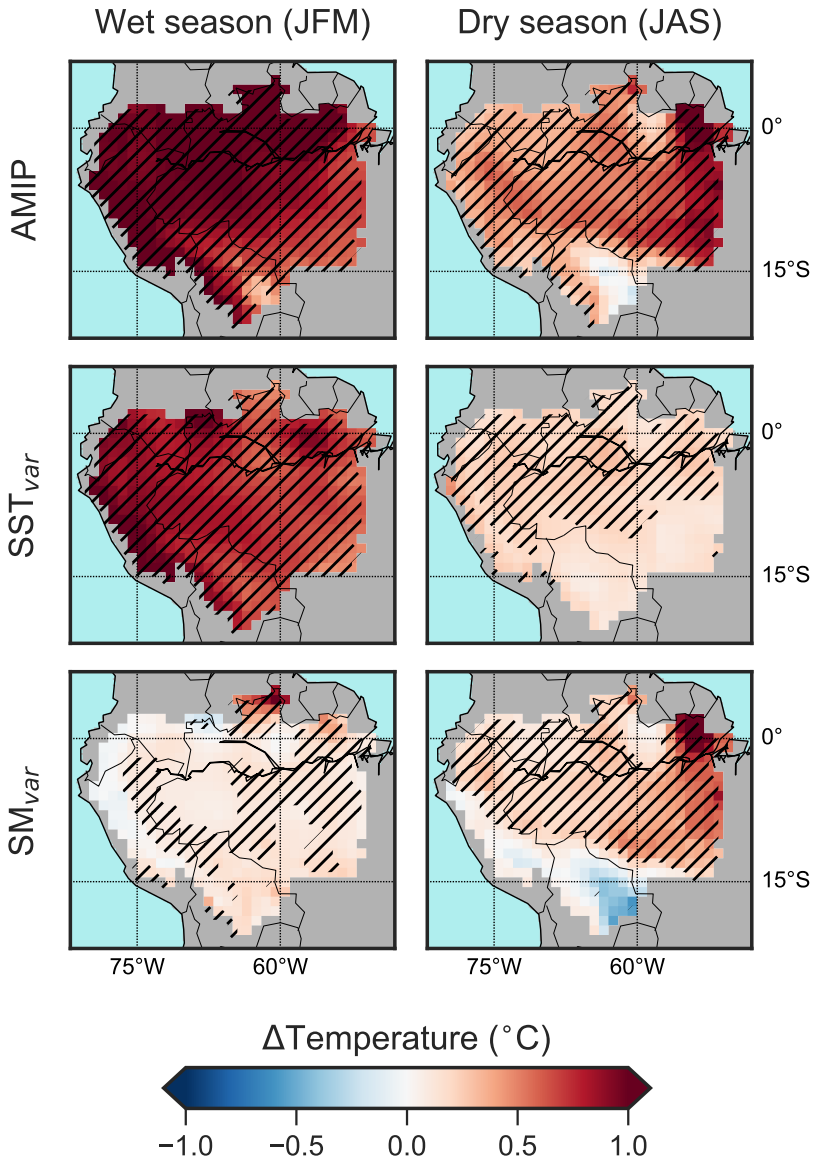

Figure 3.13: Difference between surface air temperature anomalies averaged across El Niño and La Niña years during the wet season (January–March, left column) and dry season (July–September, right column) in the AMIP, the SST$_{var}$, and SM$_{var}$ simulations. Cross-hatching indicates that the difference between El Niño and La Niña means is statistically significant at $p \leq 0.05$.

Drivers of the ENSO temperature contrast

The role of land–atmosphere coupling in the eastern Amazon was evident in the time series of evapotranspiration, downwelling shortwave radiation (insolation), and vapor pressure deficit (VPD) anomalies (Figure 3.15). During the wet season, evapotranspiration and its drivers (insolation and VPD) from the AMIP simulation exhibited a positive relationship with the ENSO phase, consistent with an evaporative regime that is not moisture-limited [*Seneviratne et al.*, 2010]. This positive relationship was also present in the SST_{var} simulation (though somewhat weaker), and there was no wet-season evapotranspiration contrast in the SM_{var} simulation, indicating that evapotranspiration was limited by radiation, rather than moisture availability.

During the dry season, the AMIP simulation maintained a positive relationship between the ENSO phase and VPD, but the relationship with evapotranspiration was reversed, consistent with a moisture-limited evaporative regime. The similarly large contrasts in the SM_{var} simulation were consistent with the interpretation that soil moisture anomalies drove the temperature contrast during

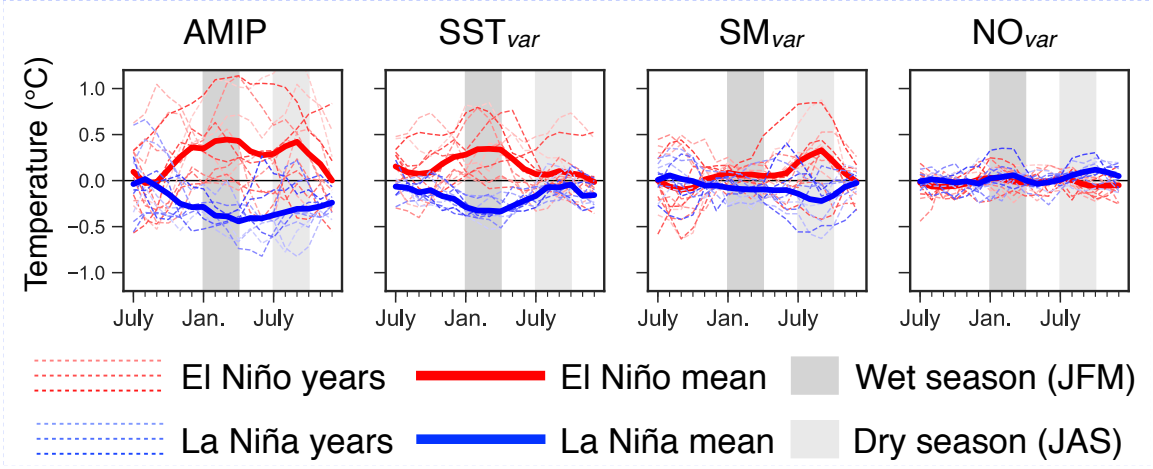

Figure 3.14: Temporal evolution of monthly temperature anomalies in the eastern Amazon in the AMIP, SST_{var}, SM_{var}, and NO_{var} E3SM simulations. Monthly surface air temperature anomalies were averaged across all grid cells in the Amazon watershed east of 60°W from the July preceding each El Niño (red) and La Niña (blue) year through the following December. Individual years are plotted with dashed lines, with a solid line for the mean of El Niño and La Niña years. Gray regions delineate the months in the wet season (January–March) and dry season (July–September). Monthly data were smoothed with a three-month centered moving average for clarity.

Table 3.2: Differences between mean El Niño and mean La Niña anomalies in eastern Amazon (east of 60°W).

	Wet season (JFM)			Dry season (JAS)		
	AMIP	SST_{var}	SM_{var}	AMIP	SST_{var}	SM_{var}
Temperature (°C)	0.81	0.67	0.15	0.71	0.13	0.48
Vapor pressure deficit (hPa)	55.1	24.2	23.0	204	31.9	162
Insolation (W/m^2)	7.69	2.04	4.34	1.83	2.13	-0.65
Evapotranspiration (mm/d)	0.12	0.08	0.03	-0.27	0.01	-0.31
Net primary production ($gC/m^2/d$)	-0.16	-0.07	-0.07	-0.47	-0.02	-0.43
Heterotrophic respiration ($gC/m^2/d$)	-0.01	0.07	-0.08	-0.16	0.00	-0.17
Net ecosystem exchange ($gC/m^2/d$)*	0.16	0.14	-0.01	0.32	0.03	0.27

*Net ecosystem exchange is a positive flux to the atmosphere. It is the balance between net primary production (positive is flux to the land surface) and heterotrophic respiration (positive is flux to atmosphere), though the rows may not sum perfectly due to rounding error.

the dry season. Insolation contrasts were relatively low across the mean composites, but there was a large degree of noise in the individual years and high contrast in the NO_{var} simulation, indicative of higher internal variability during the dry season.

ENSO-driven temperature contrast in benchmark data sets

The E3SM temperature response to ENSO variability in the eastern Amazon falls within the range of available benchmark data sets (Figure 3.16). Both the ERA-Interim and, in particular, the CRU TS4.01 data sets showed a much weaker contrast throughout the year, and particularly during the dry season, than the E3SM AMIP simulation. However, the MERRA-2 reanalysis showed a somewhat stronger temperature contrast than the E3SM AMIP simulation, and it is the only benchmark data

Table 3.3: Contribution of SST and soil moisture variability to the differences in temperature and the net CO_2 flux (see Figure 1) between mean El Niño and mean La Niña anomalies, as percent of the difference in the AMIP simulation, in eastern Amazon (east of 60°W)

	Temperature		Net ecosystem exchange	
	T_{SST}	T_{SM}	C_{SST}	C_{SM}
Wet Season	81–82%	18-19%	89–108%	-8–11%
Dry Season	18–33%	67–82%	9–15%	85–91%

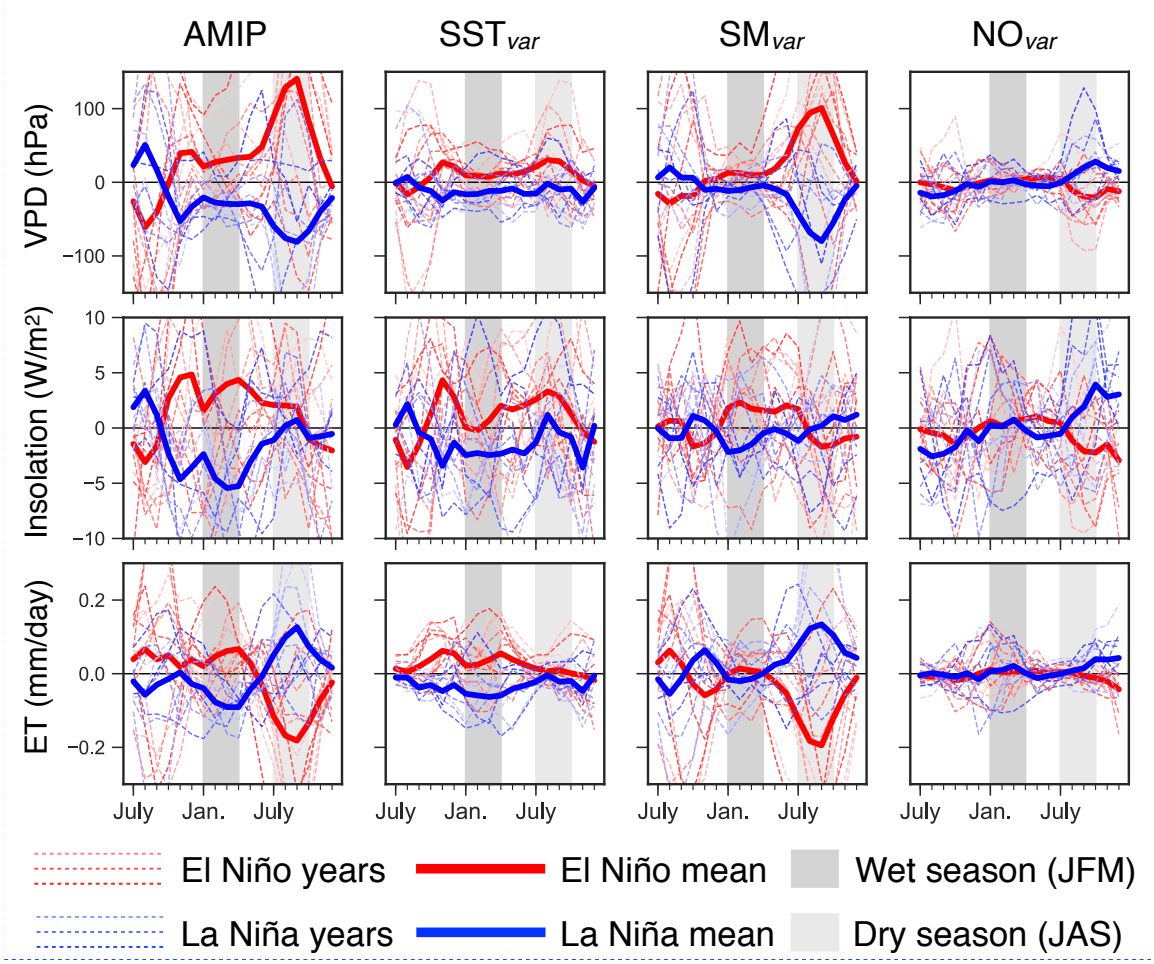

Figure 3.15: Temporal evolution of monthly anomalies of biogeophysical variables in the eastern Amazon in the AMIP, SST$_{var}$, SM$_{var}$, and NO$_{var}$ E3SM simulations. Monthly anomalies of vapor pressure deficit (VPD), downwelling shortwave radiation (insolation), and evapotranspiration (ET) were averaged across all grid cells in the Amazon watershed east of 60°W from the July preceding each El Niño (red) and La Niña (blue) year through the following December. Individual years are plotted with dashed lines, with a solid line for the mean of El Niño and La Niña years. Gray regions delineate the months in the wet season (January–March) and dry season (July–September). Monthly data were smoothed with a three-month centered moving average for clarity.

set that, like E3SM, showed a dry-season contrast as strong as the one from the wet season. The disagreement between benchmark data sets presented a challenge to evaluating the model; however, there was some indication that the temperature response to ENSO in E3SM may be too strong, particularly in response to soil moisture variability during the dry season.

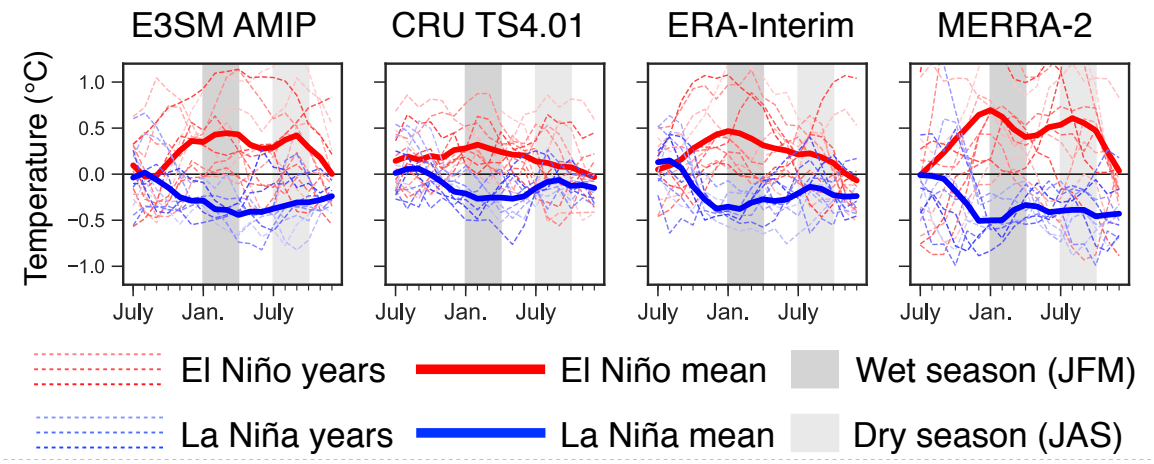

Figure 3.16: Same as Figure 3.14 but for the E3SM AMIP simulation alongside the CRU TS4.01, ERA-Interim, and MERRA-2 benchmark data sets.

Lag-correlation analysis

A lag-correlation analysis demonstrated how soil moisture variability in E3SM served to intensify and prolong the response of temperature in the eastern Amazon to SST forcing from the Niño3.4 region (Figure 3.17a). The SST_{var} simulation exhibited the maximum correlation between SST anomalies and eastern Amazon surface air temperature anomalies with a two or three months lag in air temperature. The SM_{var} simulation exhibited a weaker but still significant maximum correlation with a longer time scale, i.e., air temperature lagging SST by seven or eight months. The lag-correlation structure of the AMIP simulation was in between those of the SST_{var} and SM_{var} simulations.

For the first few months of lag time, correlations from the AMIP simulation resembled those of the SST_{var} simulation, consistent with initial forcing by SST. But the AMIP structure then attained a higher peak correlation at a longer lag time, and the predictability was prolonged consistent with support from the soil moisture interaction revealed by the SM_{var} correlation. The lagged correlations from benchmark data sets peaked somewhat earlier and lower than those from the AMIP simulation (Figure 3.17b). This could also be interpreted as an indication of temperature responding too strongly to soil moisture in E3SM. Alternatively, the timing inconsistency could have resulted

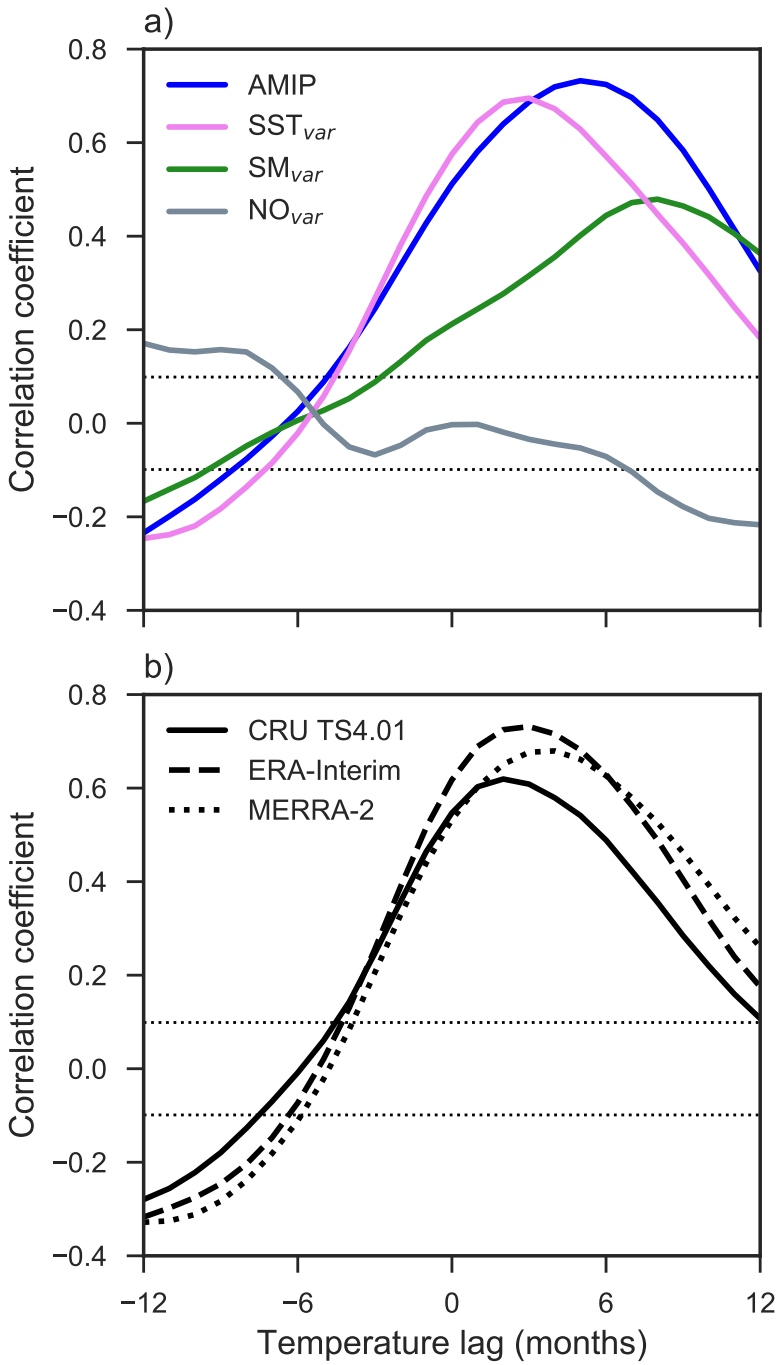

Figure 3.17: Temporal structure of Niño3.4 correlations with surface air temperature in the eastern Amazon for E3SM experiments (a) and benchmark datasets (b). Monthly temperature anomalies averaged across all grid cells in the Amazon watershed east of 60°W were correlated with the Niño3.4 index using lead and lag times up to 12 months. Dashed horizontal lines indicate correlation coefficients that are statistically significant at $p \leq 0.05$. Monthly data were smoothed with a three-month centered moving average for clarity.

from a tendency for ESMs to exhibit a delay in the seasonal cycle of precipitation in this region [*Joetzjer et al.*, 2013].

3.3.4 ENSO response of the carbon cycle

The ENSO-driven contrast in net ecosystem exchange (NEE, positive flux to atmosphere) across the E3SM experimental simulations showed a similar seasonal partitioning of drivers to that of the temperature contrast (Figure 3.18). During the wet season, the AMIP simulation showed a positive contrast throughout the Amazon basin, resulting from positive NEE anomalies (an increased source/reduced sink of atmospheric CO_2) during El Niño and negative anomalies during La Niña. This contrast was present, though reduced in strength, in the SST_{var} simulation, and was largely reversed in the SM_{var} simulation. During the dry season, there was a strong regional difference, particularly in the eastern Amazon, in both the AMIP and SM_{var} simulations that was absent in the SST_{var} simulation. This partitioning of drivers was particularly apparent in the east, which we further examined through the time series of NEE and the ecohydrologic drivers of NEE (Figure 3.19).

Interannual anomalies in net primary production (NPP) were driven primarily by soil moisture variability rather than direct SST forcing. NPP was suppressed by drier soils under El Niño conditions and enhanced by wetter soils under La Niña conditions throughout the year in both the AMIP and SM_{var} simulations, with a substantially larger contrast in the dry season than in the wet season. During the wet season, direct SST-driven temperature and radiation anomalies of the same sign (Figures 3.14 and 3.15) had opposite effects on NPP, but the influence of temperature was dominant, e.g., higher El Niño temperatures reduced NPP more than it was enhanced by increased sunlight. However, these NPP contrasts were small compared with those caused by soil moisture variability, particularly during the dry season.

Soil moisture variability was also the dominant driver of anomalies in heterotrophic respiration, but

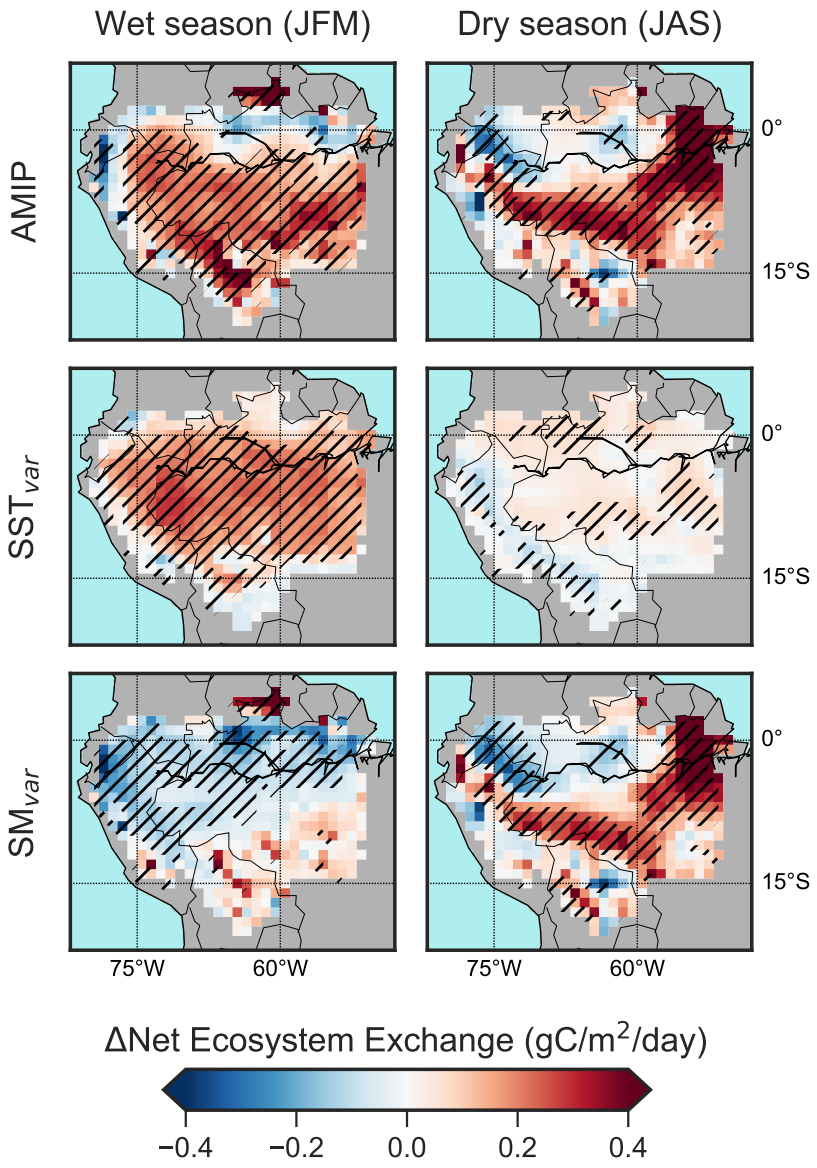

Figure 3.18: Difference between NEE (positive flux to atmosphere) anomalies averaged across El Niño and La Niña years during the wet season (January–March, left column) and dry season (July–September, right column) in the AMIP, the SST$_{var}$, and SM$_{var}$ simulations. Cross-hatching indicates that the difference between El Niño and La Niña means is statistically significant at $p \leq 0.05$.

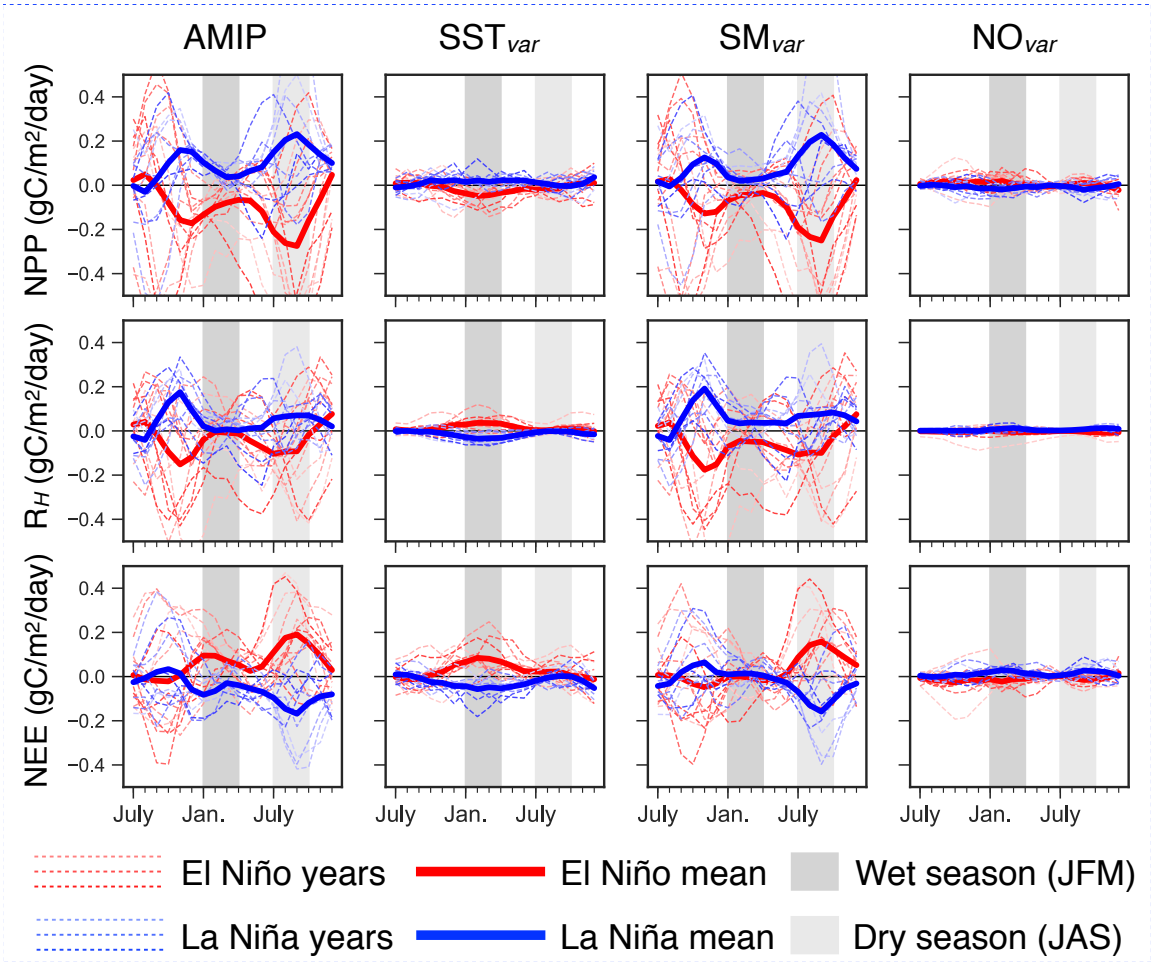

Figure 3.19: Temporal evolution of monthly anomalies of ecosystem variables in the eastern Amazon in the AMIP, SST$_{var}$, SM$_{var}$, and NO$_{var}$ E3SM simulations. Monthly anomalies of net primary production (NPP), heterotrophic respiration (R$_H$), and net ecosystem exchange (NEE) were averaged across all grid cells in the Amazon watershed east of 60°W from the July preceding each El Niño (red) and La Niña (blue) year through the following December. Individual years are plotted with dashed lines, with a solid line for the the mean of El Niño and La Niña years. Gray regions delineate the months in the wet season (January–March) and dry season (July–September). Monthly data were smoothed with a three-month centered moving average for clarity.

not necessarily for NEE. Dry El Niño soils reduced respiration and wet La Niña soils increased it throughout the year, particularly during the dry season. However, the magnitude of these effects relative to NPP led to distinct differences between the wet and dry seasons.

During the wet season, the coupled ENSO-driven contrast in eastern Amazon NEE was 0.16 gC/m^2/d (Table 3.2). Soil moisture effects on NPP and respiration were approximately equal, yielding only a small NEE contrast in the SM$_{var}$ simulation. At the same time, respiration anomalies driven by

direct SST forcing on temperature had a similar magnitude to the concurrent NPP anomalies but with the opposite sign. As a result, wet-season NEE anomalies were larger than the constituent NPP and respiration anomalies in both the AMIP and SST_{var} simulations. Therefore, despite substantial soil moisture effects on the components of NEE, wet-season C_{SM} was close to zero while C_{SST} was close to 100% (Table 3.3).

By the dry season, the fully coupled NEE anomalies grew to twice the magnitude of those occurring during the wet season, with an ENSO-driven contrast of 0.32 $gC/m^2/d$. Soil moisture effects on NPP and respiration were still of the same sign, but the effect on NPP was larger. Therefore, while respiration anomalies dampened the effect of NPP anomalies, there was still a strong NEE contrast in both the SM_{var} and AMIP runs. By this time of year, direct impact of SST forcing on temperature and radiation had weakened, and soil moisture variability was the dominant driver, with a dry season C_{SM} of 85–91%.

3.4 Discussion

3.4.1 Land-atmosphere coupling and ENSO

Our mechanism-denial experiments with E3SM illustrate how soil moisture variability served to intensify and extend the temperature response of the Amazon to forcing from ENSO. SST anomalies, which peaked during December–February, drove changes to atmospheric circulation and meteorology in the Amazon with little or no delay. The direct impact of these changes, absent any soil moisture interaction, was responsible for over four-fifths of the wet-season temperature response to ENSO in the eastern Amazon. Concurrent precipitation anomalies led to a soil moisture memory that persisted into the following dry season, resulting in a delayed temperature response via land-atmosphere coupling. This indirect impact of soil moisture was responsible for two-thirds to four-fifths of the contrast between El Niño and La Niña dry-season temperatures. These results

highlight how temperature anomalies set up by ocean-atmosphere climate modes are amplified and extended by land-atmosphere coupling.

Our results also illustrate the importance of soil moisture variability in the ENSO-modulated carbon cycle response of the eastern Amazon. During the wet season, soil moisture impacts on NPP and heterotrophic respiration counteracted each other, leading to a net neutral impact on NEE, while direct SST forcing on each of these fluxes amplified impacts on NEE. This gives direct SST forcing the appearance of driving nearly all of the wet-season ENSO-driven NEE contrast, despite soil moisture impacts on the constituent NPP and respiration anomalies. During the dry season, the carbon cycle response was driven primarily by soil moisture variability, which limited NPP more than heterotrophic respiration. At the same time, the temperature anomalies resulting from land-atmosphere coupling affected NPP with the same sign as the net soil moisture effect, with the NPP response dominating the overall NEE signal. This portion of the temperature anomaly (T_{SM}) combines with the direct control of soil moisture on NEE, leading to the dominance of C_{SM} during the dry season.

3.4.2 Uncertainties and limitations

While our experiment demonstrated the importance of land-atmosphere coupling in E3SM, extrapolation to the real Earth system must be tempered by consideration of the disagreement in the ENSO-driven temperature contrast between observations and reanalyses. The AMIP simulation was able to reproduce the mean temperature within the uncertainty of the benchmark temperature data, but it showed a stronger dry-season contrast than two of the three benchmark data sets. This may have resulted from unrealistically strong land-atmosphere coupling in this model, which would be consistent with previous studies that found land-atmosphere coupling to be too strong in the models on which E3SM is based [*Dirmeyer*, 2006; *Zeng et al.*, 2010; *Mei and Wang*, 2012; *Levine et al.*, 2016].

Land surface temperatures are more sensitive to moisture availability when that becomes the limiting factor for evapotranspiration, which mostly occurs under semi-arid conditions [*Seneviratne et al.*, 2010]. Therefore, any errors in the model that pushed the Amazon from its generally moist state towards toward drier conditions could have biased the land-atmosphere coupling strength. The dry bias in the AMIP simulation is a well-known feature of many ESMs, including the precursors to E3SM [*Joetzjer et al.*, 2013; *Yin et al.*, 2013], and could have caused excessive land-atmosphere coupling that led to an unrealistically large temperature contrast in the dry season. In addition, CLM4.5 does not include any hydraulic redistribution by deep roots [*Tang et al.*, 2015; *Wang et al.*, 2016b], which sustains dry-season plant growth and evapotranspiration in the Amazon [*Nepstad et al.*, 1994; *Lee et al.*, 2005; *Oliveira et al.*, 2005]. This may also have pushed the model toward drier conditions, by making water inaccessible by plants and unavailable for evapotranspiration.

On the other hand, there is also evidence that supports the possibility that E3SM may have underestimated the response to ENSO. The precipitation bias applied not only to the mean state, but also to the magnitude of interannual variability. Land–atmosphere coupling in the model ensured that errors in wet-season precipitation that yielded errors in subsequent dry season water storage anomalies would have led to errors in temperature, and potentially NEE as well. Insufficient precipitation variability in the eastern Amazon did not lead to insufficient water storage variability, at least over the limited time span of the GRACE record. However, in the places where water storage variability was too low, the temperature anomalies resulting from land-atmosphere would also have been too low, which could have caused the simulated response to ENSO to be too weak.

3.4.3 Implications for future research

Our results suggest that attempts to partition the relative importance of temperature and hydrology in the carbon cycle response to ENSO should consider the interdependence between these variables. In particular, studies that employ offline models that are forced by non-interactive temperature

and precipitation (or soil moisture) time series may overestimate the sensitivity of CO_2 fluxes to temperature and underestimate the sensitivity to hydrology [e.g., *Piao et al.*, 2013; *Wang et al.*, 2013]. While our experiments could not separate the direct ecosystem carbon response to soil moisture from the effect of soil moisture on temperature, they demonstrated that the largest ENSO-driven carbon cycle anomalies, in the eastern Amazon during the dry season, are spatiotemporally colocated with the strongest land-atmosphere coupling. Future work could use temperature time series from an experiment such as ours to drive an offline model, in order to further constrain how much of C_{SM} is directly due to T_{SM}.

The disagreements between the benchmark temperature data sets highlights the need to reduce uncertainties in observations and reanalyses. CRU TS4.01 is based on station observations, so in sparsely observed regions such as the Amazon, spatial interpolation could decrease interannual variability, which could have caused the overall low contrast in this data set. Reanalysis data sets showed stronger contrasts, but they are constrained at the surface by land surface models that may themselves suffer from the same bias in land-atmosphere coupling strength as E3SM. The ongoing efforts toward improving these data products are important for understanding this remote but climatically significant regions such as the Amazon.

3.5 Conclusion

We performed an experiment with a set of global E3SM simulations to decouple the direct effects of SST variability from the resulting soil moisture variability in the Amazon. We found that soil moisture anomalies served to intensify and prolong the response of the eastern Amazon climate to ENSO. The immediate component of the response to El Niño was driven directly by atmospheric circulation changes that increased temperatures and reduced precipitation during the wet season. Soil moisture anomalies persisted into the dry season, causing a delayed temperature response from land-atmosphere coupling. SST-driven soil moisture anomalies explained two-thirds to four-fifths

of the eastern Amazon's dry-season temperature response to ENSO in E3SM. Observational and reanalysis data suggest that E3SM may have overestimated the temperature response to ENSO, perhaps due to a bias in land-atmosphere coupling strength.

The drivers of the carbon cycle response to ENSO in E3SM were similar to those for temperature. In the eastern Amazon, soil moisture did not affect the net carbon cycle response to ENSO during the wet season, but it drove the majority of the dry season response. ENSO had a larger impact on dry-season carbon fluxes than those of the wet season, resulting from soil moisture limitations on ecosystem function combined with land-atmosphere coupling affecting temperature. This indicates the need to consider the interdependent relationship between temperature and the hydrologic cycle when attributing mechanisms to ENSO-driven variability in the tropical terrestrial carbon cycle.

Acknowledgements

This chapter was adapted from the following publication:

Levine, P. A., J. T. Randerson, Y. Chen, M. S. Pritchard, M. Xu, and F. M. Hoffman (2019), Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to El Niño–Southern Oscillation, *Journal of Climate*, *32*(4), 1273–1292. http://dx.doi.org/10.1175/JCLI-D-18-0150.1

We received funding support from the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation Scientific Focus Area (RUBISCO SFA), which is sponsored by the Regional and Global Model Analysis (RGMA) Program in the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research (BER) in the U.S. Department of Energy (DOE) Office of Science. This research used resources from Project m2467

of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility (DE-AC02-05CH11231), and from Project cli106bgc of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility (DE-AC05-00OR22725). P.A.L. received funding support from NASA Headquarters under the NASA Earth and Space Science Fellowship Program (NNX16AO38H). J.T.R. and Y.C. received additional funding support from the DOE Office of Science Earth System Modeling Program (DE-SC0006791) and NASA's SMAP and CMS programs. M.S.P. received funding support from the DOE Early Career Program (DE-SC0012152). M.X. and F.M.H. received additional funding support from the Energy Exascale Earth System Model (E3SM) Project and the Next Generation Ecosystem Experiments - Tropics (NGEE-Tropics) Project, sponsored by BER in the DOE Office of Science. Oak Ridge National Laboratory (ORNL) is managed by UT-Battelle, LLC, for the DOE (DE-AC05-00OR22725).

GPCP combined precipitation data were provided by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres, which develops and computes the dataset as a contribution to the GEWEX Global Precipitation Climatology Project, from their website at https://precip.gsfc.nasa.gov. CMAP Precipitation data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd. GRACE land data are available at http://grace.jpl.nasa.gov, supported by the NASA MEaSUREs Program. CRU TS data were obtained from the University of East Anglia through their website at https://crudata.uea.ac.uk. ERA-Interim data were obtained from the European Centre for Medium-Range Weather Forecasts through their website at http://apps.ecmwf.int/datasets. MERRA-2 data were obtained from the Goddard Earth Sciences Data and Information Services Center through their website at https://disc.sci.gsfc.nasa.gov. E3SM output from this experiment is archived on the National Energy Research Scientific Computing Center's High Performance Storage System and is available for download at http://portal.nersc.gov/archive/home/p/plevine/www/data_archive/e3sm_enso.

Chapter 4

The size of the global soil carbon sink constrained by radiocarbon observations and an Earth system model

4.1 Introduction

4.1.1 Uncertainty in the fate of soil carbon

Soil is the largest reservoir of biologically active carbon in the Earth system, but there is substantial uncertainty in the fate of soil carbon in a changing climate. Increasing atmospheric carbon dioxide (CO_2) concentrations enhances primary production of terrestrial vegetation [*Norby and Zak*, 2011; *Schimel et al.*, 2015], which increases carbon input into the soil. However, the rates at which soil carbon is decomposed and returned to the atmosphere as CO_2 are expected to increase with rising temperatures [*Davidson and Janssens*, 2006; *Heimann and Reichstein*, 2008; *Crowther et al.*, 2016; *Melillo et al.*, 2017]. The balance between these competing processes determines the magnitude of

the total soil carbon sink, which feeds back into the climate by exerting a control on atmospheric CO_2 concentrations [*Le Quéré et al.*, 2018].

Earth system models (ESMs) are important tools for making projections of future climate change, but there is substantial uncertainty in their simulations of soil carbon dynamics [*Todd-Brown et al.*, 2013; *Bradford et al.*, 2016; *Todd-Brown et al.*, 2018]. ESMs do not agree on the magnitude, or even the sign, of the projected 21st-century soil carbon sink [*Anav et al.*, 2013; *Todd-Brown et al.*, 2014]. Poor model representation of feedbacks introduces additional uncertainty into projections of future atmospheric CO_2 concentrations, and feeds back into the climate in biogeochemically coupled ESMs [*Friedlingstein et al.*, 2014].

The large uncertainties in ESMs stem from the difficulty of using information that was collected mostly at the site level and extrapolating it to the global scale [*Reichstein and Beer*, 2008; *Manzoni and Porporato*, 2009]. Soil organic carbon (SOC) is controlled by a complex network of interacting biogeochemical processes [*Ahrens et al.*, 2015; *Dwivedi et al.*, 2017]. However, in order to maintain computational efficiency, ESMs typically simplify these processes using first-order kinetics with parameterized decay rate constants. As such, there are large discrepancies not only between ESMs and observations, but also between the lumped parameter approach used in ESMs and more computationally expensive approaches that explicitly account for important processes such as microbial dynamics and mineral sorption [*Riley et al.*, 2014; *Ahrens et al.*, 2015]. While multi-compartment first-order decay formulations have proven capable of reproducing observed SOC states and fluxes, the issue of equifinality emphasizes the need for confronting models with multiple observational constraints[*Reichstein and Beer*, 2008; *Luo et al.*, 2016].

4.1.2 Radiocarbon as a model constraint

One of the most important constraints to modeled representations of SOC cycling, not only in ESMs but in all models, is measurements of the radioactive nuclide carbon-14 (^{14}C, or radiocarbon)

[*Trumbore*, 2009]. Radiocarbon naturally forms in the Earth's atmosphere when nitrogen atoms are bombarded by cosmic rays, and is then absorbed in the form of $^{14}CO_2$ by plants via photosynthesis. Living plants have a ^{14}C to ^{12}C that is in equilibrium with the atmosphere, but when the plant dies, the ^{14}C undergoes radioactive decay, and the ratio decreases over time. By measuring this ratio in a sample and comparing it with the background ratio, the sample can be "dated," i.e., the amount of time that has passed since the plant matter was in equilibrium with the atmosphere may be measured.

This form of chronometry is widely applied to archaeology, where it can be used to date material up to ~50,000 years old. The application to soil science is somewhat different, as soil contains a mixture of decomposed plant matter that spans a range of ages. As such, the "age" of a soil sample does not indicate a single date on which the soil formed, but instead indicates the mean age, which can be used to determine the mean transit time of carbon atoms in that soil [*Sierra et al.*, 2018].

Rates of organic matter decomposition are key parameters in models of soil biogeochemistry, and are related to the mean transit times. While radiocarbon does not directly measure decomposition rates, models that include a radiocarbon tracer can be constrained with radiocarbon observations. In other words, radiocarbon is something that models predict and can be observed.

Importantly, both model predicted and observed radiocarbon are related to processes that may take place across tens of thousands of years. It is, therefore, the only known way to constrain rates of soil decomposition that act at timescales beyond those that can be observed in the laboratory or field. In addition, above-ground thermonuclear weapons testing around the mid-1950s generated levels of radiocarbon in the atmosphere that were nearly double the background. This so-called "bomb spike" serendipitously provided scientists with an impulse-response tracer experiment for the entire Earth, and has been useful for constraining faster decay rates that act upon timescales of decades or less [*Schrumpf and Kaiser*, 2015; *Balesdent et al.*, 2018].

Radiocarbon dating has demonstrated that organic matter in soil can be very old (millennia or greater), particularly at depth. The persistence of SOC, i.e., the tendency of carbon to remain in

soil for long periods of time without being decomposed, is an area of active research [*Schmidt et al.*, 2011; *Riley et al.*, 2014; *Castellano et al.*, 2015; *Mathieu et al.*, 2015; *Sierra et al.*, 2018]. Previous hypotheses that microbes transform plant matter into chemically stable forms of SOC (humus) have been rejected due to evidence that soil carbon can become old while remaining in a biogeochemically available form *Campbell and Paustian* [2015].

The current understanding is that multiple environmental factors contribute to SOC persistence [*Schmidt et al.*, 2011]. When carbon enters the soil, much of it gets remineralized very quickly, while some of it is stabilized via physical protection in micro-aggregates and chemical protection by minerals [*Riley et al.*, 2014; *Paul*, 2016]. Observational evidence suggests that while climate, and particularly temperature, explains most of the variability in radiocarbon ages at the surface, persistence of SOC in deeper soil is more strongly controlled by mineral associations [*Koarashi et al.*, 2012; *Kramer et al.*, 2017], with the most important climatic driver being soil moisture [*Kramer and Chadwick*, 2018; *Balesdent et al.*, 2018].

4.1.3 Radiocarbon in Earth system models

While a small set of ESMs have recently incorporated radiocarbon as a tracer in simulated SOC [*Koven et al.*, 2013; *Tifafi et al.*, 2018a; *Chen et al.*, 2019], none did so in time to be included in the fifth phase of the Coupled Model Intercomparison Project (CMIP5) [*Taylor et al.*, 2011]. However, uncertainty in the fate of SOC in CMIP5 ESMs [*Todd-Brown et al.*, 2013] prompted *He et al.* [2016] to fit their output to a set of radiocarbon-enabled reduced complexity models (RSMs). *He et al.* [2016] compared these RSMs to a large set of radiocarbon observations from around the world, and found that the models tended to underestimate observed radiocarbon ages. They posited that the age bias was the result of carbon cycling through the soil too quickly, which they hypothesized was causing the models to overestimate the response of the soil carbon sink to increased atmospheric CO_2 concentrations. This hypothesis was supported by constraining the RSMs with the radiocarbon

observations, which reduced the 21st-century soil carbon sink relative to the unconstrained models [*He et al.*, 2016].

The first ESM to explicitly include terrestrial radiocarbon was the Community Earth System Model (CESM), by way of its land surface model, version 4.5 of the Community Land Model (CLM4.5) [*Koven et al.*, 2013]. Since then, the CESM developers have released version 5 of CLM (CLM5), and will include terrestrial radiocarbon in their submissions to the sixth phase of the CMIP experiment (CMIP6). In addition, the United States Department of Energy has been actively developing its own ESM, the Energy Exascale Earth System Model (E3SM). E3SM began as a branch of CESM that included the radiocarbon-enabled CLM4.5, and will also be including terrestrial radiocarbon in its submissions to CMIP6. Finally, Organising Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE), the land surface model for the Institut Pierre Simon Laplace (IPSL) ESM, now also includes radiocarbon [*Tifafi et al.*, 2018a], though it will not be included in submissions to CMIP6 [*Charles Koven, personal communication*].

The radiocarbon tracer was an important component in the development of vertically resolved soil decomposition in CLM4.5 [*Koven et al.*, 2013]. However, the developers had access to only a fairly limited number of soil radiocarbon observations. The time since then has seen multiple efforts to compile a synthesis of soil radiocarbon observations from around the globe [*Mathieu et al.*, 2015; *He et al.*, 2016], which have been combined into the International Soil Radiocarbon Database (ISRaD) [*Lawrence et al.*, 2019]. *Chen et al.* [2019] used the synthesis compiled from *He et al.* [2016] to evaluate version 1.0 of the E3SM Land Model (ELM1.0), and they found that while the model simulated radiocarbon ages fairly well at the surface, they strongly underestimated radiocarbon age at depth [*Chen et al.*, 2019]. However, it was later discovered that ELM1.0 contained a coding error that resulted in an unrealistically low rate of radioactive decay, which would have contributed to underestimation of the older soil layers. This error was subsequently fixed in ELM1.1.

Here, we used the released version 1.0 of ISRaD (Figure 4.1) to constrain the size of the soil carbon sink in ESMs. We evaluated vertically-resolved soil radiocarbon profiles in both CLM5 and ELM1.1.

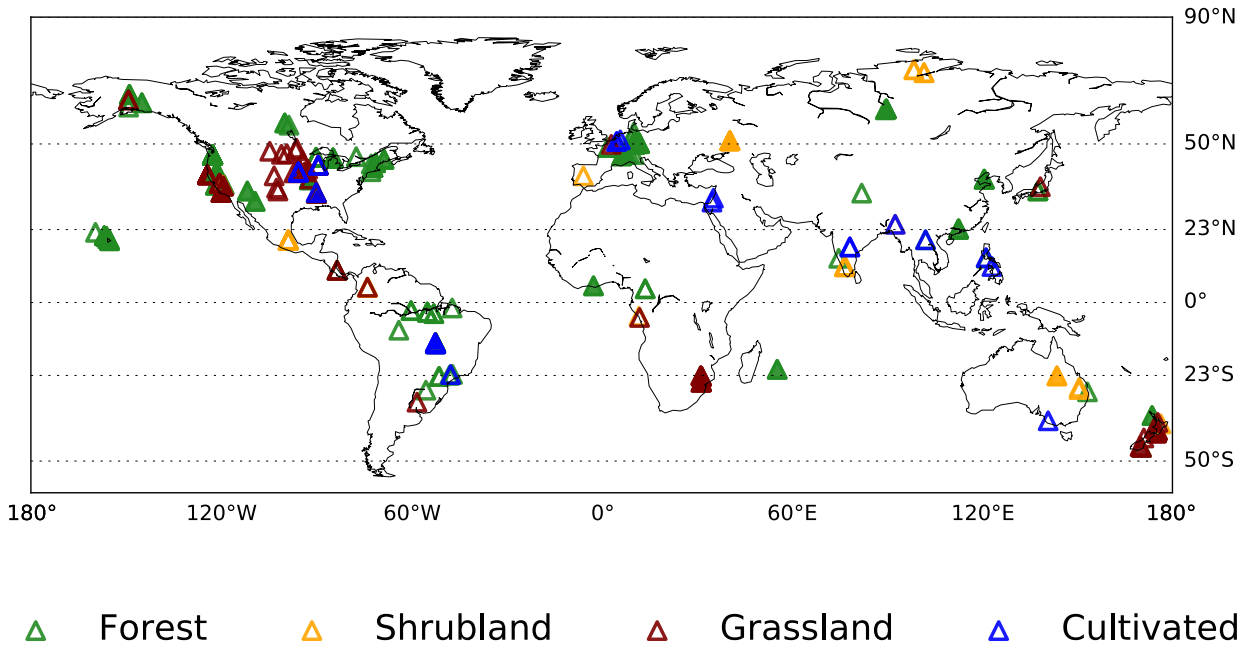

Figure 4.1: Location of observational radiocarbon profiles separated by land cover type. Filled symbols indicate locations that also include data for determining carbon stocks, which were used in the calibration of ELM1.1.

We then used the observations from ISRaD to calibrate some key soil decomposition parameters in ELM1.1. We describe the theory that formed the basis of the calibration and how it was implemented in ELM1.1 in the next section, followed by a detailed description of the calibration methods, a presentation of the results, and an analysis of their significance.

4.2 Methods

4.2.1 Theory and model implementation

The simulation of soil decomposition processes in both CLM5 and ELM1.1 was inherited from CLM4.5, which was based on the CENTURY model [*Parton et al.*, 1993]. CENTURY is a multi-compartment first-order decay model in which the compartments represent conceptual "pools" of carbon. Each pool is each parameterized with a kinetic rate constant that determines how much

carbon leaves via microbial decay at each time step, and a transfer coefficient that determines how much of the departing carbon gets shifted to downstream pools versus how much is lost to respiration. The transfer coefficient of each pool (referred to here as the respired fraction) can be subtracted from one to determine the inherent carbon use efficiency of each pool, i.e., the fraction of transferred carbon that is consumed via respiration.

The implementation of CENTURY in both CLM and ELM comprises three SOC pools, which are colloquially referred to as the "fast," "slow," and "passive" pools to reflect their characteristic time scales of months-to-years, decades, and centuries-to-millennia, respectively. Each of these three soil pools, along with three litter pools and a pool for coarse woody debris (CWD) that comprise the rest of the CENTURY-derived decomposition cascade, are represented in each of the 10 biologically and hydrologically active soil layers, for a total of 30 individual soil pools. Diffusive transport allows vertical exchange of carbon between the same pool types in different layers. Within each layer, the parameterized rate constants for each pool are modified by scalars derived from temperature and soil moisture, reflecting the influence of these environmental factors on decay rates.

Despite their relative simplicity, multi-compartment approaches to modeling soil decomposition have been known for some time to reproduce observed radiocarbon ages [*Jenkinson*, 1990]. However, they are limited by the conceptual nature of the distinctions between pools. While there have been attempts to relate individual pools in multi-compartment models to measurable soil fractions [*Skjemstad et al.*, 2004; *Zimmermann et al.*, 2007], detailed radiocarbon measurements have indicated that even the fractions of soil carbon are mixtures of fast and slow carbon [*Schrumpf and Kaiser*, 2015].

In our analysis, we use output from both models based on their "offline" mode, in which the land model was forced with atmospheric boundary conditions derived from observations and reanalysis instead of being coupled with an interactive atmospheric circulation model. If offline atmospheric forcing is fixed to a relatively short climatology (one or a few years), the climatological means of the temperature and moisture scalars become fixed, and therefore so do the decay rates in each

pool. Furthermore, the rates of input from primary production via the upstream litter and CWD pools remain constant bewtween cycles of the climatology, meaning that the model behaves like an autonomous compartmental system as described in *Rasmussen et al.* [2016]. In this case, the quotient of the fluxes into each pool and the decay rates determine the "steady state" stocks, which would be quasi-equilibrated to the specific conditions of the (quasi-constant) climatology.

In a transient simulation, in which the atmospheric forcing is not fixed, the system is no longer autonomous, and the stocks would be expected to change in response to changing inputs and changing environmental scalars. However, if a transient simulation begins in a state that is not equilibrated to the initial climatological conditions, there will be drift in the stocks that would confound the signal of interest. Because stocks in ELM get initialized arbitrarily, we were required to "spin up" the model by fixing atmospheric forcing to a climatology that resembles the first several years of the transient simulation, and integrating it forward in time long enough for the stocks to equilibrate to those conditions.

After spinning the model up to sufficiently equilibrate the stocks, changes in the mass of the stocks can be interpreted as the response to the transient forcing, both of atmospheric CO_2 concentrations climatic conditions. The response of the stocks depends on the rates at which carbon cycles through the individual pools, which itself is determined by the rate constants and respired fractions used to parameterize the soil pools. We demonstrated the dependence of stock responses to these model parameters using an idealized conceptual model (Figure 4.2).

The initial goal of this research was to modify the model in a way that would make radiocarbon ages older at depth while maintaining the well-simulated SOC stocks. Using our idealized conceptual model, we demonstrated how simultaneously changing the decay rates and respired fractions can allow a model to simulate the same SOC stocks with older radiocarbon ages. In this demonstration, models A and B are idealized two-compartment models, each with a fast and slow carbon pool that are parameterized with a turnover time (τ) that is the multiplicative inverse of the decay rate constant (Figure 4.2a).

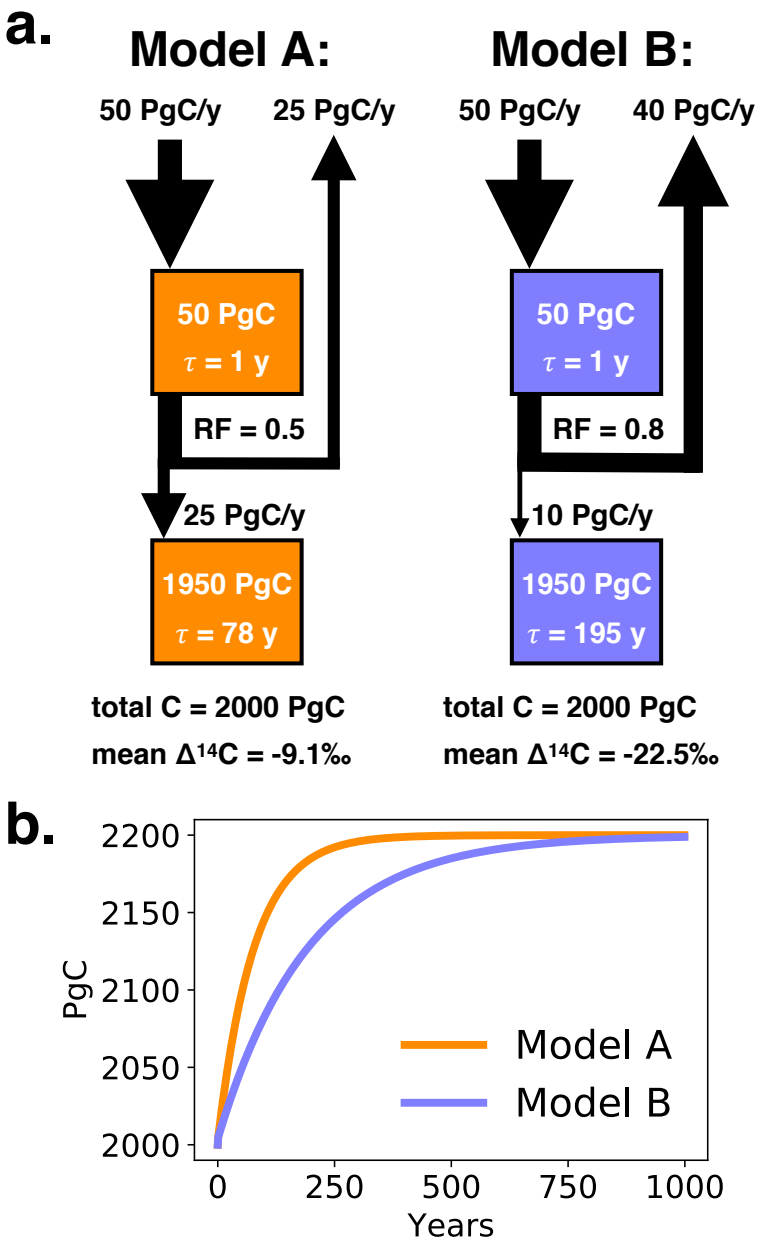

Figure 4.2: Conceptual demonstration of calibration strategy. a) Two idealized models with the same inputs and steady state stocks. b) Both models eventually equilibrate to the same steady state stock, but the slower model (B) takes longer to reach this state.

Both models have the same steady-state stock of 2000 PgC and the same total input of 50 PgC per year, with differences in the transfer coefficients (RF) between the fast and slow pools affecting the inputs to the slow pools. The decay rate of the slow pool in Model B is slower (longer turnover time) than that of Model A, leading to an older radiocarbon age for the system as a whole. The different parameterizations caused the two models to respond differently to an instantaneous increase in the total input of 50 PgC per year (Figure 4.2b). While both models eventually equilibrated to the same SOC stocks in response to increased input, the slower model (with the older radiocarbon age) responds more slowly, thereby providing a smaller carbon sink in the short term.

4.2.2 Evaluation of ESM radiocarbon

ISRaD is a compilation of soil radiocarbon measurements that have been collected from published literature and submitted by community members. ISRaD includes radiocarbon measurements from bulk soil, soil fractions, soil interstitial gas and liquid, and soil–atmosphere fluxes. For the present study, we used bulk soil observations from mineral soil layers, as this provides the best analog for comparison with an ESM.

We compared the full set of observed radiocarbon profiles against the output from CLM5 and ELM1.1. The simulated profiles from CLM5 were biased young at the surface in some biomes, and were biased young at depth throughout the globe (Figure 4.3). Although ELM1.1 was generally less biased than CLM, and was able to reproduce the radiocarbon profile fairly well for temperate woody vegetation (trees and shrubs), it still had a young age bias at depth in the tropical and boreal latitudes (Figure 4.4).

In addition to examining profiles at the site level, we evaluated the global mass-age distribution of carbon against a benchmark constructed from two different machine-learning products. We obtained SOC density and soil bulk density from LandGIS (`https://landgis.opengeohub.org`), a successor to the SoilGrids [*Hengl et al.*, 2017], that globally extrapolates an extensive set of soil

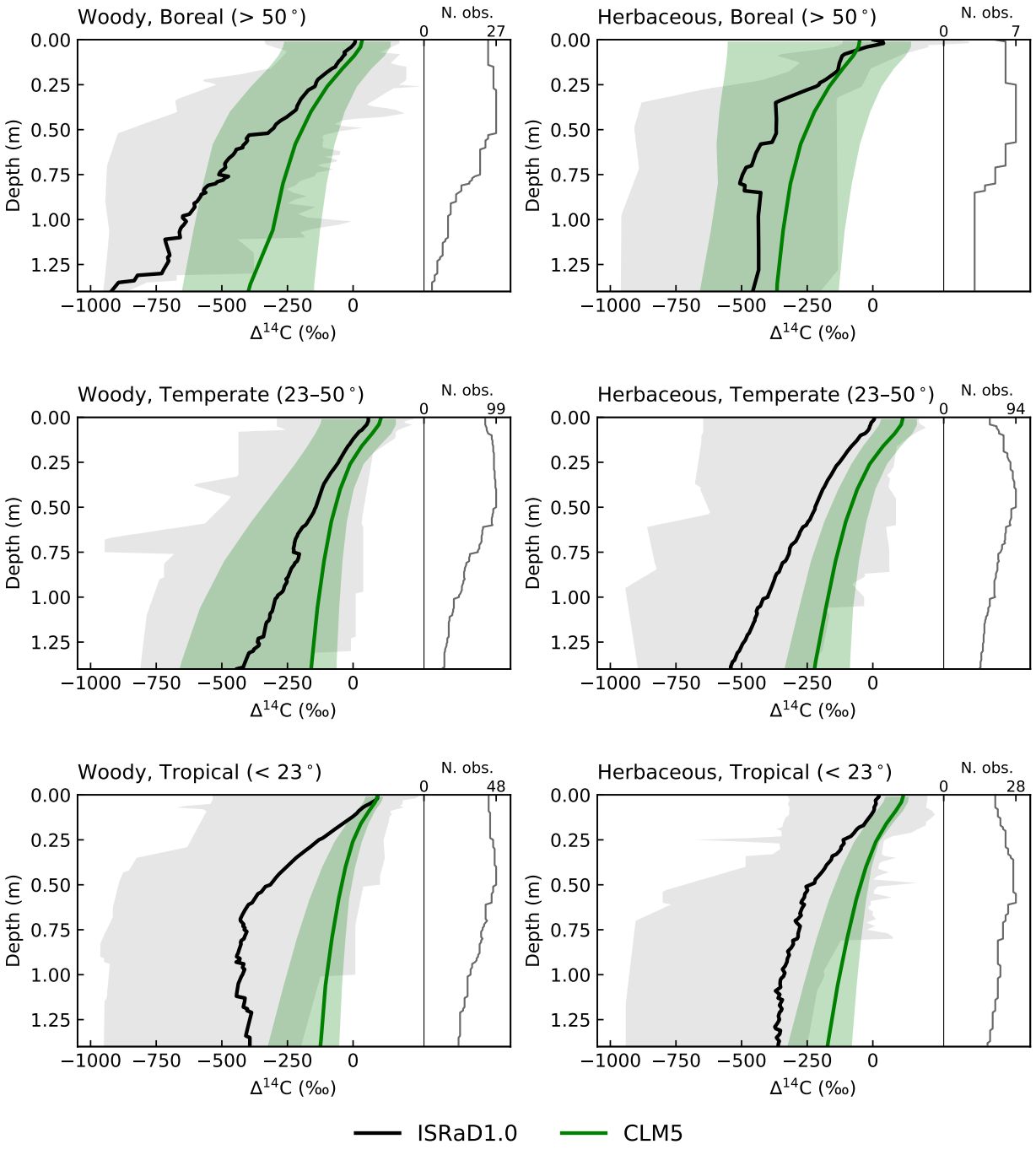

Figure 4.3: Comparison of observed radiocarbon profiles with CLM5. Shaded regions indicate the minimum and maximum values. Woody vegetation consists of all forest and shrubland, and herbaceous vegetation consists of grassland and cropland.

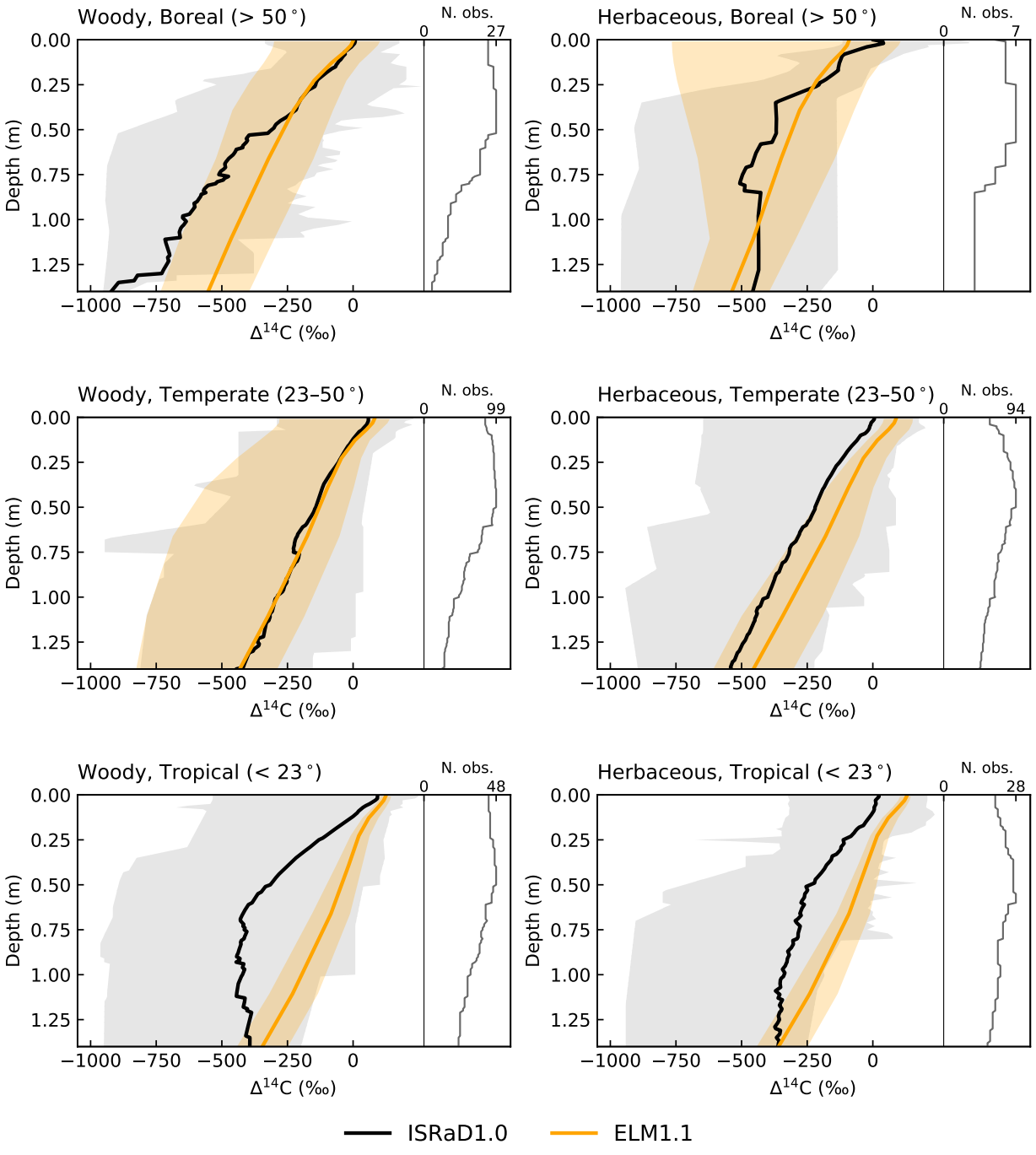

Figure 4.4: Comparison of observed radiocarbon profiles with ELM1.1. Shaded regions indicate the minimum and maximum values. Woody vegetation consists of all forest and shrubland, and herbaceous vegetation consists of grassland and cropland.

observations using random forest regression. An additional random forest regression was run to globally extrapolate radiocarbon ages at two depth intervals (0–30 cm, and 30–100 cm) based on environmental factors, mainly temperature, precipitation, and soil texture [*Shi et al.*, 2019]. We used the combination of the two machine-learning products to determine the distribution of SOC by age within the two depth intervals, and compared this with the equivalent distribution from CLM5 and ELM1.1 (Figure 4.5). This global comparison indicated that SOC in both models is younger than in the benchmark data, consistent with the site-level comparisons.

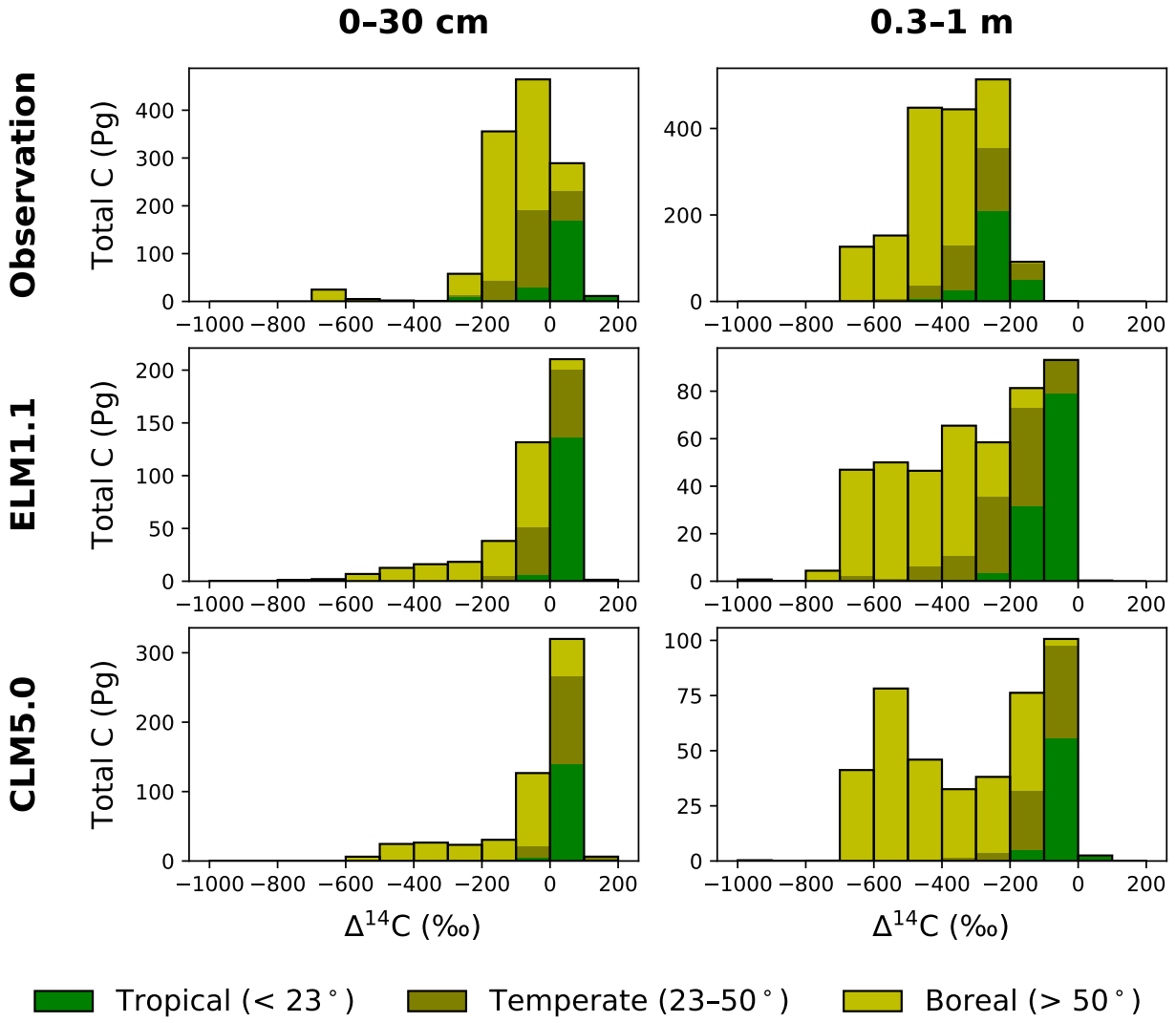

Figure 4.5: Mass-age distribution of soil carbon comparing an observational benchmark product (LandGIS and *Shi et al.* [2019]) with CLM5 and ELM1.1.

4.2.3 Radiocarbon-constrained ESM calibration

Data selection

The importance of multiple constraints [*Luo et al.*, 2016] prompted us to perform the ESM calibration using only the observations that include enough information to also determine the amount of carbon in the soil. In some cases, SOC stocks were reported directly, but in most cases they were derived from measurements of soil bulk density and organic carbon concentration. We included these data as objectives in our calibration in order to insure that we did not improve model representation of radiocarbon ages at the expense of being able to reproduce SOC stocks.

For the purpose of the calibration, we extracted all layers from ISRaD that contained both SOC stocks (or the information required to calculate them) and radiocarbon ages for bulk soil layers. Layers were screened to eliminate those with a bulk density less than 0.5 $g\,m^{-3}$ or organic carbon concentrations above 20%, as we considered those to be more characteristic of organic rather than mineral soil horizons. We also eliminated layers with a bulk density greater than 2 $g\,m^{-3}$, as we considered those to exceed the normal density of mineral soil and likely to yield unrealistically high SOC stocks.

In addition, several observational profiles included layers across which radiocarbon ages got younger with depth. While there are several processes that could lead to such conditions, such as vertical mixing (particularly cryoturbation) or disturbance (e.g., tilling, fire, deforestation), the model does not simulate these processes in a manner that would allow it to reproduce such profiles. We determined that including such conditions degraded the performance of the calibration algorithm, therefore we eliminated all layers for which the radiocarbon age in $\Delta^{14}C$ exceeded that of the layer above by more than 100‰. This threshold was chosen to balance the number of observations with which to constrain the model with the extent to which larger differences caused greater performance degradation.

We defined valid profiles as those that comprised three or more valid layers, and that extended to a depth of at least 50 cm. The final set of valid profiles for calibration comprises a total of 102 sites, with more than half of the profiles coming from forested land cover, and nearly two thirds of the profiles located in temperate latitudes (Figure 4.1). We performed a site-level comparison against ELM1.1 using only the 102 selected sites, and determined that the general patterns evident in Figure 4.4 are maintained with the reduced data set (Figure 4.6). We additionally compared SOC stocks from ELM1.1 with the observations, and found that the model tended to underestimate the amount of carbon across all depths in the tropical and temperate latitudes (Figure 4.7).

Calibration parameters

Based on the theory presented in Section 4.2.1, we developed a method to calibrate both the decay rates and the respired fractions simultaneously. Instead of calibrating each parameter for each pool separately, which would risk over-fitting the data, we created ad-hoc meta-parameters that scaled each pool by the same amount. We designated the parameter $k*$ to scale decay constants and $rf*$ to scale the respired fractions for all soil pools.

In addition to the environmental scalars that limit decomposition based on temperature and soil moisture, ELM1.1 also includes an exponential decrease of decay rates with depth that was inherited from CLM4.5 [*Koven et al.*, 2013]. This decrease is designed to account for processes that are not explicitly represented in the model, and is parameterized with an e-folding depth (referred to as z_τ) that has a default value of 0.5 m. The shape of radiocarbon profiles in ELM1.1 is known to be sensitive to z_τ [*Chen et al.*, 2019], so we included this parameter in our calibration as well.

Optimization algorithm

We conducted a global offline transient simulation on a 1.9°x2.5° grid, using atmospheric forcing data from the Global Soil Wetness Project (GSWP3) [*Dirmeyer et al.*, 2006b] for the years 1901 through

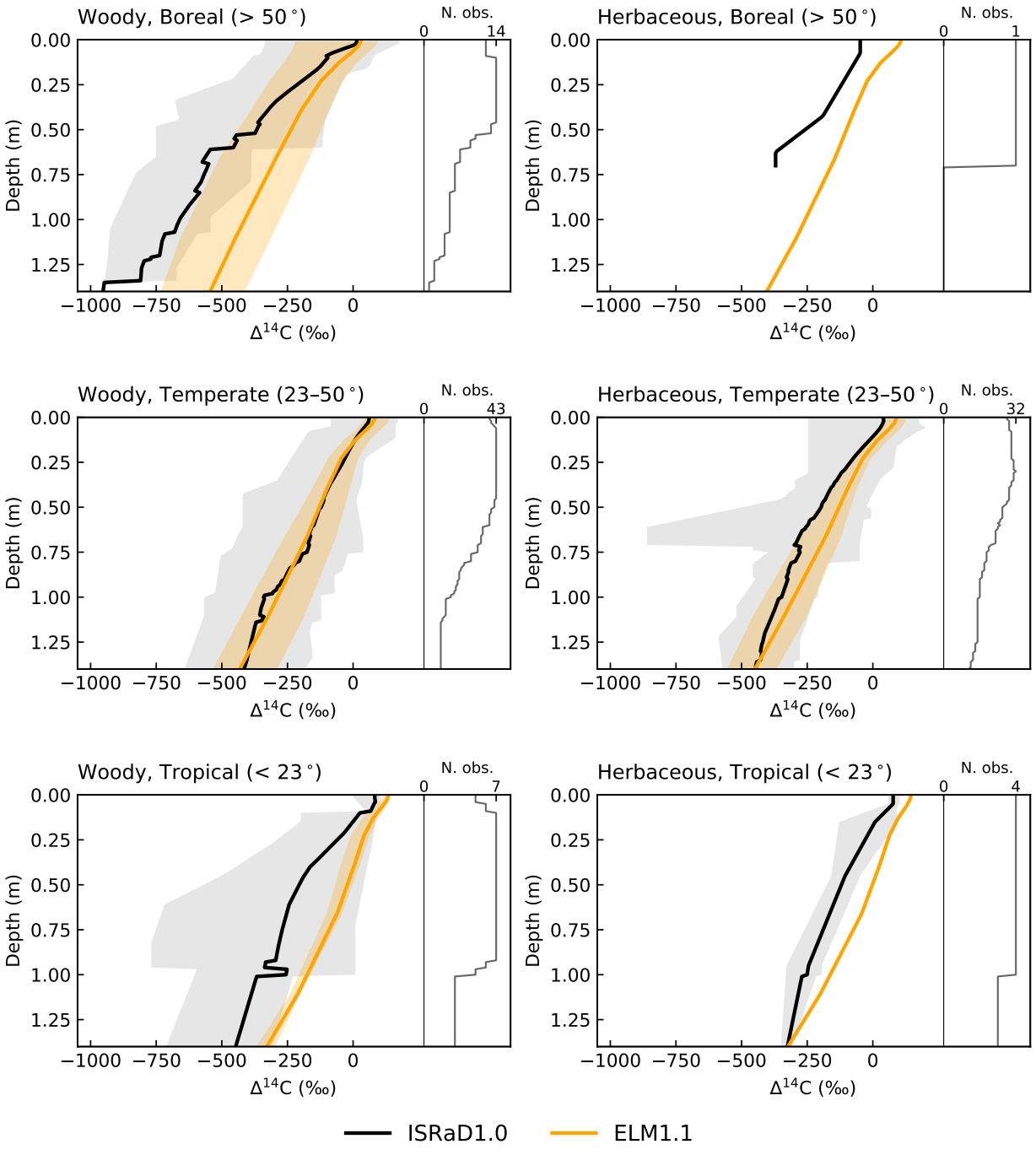

Figure 4.6: Comparison of observed radiocarbon profiles selected for calibration with ELM1.1. Shaded regions indicate the minimum and maximum values. Woody vegetation consists of all forest and shrubland, and herbaceous vegetation consists of grassland and cropland.

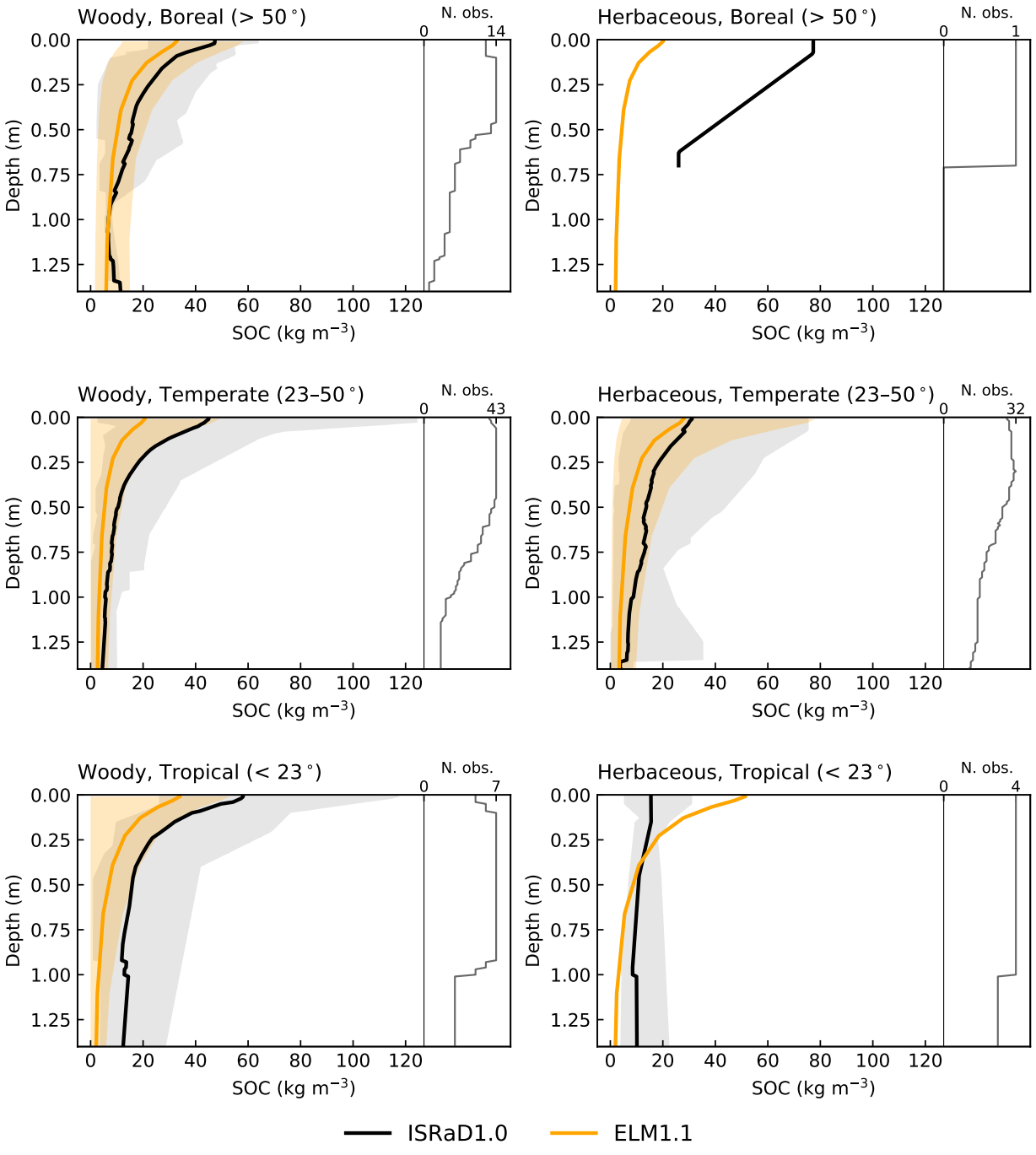

Figure 4.7: Comparison of observed SOC profiles selected for calibration with ELM1.1. Shaded regions indicate the minimum and maximum values. Woody vegetation consists of all forest and shrubland, and herbaceous vegetation consists of grassland and cropland.

2010. We configured ELM to use the Equilibrium Chemistry Approximation (ECA) approach [*Zhu et al.*, 2017] for nitrogen and phosphorous biogeochemistry, and used historical records of atmospheric CO_2 and $^{14}CO_2$. We used the first 20 years of atmospheric data to spin up the model until the total global soil SOC stock was changing by less than 0.1% annually. However, individual locations and layers were not always as close to equilibrium. To account for this, we conducted an additional control simulation, which amounted to continuing to spin up for an additional 110 years, and subtracted changes to each pool in the control from the transient simulation. Because we did not modify any parameter values from the ELM defaults, we refer to this as the default simulation.

Because the parameters we chose to calibrate determine the equilibrium stocks, all parameter combinations required undergoing the full spin up. This precluded any sort of iterative optimization algorithm, such as Markov chain Monte Carlo (MCMC) or evolutionary approaches. Instead, the model was run simultaneously across a range of seven values for each parameter (Table 4.1) for a total of 343 (7^3) combinations.

Table 4.1: Minimum and maximum values for each parameter, the increment it was sampled along, its default value in the model, and its description.

	Min.	Max.	Inc.	Def.	Description (and units)
$k*$	0.7	1.3	0.1	1*	Scaling factor for soil pool decay rate (unitless)
$rf*$	0.7	1.3	0.1	1*	Scaling factor for soil pool respired fraction (unitless)
z_τ	0.2	0.8	0.1	0.5	Decay rate decrease e-folding depth (m)

*$k*$ and $rf*$ are not part of the original model, but have no effect when set to 1.

For each of the 102 profiles in the study, we determined the plant functional type (PFT) from ELM that most closely matched the observed vegetation and land cover. We included all model grid cells that coincide with the observed profiles, and assigned each grid cell to the PFT associated with the observations. If two or more observational profiles with different vegetation types were located on the same model grid cell, we included that grid cell multiple times, once for each vegetation type. The total number of unique grid cells in the calibration simulation was 72, and each was simulated for all 343 parameter combinations, for a total of 24,696 grid cells.

In order to facilitate comparison between observations and model output at different layer depths, we linearly interpolated each profile to a common 1-cm vertical resolution. We chose 1.4 m as the maximum depth for the optimization, as the number of observations below that depth were too few for a robust comparison. If $C_o(i, p, z)$ is the observed carbon content (in g m^{-3}) in each of the 150 1-cm vertical layers (z) for all 343 parameter combinations (p) at each of the 102 profiles (i), and $C_m(i, p, z)$ is the modeled carbon content, the root mean squared error for the carbon stock each profile and parameter combination ($RMSE_C(i, p)$) was calculated as

$$RMSE_C(i, p) = \sqrt{\frac{\sum_{z=0}^{150} \left([C_m(i, p, z) - C_o(i, p, z)]^2 \right)}{150}} \tag{4.1}$$

For the observed and modeled radiocarbon ages ($\Delta^{14}C_o(i, p, z)$ and $\Delta^{14}C_m(i, p, z)$, respectively, in units of ‰), the root mean squared error ($RMSE_{\Delta^{14}C}(i, p)$) was weighted by the the observed carbon content and calculated as

$$RMSE_{\Delta^{14}C}(i, p) = \sqrt{\frac{\sum_{z=0}^{150} \left(C_o(i, p, z) * \left[\Delta^{14}C_m(i, p, z) - \Delta^{14}C_o(i, p, z) \right]^2 \right)}{\sum_{z=0}^{150} C_o(i, p, z)}} \tag{4.2}$$

Because each latitude band had a unique combination of stock and age biases, we conducted a separate calibration for each. For each latitude band, we considered woody vegetation types only (trees and shrubs), due to the different biases between these and herbaceous (grass and crop) vegetation types and the relatively small number of valid profiles available for the latter (Figures 4.6 and 4.7). The modeled radiocarbon profiles for woody land cover in the temperate latitudes matched the observations fairly well (Figures 4.4 and 4.6), therefore, we excluded this region while performing separate calibrations for the tropical and boreal latitude bands.

For each of the calibrated latitude bands, a total RMSE was calculated for each parameter combination

as a weighted mean of the N sites in that region, where the weights were the inverse of the number of profiles sharing the same model grid cell as profile i ($S(i)^{-1}$). For example, the region-wide RMSE of carbon content ($\mathrm{RMSE}_{\mathrm{C}}(p)$) was calculated as

$$\mathrm{RMSE}_{\mathrm{C}}(p) = \frac{\sum_{i=0}^{N} \left(S(i)^{-1} * \mathrm{RMSE}_{\mathrm{C}}(i,p) \right)}{\sum_{i=0}^{N} S(i)^{-1}} \tag{4.3}$$

and the region-wide RMSE for radiocarbon ages ($\mathrm{RMSE}_{\Delta^{14}\mathrm{C}}(p)$) was calculated equivalently. Weighting by $S(i)^{-1}$ insured that each model grid cell was weighted equally, and was numerically equivalent to averaging all observed profiles that share a common grid cell.

Because the RMSEs are in different units (mass of carbon for SOC stocks, and Δ^{14}C for radiocarbon ages), we normalized each of the two global objective functions by the standard deviation of RMSEs across all parameter combinations ($\sigma_{\mathrm{RMSE}_{\mathrm{C}}}$ and $\sigma_{\mathrm{RMSE}_{\Delta^{14}\mathrm{C}}}$ for carbon content and ages, respectively) for all 343 parameter combinations. The two normalized RMSEs were then summed to create a combined objective function, and in each latitude band, the parameter combination with the smallest combined objective function was designated as the optimum. Mathematically, the optimum parameter combination (p_{opt}) was found as

$$p_{opt} = \arg\min \left(\frac{\mathrm{RMSE}_{\mathrm{C}}(p)}{\sigma_{\mathrm{RMSE}_{\mathrm{C}}}} + \frac{\mathrm{RMSE}_{\Delta^{14}\mathrm{C}}(p)}{\sigma_{\mathrm{RMSE}_{\Delta^{14}\mathrm{C}}}} \right) \tag{4.4}$$

for each of the latitude bands in our calibration.

The two sets of p_{opt} for boreal and tropical latitudes were used as the basis of a constrained simulation. We assigned all model grid cells with more than 50% of the vegetated area assigned to boreal forest PFTs to the boreal p_{opt}, and likewise for the tropical forest PFTs and the tropical p_{opt}. We note that a small number of model grid cells in the temperate latitudes (23–50°) were modified, by virtue of containing boreal and tropical forest PFTs near the highest and lowest latitudes, respectively. The constrained simulation repeated the same offline spinup and transient 20^{th}-century climate as the

default simulation, with the calibrated model parameters.

4.3 Results

4.3.1 Optimum model parameters

In our optimization, neither of the individual latitude band optimizations yielded a single parameter value that minimized the RMSE for both SOC stocks and radiocarbon ages. As is the case with any multi-objective optimization, we needed to decide how to balance the trade-offs between the the two objective functions. Within each latitude band, we considered the Pareto frontier, which is the set of parameter values in which improvements to either objective function can be made only at the expense of the other (Figure 4.8). The member of the Pareto frontier selected as the optimum was determined by minimizing the combined objective function, which was the sum of normalized RMSE values as described in Equation 4.4.

Examination of the Pareto frontier indicated that a wide range of possible parameter combinations can yield a fairly narrow range of combined objective functions. This highlights the issue of equifinality, in which multiple solutions (combinations of parameters) were able produce results that may be considered equally desirable. It also demonstrates the interdependence between parameters, as alternative values in one parameter could be close to each other in objective space as a result of differences in the other parameters.

The $k*$ parameter in particular was not well-constrained by the observational data, as evidenced by the wide range of values within the Pareto frontier for this parameter across both latitude bands. This was especially noticeable in the boreal latitudes, in which the optimum combination had a $k*$ value of 1.1, but alternatives with values of 0.7 and 1.2 had nearly the same combined objective function. While this parameter is constrained somewhat better in the tropical latitudes, it still varied

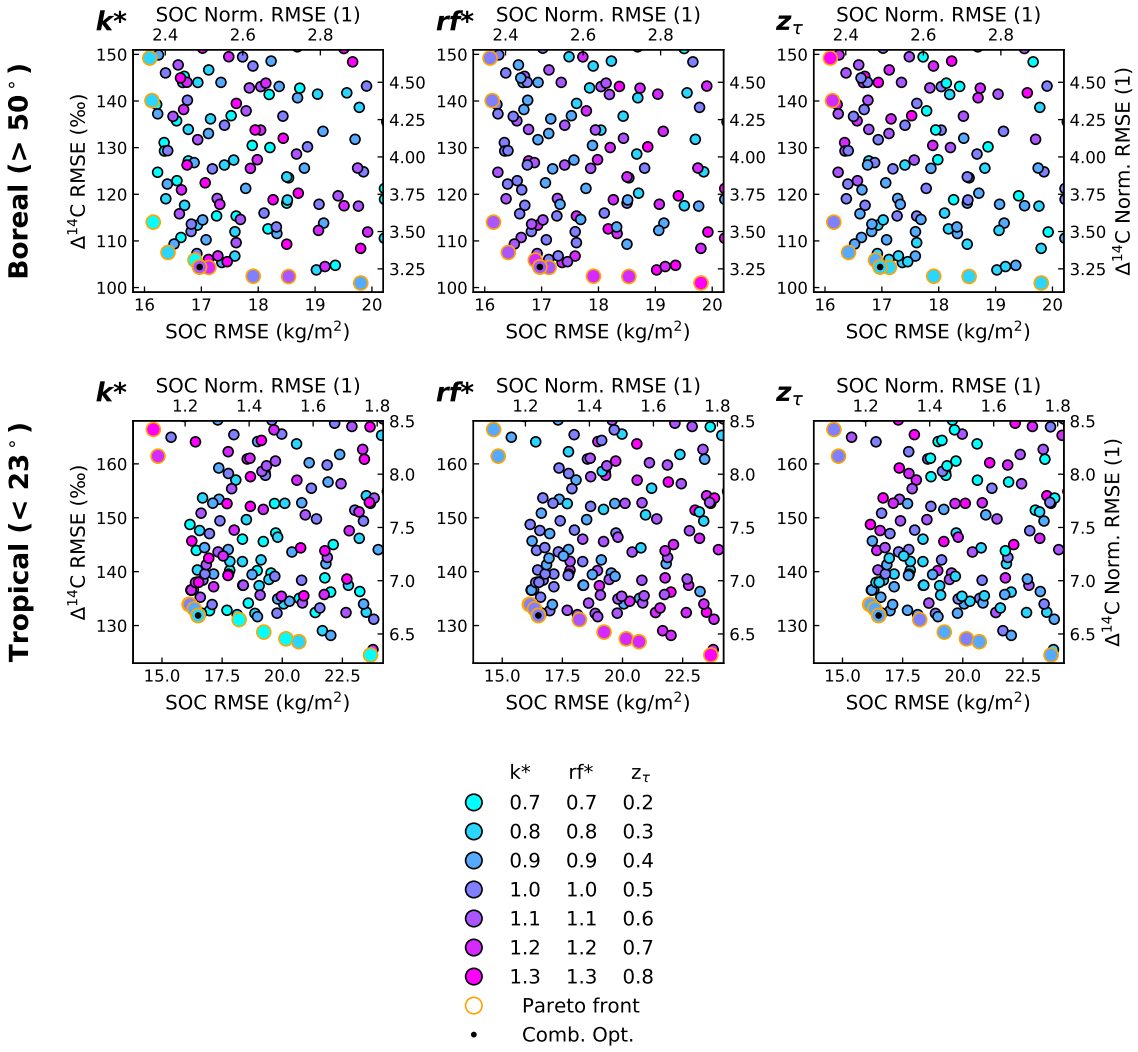

Figure 4.8: Parameter values in objective space. The horizontal axes are the RMSE for carbon stocks, in absolute units on the bottom and normalized units on the top. The vertical axes are the RMSE for radiocarbon ages, in absolute units on the left and normalized units on the right. Each point indicates the RMSEs for one of the 343 parameter combinations, with color coded values for each parameter in the columns. Each row is for a single latitude band. Points that are on the Pareto front are outlined in gold, and the optimum value (minimum of the sum of the normalized RMSEs) is indicated with a black dot.

more across the Pareto frontier than either of the other two parameters.

The interdependence between parameters was also evident upon examination of the combined objective function in parameter space (Figure 4.9). Within both latitude bands, the $k*$ and $rf*$

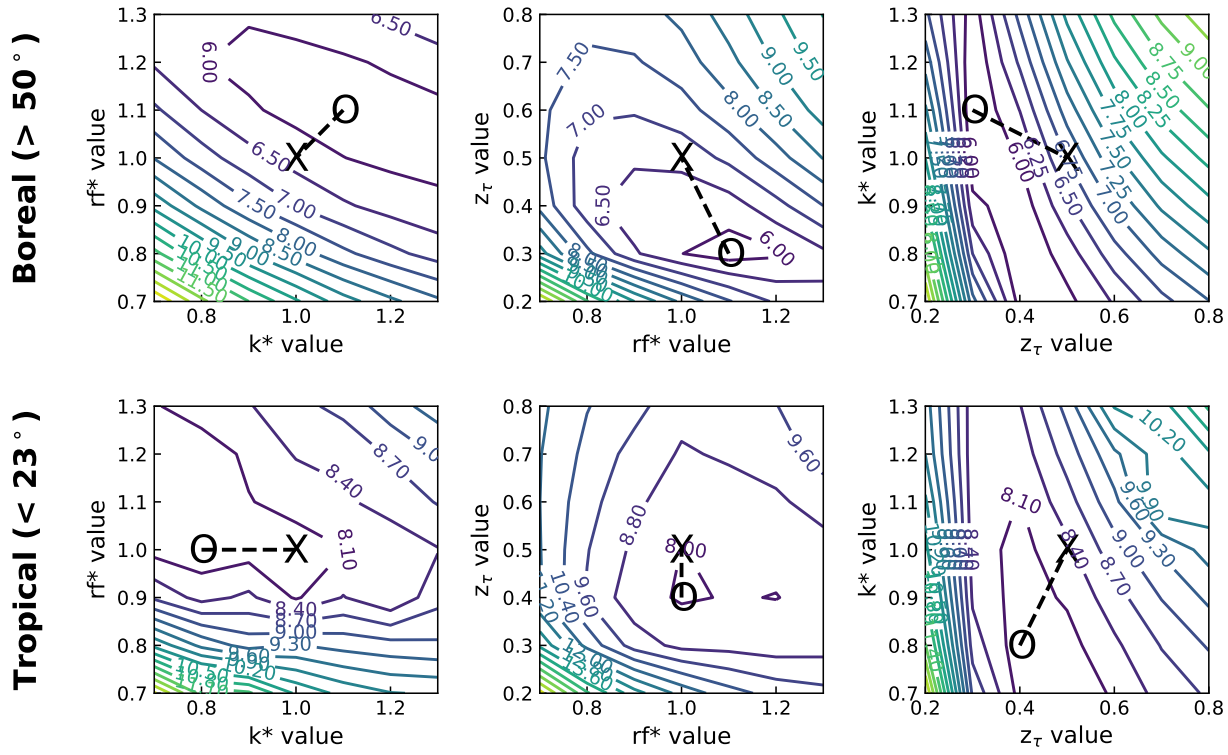

Figure 4.9: Combined objective function in parameter space. The contour lines are the sum of the normalized RMSEs for SOC stocks and radiocarbon. The X in the center of each denotes the model's default value, and the O denotes the minimum combined RMSE value.

parameters were negatively correlated, as the combination of high $rf*$ and low $k*$ values produced the having the same combined objective function as low $rf*$ and high $k*$. The $k*$ and z_τ parameters were also negatively correlated, which was to be expected, considering that each of these parameters affected decomposition rates with the same sign, though the former affected the entire soil column equally while the latter had a larger effect at depth.

4.3.2 Constrained model output

The optimized parameter values yielded a noticeable improvement to the simulated soil profiles (Figure 4.10). Carbon stocks in the constrained simulation remained close to those in the default simulation in the boreal latitudes, but improved in the tropics by simulating increased SOC at all depths. Radiocarbon ages showed improvements–with ages increasing (Δ^{14}C decreasing) at depth–for both latitude bands.

The constrained simulation also yielded improvements to the global mass-age distributions compared to the machine-learning benchmark data (Figure 4.11). The constrained simulation was slightly older at 0–30 cm, and substantially older at the 30–100 cm depth interval, compared with the default simulation. However, the constrained simulation was still younger than the benchmark product in both depth ranges, indicating the need for continued improvement of the optimization process.

4.3.3 The twentieth-century soil carbon sink

Constraining ELM1.1 with the observations from ISRaD led to changes in both absolute SOC stocks and the change in those stocks in response to 20th century climatic and anthropogenic forcing (Table 4.2). The constrained model began its transient simulation with a larger SOC stock than the default model, reflecting the higher equilibrium values of the optimized parameters achieved during spin-up. In 1910, SOC represented 74% and 75% of the total terrestrial carbon stock in the default and constrained models, respectively.

The total 20$^{\text{th}}$-century soil carbon sink in the constrained model was 0.3 PgC lower than the default model. This represented both a smaller portion of the 1910 value and a smaller portion of the total terrestrial carbon sink. In the constrained simulation, all pools other than vegetation began the 20^{th} century with higher carbon stocks, and all pools other than soil increased slightly more, compared with the default simulation.

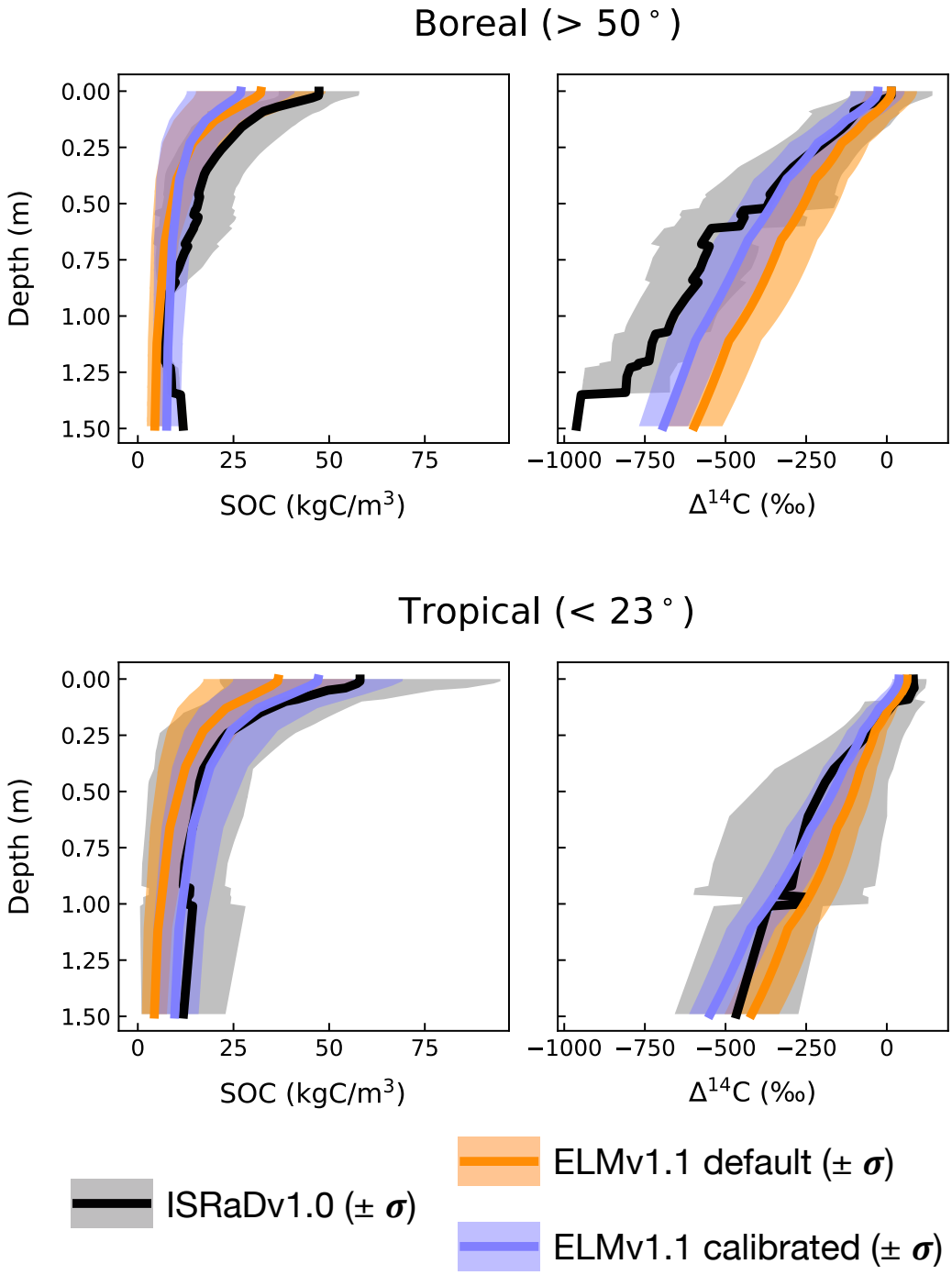

Figure 4.10: Observed profiles with default and constrained ELMv1.1 output. Observed profiles are the same as the "Woody" columns of Figures 4.6 and 4.7. Modeled profiles are from the calibration run, in which the grid cells were assigned to 100% of the observed PFT. Shaded regions are $\pm 1\sigma$.

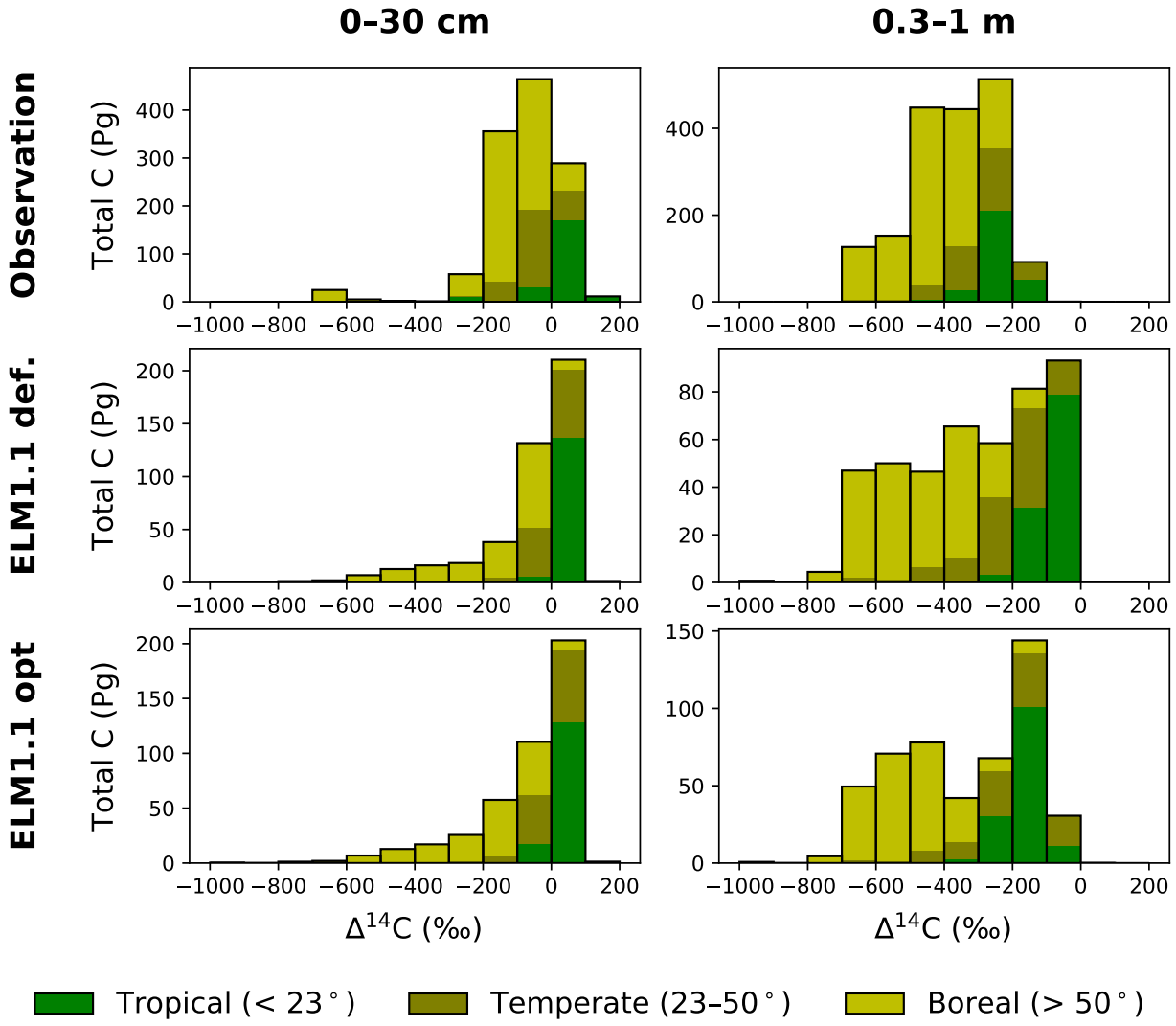

Figure 4.11: Mass-age distribution of soil carbon comparing an observational benchmark product (LandGIS and *Shi et al.* [2019]) with the default and constrained simulations from ELM1.1.

In both the default and constrained simulations, the boreal latitude band contained more than half of the global SOC stock, but 20th-century changes were very small in magnitude (nearly 0% of the global sink, and less than 2% of the beginning-of-the-century stock). The temperate latitudes remain nearly unchanged between the simulations, owing to the very small number of model grid cells in which more than 50% of the vegetated surface consists of either boreal or tropical forest PFTs. The tropical latitudes saw the largest change in the 20th-century sink, with the constrained model simulating a smaller sink due to a slower soil carbon cycle.

Table 4.2: Changes in carbon stocks between 1910 and 2010. 1910 and Δ2010 columns are in units of PgC, and are the 1910 stocks and the change to those stocks in 2010, respectively. %Δ columns express Δ2010 as a percent of the 1910 stock. %Σ columns express Δ2010 as a percent of the Δ2010 for all rows in the grouping, which is the total column Δ2010 for the first grouping, and the total soil Δ2010 for the second and third groupings.

	ELM1.1 default				ELM1.1 constrained			
Total C:	1910	Δ2010	%Δ	% Σ	1910	Δ2010	%Δ	% Σ
Vegetation	356.1	51.2	14.4	57.4	352.9	51.9	14.7	55.5
CWD	227.0	19.6	8.6	22.0	280.5	22.9	8.2	24.5
Litter	29.8	2.2	7.4	2.5	33.9	2.7	8.1	2.9
Soil	1755.8	16.2	0.9	18.1	2010.7	15.9	0.8	17.0
Soil C:	1910	Δ2010	%Δ	% Σ	1910	Δ2010	%Δ	% Σ
> 50°	906.8	-0.3	-0.0	-1.6	1009.1	-0.3	-0.0	-1.8
23–50°	438.1	6.5	1.5	39.9	420.4	6.5	1.5	40.9
< 23°	410.9	10.0	2.4	61.6	581.1	9.7	1.7	60.9
Soil C:	1910	Δ2010	%Δ	% Σ	1910	Δ2010	%Δ	% Σ
< 30 cm	476.1	13.4	2.8	82.5	474.0	13.1	2.8	82.2
30–100 cm	446.4	2.8	0.6	17.1	486.6	2.8	0.6	17.8
> 1 m	833.3	0.1	0.0	0.8	1050.1	0.0	0.0	0.0

In the default simulation, almost half of the global SOC stock at the beginning of the 20th century were below 1 m deep, and was more than half in the constrained simulation. While the default simulation had a very small 20th-century sink at this depth (0.1 PgC), that of the constrained simulation was even smaller. In both simulations, more than 82% of the 20th-century sink was shallower than 30 cm.

The simulated response to 20th-century climate change (embedded in the GSWP3 forcing data) and to increased atmospheric CO_2 concentrations exhibited two distinct phases in the time series (Figure 4.12). In the first half of the century, there was little accumulation in any of the pools, but the second half of the century brought relatively rapid increases across most pools. Notable exceptions were the litter pool, which never sustained meaningful accumulation due to rapid decay rates, and the boreal soil, which actually began the century accumulating carbon but ended up losing it to the atmosphere in the latter half of the century. As evident in Table 4.2, the 20th-century soil carbon

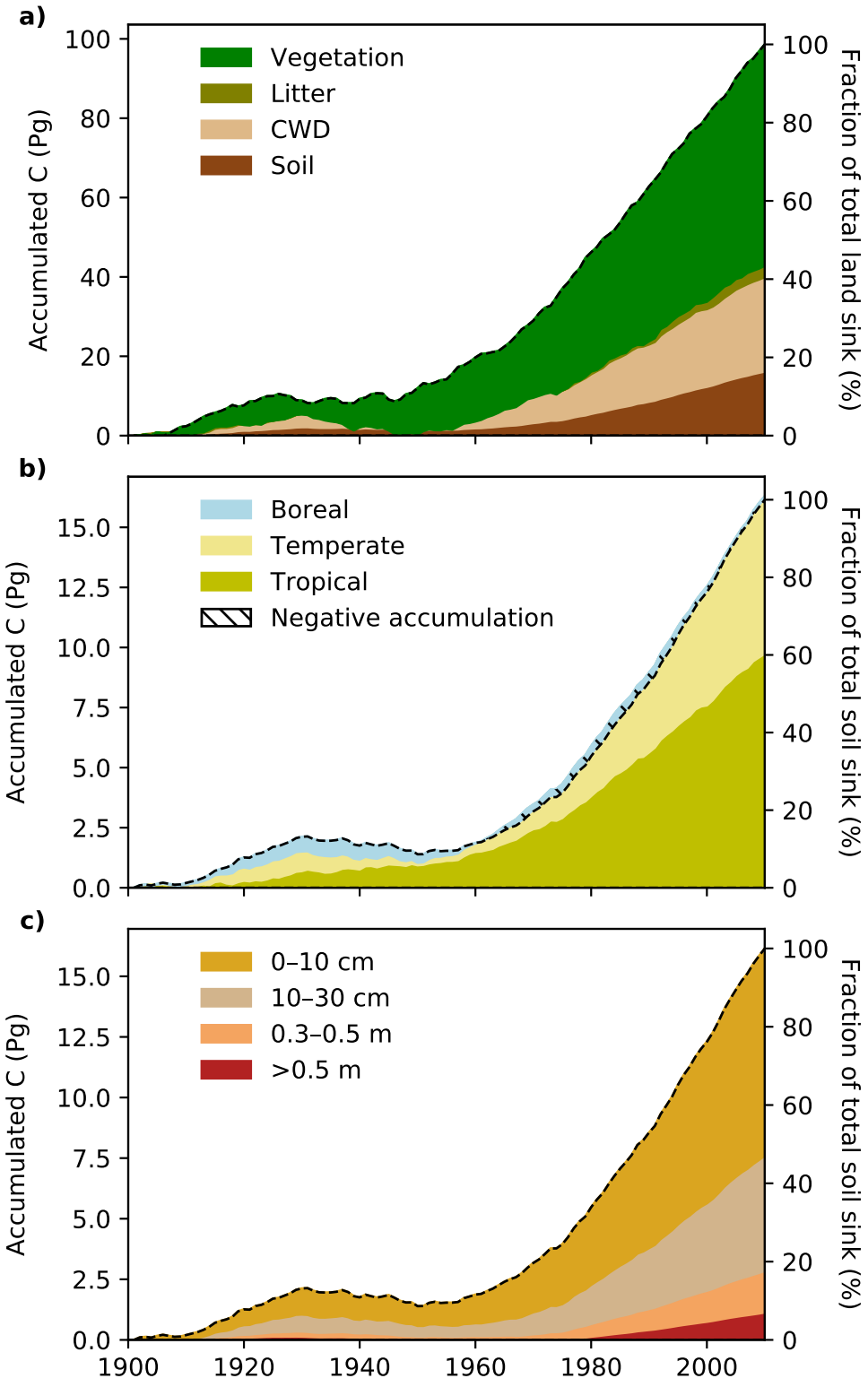

Figure 4.12: Constrained 20th-century terrestrial carbon sink. The response to the constrained transient simulation across terrestrial carbon pools (a) and in the soil pool across latitude bands (b) and depth intervals (c).

sink was a relatively small component of the total terrestrial sink, and it took place primarily in the tropics and primarily in the shallowest layers.

4.4 Discussion

4.4.1 Implications for the future soil carbon sink

Both land surface models we examined here had soil carbon ages that were too young relative to the observations from ISRaD, consistent with previous research [*He et al.*, 2016; *Chen et al.*, 2019]. This suggests that the models were cycling carbon too quickly, which would cause them to overestimate how rapidly it could sequester carbon in response to elevated atmospheric CO_2 concentrations. The constrained simulation, in which decomposition rates were reduced in the boreal and tropical latitudes along with inputs to slower pools (by virtue of larger respired fractions), supports this by simulating a smaller 20^{th}-century soil carbon sink.

Even the default ELM1.1 simulates radiocarbon profiles better than CLM5, and the difference between the default and constrained simulations are less than the difference between models. The carbon sink in CLM5 could not be directly compared with the ELM simulations due to the inclusion of land cover change, which drove a large loss of soil carbon. Both of these models are descendants of earlier versions of CLM, which was known for its rapid soil and litter decay rates, being the only land surface model in the CMIP5 ensemble (in CESM1-BGC and NorESM1) to simulate a terrestrial carbon source [*Anav et al.*, 2013; *Friedlingstein et al.*, 2014]. But it was also the land surface model associated with the smallest bias and smallest 21^{st}-century soil carbon sink in the reduced complexity models from *He et al.* [2016]. This encourages us to interpret the generally small soil sink in both ELM1.1 simulations as being well constrained by the observations.

This interpretation is in agreement with the current understanding of soil carbon age distribution with

depth. A large portion of carbon is remineralized soon after entering the soil, while a smaller portion is protected physically through micro-aggregates and chemically through mineral associations, and transported in solution to depth where it remains for centuries to millennia [*Schmidt et al.*, 2011; *Riley et al.*, 2014; *Paul*, 2016]. While ELM1.1 does include vertical transport via diffusion, it does not account for the mineral sorption and microbial interactions that strongly influence soil age structure [*Ahrens et al.*, 2015; *Dwivedi et al.*, 2017], the only land surface model for an ESM that accounts for these processes being the Geophysical Fluid Dynamics Laboratory's Land Model 4 (LM4) when using the Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment (CORPSE) module [*Sulman et al.*, 2014]. The z_τ parameter was designed to account for all of these un-simulated processes in CLM [*Koven et al.*, 2013], and allows ELM1.1 to accurately reproduce radiocarbon ages, especially when constrained by the calibration.

These results to not inspire optimism in the efficacy of soil to serve as a solution to anthropogenic CO_2 emissions. Even though atmospheric CO_2 will enhance carbon input to the soil through primary production [*Cox et al.*, 2013; *Keenan et al.*, 2016], this will not necessarily lead to more storage. Uptake capacity depends on saturation of the physical and chemical substrates that protect soil carbon and allow it to persist, and priming effects could even result in increased input enhancing decomposition [*Castellano et al.*, 2015]. Increased inputs would also have to overcome enhanced heterotrophic respiration driven by global warming [*Heimann and Reichstein*, 2008; *Crowther et al.*, 2016; *Melillo et al.*, 2017], as well as soil carbon loss due to land cover change [*Sanderman et al.*, 2017]. On the other hand, this could potentially be mitigated if enhanced decomposition has sufficient nutrient availability to increase the conversion of carbon to protected forms [*Tang and Riley*, 2014; *Crowther et al.*, 2015]. On the whole, there is a strong case to be made that soil can only serve as a useful carbon sink if managed explicitly to do so [*Stockmann et al.*, 2013; *Minasny et al.*, 2017].

4.4.2 Limitations, uncertainties, and future improvements

Despite the improved simulation using the constrained parameters, and the increased confidence in its representation of soil carbon uptake, there are several limitations that highlight the need for refinement of the calibration process. The primary limitations are imposed by the observational data set, but others stem from the model itself as well as the optimization methodology. These limitations should motivate future efforts toward reducing the uncertainty they impose.

While the value of ISRaD is unprecedented for examining radiocarbon ages in soil, as was demonstrated in this study, it is limited in its ability to provide a high quality and quantity of information regarding SOC stocks. A majority of ISRaD profiles we included (those extending below 50 cm depth) did not include sufficient information to determine SOC stocks, which limited their use in our multiple-constraint approach. The small number of SOC observations exacerbated the issues introduced by observational uncertainty, stemming mostly from measurements of soil bulk density [*Todd-Brown et al.*, 2013; *Tifafi et al.*, 2018b].

An additional limitation inherent to ISRaD is the reduction in the number of observations with increasing depth (Figures 4.3–4.7). While we limited our optimization to the top 1.5 m, we found the calibration algorithm to be fairly insensitive to the maximum depth beyond ~75 cm, where the number of observations declines precipitously, due mostly to the higher weighting of the higher mass layers at the surface. However, both the models and the benchmark data from LandGIS include substantial amounts of soil carbon below 1 m, which reflects uncertainty in global soil carbon stocks [*Batjes*, 2014; *Jackson et al.*, 2017; *Tifafi et al.*, 2018b] and the maximum depth to which SOC extends [*Harper and Tibbett*, 2013; *Jackson et al.*, 2017; *Balesdent et al.*, 2018].

Further uncertainties are introduced by the mismatch in spatial scales between the model and observations. SOC stocks and radiocarbon ages are highly spatially variable [*Schrumpf and Kaiser*, 2015], and many of the observations in ISRaD were collected to maximize this variability [*Lawrence et al.*, 2019]. This limitation could be overcome by field campaigns with a sampling strategy

designed to obtain observations that are more representative of larger landscapes. While some drivers of variability are included in ELM1.1, specifically the temperature and moisture scalars, the model does not account for geologic variability (information about soil texture and content are included indirectly through their use in the soil moisture and soil temperature portions of the model), nor can it account for spatial heterogeneity in microbial communities [*Don et al.*, 2017]. This highlights the importance of models that represent processes at observable scales [*Riley et al.*, 2014].

Finally, our calibration protocol introduced additional uncertainty in the selection of the optimum parameter combination. For the purpose of performing a constrained transient simulation, we were satisfied with obtaining a single set of values for each of the three major latitude bands. The variability within the Pareto frontiers demonstrated the sensitivity of the constrained parameters to the normalization of the objective functions. Assigning a value that determines the relative importance of each RMSE is an unavoidable step in a multi-objective optimization. Future work will consider sampling multiple solutions from the Pareto frontier, in order to produce a distribution of constrained models rather than a single realization.

4.5 Conclusion

We used global observations from ISRaD to evaluate and constrain the soil carbon cycle in ELM1.1. We found that while ELM1.1 slightly overestimated radiocarbon ages in temperate latitudes, it underestimated them in boreal and tropical latitudes, particularly at depth, enough to lead to a young soil carbon age bias globally. CLM5, which we also evaluated, had an even stronger young bias. This suggests that carbon was cycling through the soil too rapidly in the models, leading them to overestimate the rate at which soils could sequester carbon in response to increasing atmospheric CO_2 concentrations.

We used the observed radiocarbon profiles to constrain kinetic rate constants and transfer coefficients in ELM1.1, which improved the young age bias and improved a low bias in the tropical soil carbon stock. Constraining the model increased the global soil carbon stock as a result of the improvements in the tropics, and reduced the contribution of soil to the terrestrial carbon sink. This suggests that soil is likely to respond very slowly to increasing atmospheric CO_2 concentrations, and is unlikely to serve as a carbon sink in the near term. To the extent that anthropogenic CO_2 emissions may stimulate terrestrial primary production, the additional carbon will be stored in the vegetation and CWD, which are highly vulnerable to carbon loss and therefore much less effective as a long-term sink.

Chapter 5

Conclusions

5.1 Implications of research results

The research presented in this dissertation contributes toward the general scientific knowledge of land-atmosphere feedbacks in the Earth system, and demonstrates the importance of Earth system models (ESMs) as tools in this effort. I focused on two important feedbacks between the land surface and the atmosphere: soil moisture feedbacks via land–atmosphere coupling, and the long-term portion of the carbon-concentration feedback. I demonstrated how ESMs can be evaluated for and improved in their representation of land-atmosphere feedbacks, and how they can be manipulated to investigate mechanisms that cannot be directly observed in nature.

In Chapter 2, I established that ESMs from the fifth phase of the Coupled Model Intercomparison Project (CMIP5) tend to overestimate the strength of feedbacks between soil moisture and the atmosphere. This is probably due to overly simplistic "big-leaf" representations of the land surface [e.g., *Dickinson et al.*, 1998] with insufficient representation of bare soil evaporation, leading to a land surface that is too sensitive to soil moisture variability. By demonstrating that models overestimate both the forcing and response limbs, I highlighted how warm/dry biases in ESMs

[*Cheruy et al.*, 2014; *Mueller and Seneviratne*, 2014; *Wehrli et al.*, 2018] could be exacerbated, because even though the forcing limb is relatively small, the large response limb sets up even larger biases for subsequent forcing. These results emphasize the extent to which uncertainty in projections of the hydrologic cycle can degrade the simulation of temperature trends under climate change. For example, while *Berg et al.* [2016b] found that soil moisture feedbacks exacerbate temperature trends in CMIP5 models, their results are applicable only to the extent that the projected soil moisture increases are accurate.

In Chapter 3, I demonstrated the importance of this feedback in the response of the Amazon Rainforest to climate variability associated with the El Niño-Southern Oscillation (ENSO). I showed that while the immediate response to ENSO is driven directly by atmospheric circulation, soil moisture feedbacks to temperature amplify this response and extend it by several months, particularly in the eastern Amazon. The results of Chapter 2 suggest that this effect may be unrealistically strong in the ESM I used. However, it would also be expected to be uncrealistically strong across most ESMs, which only further emphasizes the need to consider hydrologic variability and its relationship with temperature when attributing the contribution of these factors to the tropical terrestrial carbon cycle. This suggests that efforts to constrain the long-term carbon-climate feedback using the present-day relationship between temperature and CO_2 may need to be revisited.

In Chapter 4, I describe a method for constraining a radiocarbon-enabled ESM with data from the International Soil Radiocarbon Database (ISRaD). The constrained model reduces some of the bias in the modeled radiocarbon ages, and suggests that soil is unlikely to serve as a major sink for anthropogenic CO_2 emissions in the 21^{st} century. The methodology that I developed for this chapter can be applied to other ESMs with vertically resolved radiocarbon, and can be used to calibrate other parameters beyond those that I demonstrated. I plan to continue working with this method in order to refine the constraints on the Energy Exascale Earth System Model (E3SM) Land Model (ELM) and to impose similar constraints on the Community Land Model (CLM) prior to continuing with the next steps described below in Section 5.2.

5.2 Future research directions

There are two general categories of future research that would extend and/or complement the work presented here. The first category is more methodological, and involves formalizing the techniques I developed in Chapters 2 and 4 so that they can be applied more broadly toward benchmarking ESMs. The second category is more scientific, and involves pursuing the questions from Chapter 3 with additional data, and extending those questions to shed light on the carbon-climate feedback [*Cox et al.*, 2013].

5.2.1 Formalizing new methods

For the first category, I plan to implement my methods from Chapters 2 and 4 within the framework of the International Land Model Benchmarking (ILAMB) system [*Collier et al.*, 2018]. ILAMB provides a platform for systematically benchmarking ESMs against a wide variety of observational data at both the global and site level. In addition to assessing the mean state and variability of individual variables, ILAMB was specifically designed to include functional response benchmarks [*Randerson et al.*, 2009] that evaluate the relationship between multiple variables. The metrics I developed for Chapter 2 would serve to complement existing functional response benchmarks in ILAMB.

When ILAMB was originally designed, few if any ESMs included vertically resolved soil profiles, so there is currently no functionality for benchmarking soil characteristics as a function of depth. However, it has become increasingly common for ESMs to simulate vertically resolved soil processes [*Koven et al.*, 2013; *Camino-Serrano et al.*, 2018]. If added to ILAMB, the methods I developed in Chapter 4 would enable the benchmarking of any vertically resolved soil characteristics, including not only carbon stocks and radiocarbon ages as I compared in Chapter 4, but any other observations at multiple depths, such as soil moisture and temperature.

5.2.2 Further investigating terrestrial carbon cycle response to ENSO

The total terrestrial carbon cycle feedback is often conceptualized as the sum of two separate feedback processes: the carbon-concentration feedback, which I considered in Chapter 4, and the carbon-climate feedback. The carbon-climate feedback comprises the sum of various terrestrial responses to climate change that can complete both positive and negative feedback loops. In general, higher temperatures from global warming would promote an increased CO_2 flux to the atmosphere from both respiration [*Jones et al.*, 2011] and fire [*Flannigan et al.*, 2013], but could also enhance CO_2 uptake via gross primary production (GPP) in temperature-limited environments [*Winkler et al.*, 2019].

Cox et al. [2013] developed a method for reducing uncertainty in the long-term carbon-climate feedback based on the "emergent constraint" approach. An emergent constraint [*Hall and Qu*, 2006] is a linear relationships within an ensemble of models between some unknown feature, typically either a future state or a long-term sensitivity, and some other feature that can be validated with observational evidence. The emergent relationship enables the observationally derived probability distribution of the ensemble to constrain the probability distribution of the unknown feature. This approach has been used to constrain the equilibrium climate sensitivity [*Huber et al.*, 2011], the snow-albedo feedback [*Qu and Hall*, 2014], cloud feedbacks [*Klein and Hall*, 2015], and atmospheric CO_2 projections [*Hoffman et al.*, 2014].

The emergent constraint used by *Cox et al.* [2013] was the relationship between the present-day sensitivity of global atmospheric CO_2 to the interannual variability of tropical temperature ($\gamma_{\,\mathrm{IAV}}$) and the long-term carbon-climate sensitivity ($\gamma_{\,\mathrm{LT}}$). By calculating a value for $\gamma_{\,\mathrm{IAV}}$ based on observations, they applied the emergent constraint to reduce the uncertainty in the ensemble's probability distribution for $\gamma_{\,\mathrm{LT}}$. Based on their constrained $\gamma_{\,\mathrm{LT}}$ for the Coupled Climate Carbon Cycle Model Intercomparison Project (C^4MIP) ensemble [*Friedlingstein et al.*, 2006], they determined the carbon-climate feedback is likely to reduce net terrestrial uptake as the climate

warms, but is unlikely to cause massive dieback of tropical forests, as some of the more extreme models projected. *Wenzel et al.* [2014] subsequently formalized the methodology and applied it to ESMs from the CMIP5 experiment, demonstrating its applicability across ensembles.

The utility of this emergent constraint approach is rooted in the relationship between the interannual variabilities of global atmospheric CO_2 and tropical temperature. Global CO_2 concentrations are strongly correlated with the ENSO, and the terrestrial carbon cycle has been identified as the dominant driver using both isotopic observations [*Rayner et al.*, 2008] and process-based models [*Jones et al.*, 2001]. The tropical land surface exhibits a relatively consistent response to ENSO, with warm and dry El Niño conditions reducing NEP by limiting GPP while enhancing respiration and fire, and the opposite effects under cool and moist La Niña conditions [*Hashimoto et al.*, 2004; *Qian et al.*, 2008]. The large difference between ENSO phases within a fairly short time frame (three to seven years) enables a robust calculation of γ_{IAV} even with a limited observational record.

The original methodology [*Cox et al.*, 2013; *Wenzel et al.*, 2014] calculated γ_{IAV} and γ_{LT} in reference to temperature, neglecting other climate variables under the assumption that they linearly scale with temperature. However, this assumption has been challenged by evidence that the temperature sensitivity of the terrestrial ecosystem is modulated by drought stress [*Wang et al.*, 2014]. While correlations between CO_2 and precipitation are low in observations and overestimated in models [*Piao et al.*, 2013; *Wang et al.*, 2013], stronger relationships emerge when CO_2 lags precipitation by several months [*Wang et al.*, 2016a], indicating the importance of soil moisture memory, which can also feed back to temperature. An active debate persists in the literature regarding the importance of hydrology relative to temperature as a driver of CO_2 interannual variability [*Jung et al.*, 2017; *Bastos et al.*, 2018], and served as part of the motivation for the research in Chapter 3.

One of the challenges to accounting for hydrology is that the lags vary spatially, with a time-evolving cascade of responses to ENSO drivers across the globe and, particularly, in the tropics [*Chen et al.*, 2017; *Levine et al.*, 2019]. This spatial heterogeneity is obscured by a single global time series of observational CO_2, as in the original studies [*Cox et al.*, 2013; *Wenzel et al.*, 2014]. The

use of a single global time series was due to spatially sparse observations of atmospheric CO_2 concentrations. However, the Greenhouse Observing Satellite (GOSAT) and Orbiting Carbon Observatory-2 (OCO-2) missions now provide spatially explicit observations of column-averaged atmospheric CO_2 concentrations [*Kataoka et al.*, 2017]. In addition, these satellites, along with the Global Ozone Monitoring Experiment-2 (GOME-2), provide measurements of gross primary production (GPP) constrained by solar induced fluorescence (SIF) [*Sun et al.*, 2018]. The CMS-Flux framework assimilates these and other data sources into an inverse atmospheric transport model to provide observationally constrained estimates of the net land-atmosphere CO_2 flux (net biome exchange, NBE), as well as the constituent gross fluxes (GPP, respiration, and fire) [*Liu et al.*, 2014; *Parazoo et al.*, 2014].

Despite the relatively short time series of available data, *Liu et al.* [2017] were able to use the CMS-Flux framework to measure the response of the tropical land surface to the 2015-2016 El Niño event. They showed that the response of the terrestrial land surface was far from homogeneous, and that CO_2 fluxes varied highly across the tropical continents as the result of spatial variability in the component fluxes (GPP, respiration, and fire). However, comparison with the E3SMv0.3 simulation I conducted for Chapter 3 demonstrates that the model does not correctly capture the spatial variability of either the component or net fluxes across tropical continents (Figure 5.1).

The satellite observations from GOME-2, GOSAT, and OCO-2, combined with the CMS-Flux framework, are powerful sources of data for evaluating, understanding, and improving ESMs. The spatially resolved observations of atmospheric CO_2 concentrations could be very useful for testing the assumptions and limitations of emergent constraints based on globally integrated data, and reducing the uncertainty they introduce [*Keppel-Aleks et al.*, 2018]. SIF-constrained GPP observations and inverted fluxes will reveal mechanistic relationships, which will reveal not only whether the models are getting these right, but also whether they are doing so for the right reasons (Figure 5.1). When combined with additional satellite observations of radiation and the hydrologic cycle, these data can be used to observe the entire time-evolving cascade of drivers and responses in

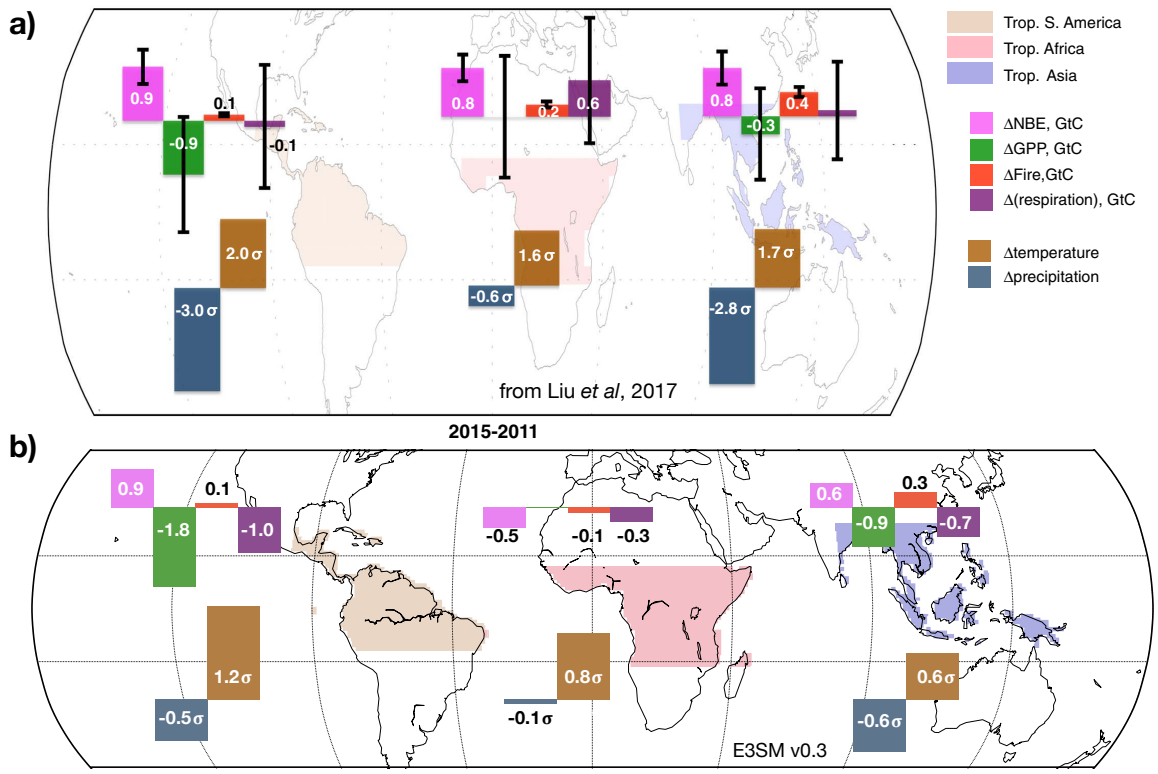

Figure 5.1: Observed and simulated tropical responses to the 2015 El Niño. Top (a) is Figure 2 from *Liu et al.* [2017]. Bars show the difference between 2015 and 2011 for CO_2 fluxes (PgC) and climate (relative to 1981–2010 σ). Bottom (b) is a reproduction of this figure using output from the Energy Exascale Earth System Model (E3SM) v0.3, forced with observed sea surface temperature. Even where the modeled NBE is accurate (S. America), there are large errors in both climate forcing and component flux response.

the tropical ENSO cycle [*Chen et al.*, 2017].

The spatial variability of functional relationships between environmental factors and land surface fluxes will provide new insights into the carbon-climate feedback. Uncertainties and assumptions in globally derived constraints on the long-term feedback can be critically examined and minimized to the extent possible. New functional response metrics that focus on the response of multiple land surface variables to ENSO can be incorporated into ILAMB, where they will reveal the conditions under which they can be used to constrain spatially resolved carbon-climate feedbacks, along with where ESMs need the most urgent improvements.

The atmospheric CO_2 observations will be of limited utility for the CMIP5-era models, as relatively

few include a time-varying CO_2 tracer that gets transported through the atmosphere, but will be quite valuable for assessing the newer generation of models from the sixth phase of the project (CMIP6). Many of the newer models are expected to have incorporated previously unrepresented information about the role of temperature and moisture in soil dynamics [*Bradford et al.*, 2016], which will require carefully evaluating functional relationships to determine where ESMs are falling short, and help model developers know where to focus there efforts. This research direction has the potential to not only help improve ESMs, but also to improve the general understanding of how the terrestrial carbon cycle responds to multiple drivers associated with ENSO.

Bibliography

Ahrens, B., M. C. Braakhekke, G. Guggenberger, M. Schrumpf, and M. Reichstein (2015), Contribution of sorption, DOC transport and microbial interactions to the ^{14}C age of a soil organic carbon profile: Insights from a calibrated process model, *Soil Biology and Biochemistry*, *88*, 390–402, doi:https://doi.org/10.1016/j.soilbio.2015.06.008.

Alfieri, L., P. Claps, P. D'Odorico, F. Laio, and T. M. Over (2008), An analysis of the soil moisture feedback on convective and stratiform precipitation, *Journal of Hydrometeorology*, *9*(2), 280–291, doi:10.1175/2007JHM863.1.

Anav, A., P. Friedlingstein, M. Kidston, L. Bopp, P. Ciais, P. Cox, C. Jones, M. Jung, R. Myneni, and Z. Zhu (2013), Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth system models, *Journal of Climate*, *26*(18), 6801–6843, doi:10.1175/JCLI-D-12-00417.1.

Arora, V. K., G. J. Boer, P. Friedlingstein, M. Eby, C. D. Jones, J. R. Christian, G. Bonan, L. Bopp, V. Brovkin, P. Cadule, T. Hajima, T. Ilyina, K. Lindsay, J. F. Tjiputra, and T. Wu (2013), Carbon–concentration and carbon–climate feedbacks in CMIP5 Earth system models, *Journal of Climate*, *26*(15), 5289–5314, doi:10.1175/JCLI-D-12-00494.1.

Bacastow, R. B. (1976), Modulation of atmospheric carbon dioxide by the Southern Oscillation, *Nature*, *261*, 116–118, doi:10.1038/261116a0.

Balesdent, J., I. Basile-Doelsch, J. Chadoeuf, S. Cornu, D. Derrien, Z. Fekiacova, and C. Hatté (2018), Atmosphere–soil carbon transfer as a function of soil depth, *Nature*, *559*(7715), 599–602, doi:10.1038/s41586-018-0328-3.

Banzon, V., T. M. Smith, T. M. Chin, C. Liu, and W. Hankins (2016), A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies, *Earth System Science Data*, *8*(1), 165–176, doi:10.5194/essd-8-165-2016.

Bastos, A., P. Friedlingstein, S. Sitch, C. Chen, A. Mialon, J.-P. Wigneron, V. K. Arora, P. R. Briggs, J. G. Canadell, P. Ciais, F. Chevallier, L. Cheng, C. Delire, V. Haverd, A. K. Jain, F. Joos, E. Kato, S. Lienert, D. Lombardozzi, J. R. Melton, R. Myneni, J. E. M. S. Nabel, J. Pongratz, B. Poulter, C. Rödenbeck, R. Séférian, H. Tian, C. van Eck, N. Viovy, N. Vuichard, A. P. Walker, A. Wiltshire, J. Yang, S. Zaehle, N. Zeng, and D. Zhu (2018), Impact of the 2015/2016 El Niño on the terrestrial carbon cycle constrained by bottom-up and top-down approaches, *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20170304, doi:10.1098/rstb.2017.0304.

Bates, J. R. (2007), Some considerations of the concept of climate feedback, *Quarterly Journal of the Royal Meteorological Society*, *133*(624), 545–560, doi:10.1002/qj.62.

Batjes, N. H. (2014), Total carbon and nitrogen in the soils of the world, *European Journal of Soil Science*, *65*(1), 10–21, doi:10.1111/ejss.12114_2.

Berg, A., B. R. Lintner, K. Findell, S. I. Seneviratne, B. van den Hurk, A. Ducharne, F. Chéruy, S. Hagemann, D. M. Lawrence, S. Malyshev, A. Meier, and P. Gentine (2015), Interannual coupling between summertime surface temperature and precipitation over land: Processes and implications for climate change, *Journal of Climate*, *28*(3), 1308–1328, doi:10.1175/JCLI-D-14-00324.1.

Berg, A., J. Sheffield, and P. C. D. Milly (2016a), Divergent surface and total soil moisture projections under global warming, *Geophysical Research Letters*, *44*(1), 226–234, doi:10.1002/2016GL071921, 2016GL071921.

Berg, A., K. Findell, B. Lintner, A. Giannini, S. I. Seneviratne, B. van den Hurk, R. Lorenz, A. Pitman, S. Hagemann, A. Meier, F. Cheruy, A. Ducharne, S. Malyshev, and P. C. D. Milly (2016b), Land-atmosphere feedbacks amplify aridity increase over land under global warming, *Nature Climate Change*, *6*(9), 869–874, doi:10.1038/nclimate3029.

Betts, A. K. (2009), Land-surface-atmosphere coupling in observations and models, *Journal of Advances in Modeling Earth Systems*, *1*(4), doi:10.3894/JAMES.2009.1.4.

Betts, A. K., R. Desjardins, D. Worth, and B. Beckage (2014), Climate coupling between temperature, humidity, precipitation, and cloud cover over the Canadian Prairies, *Journal of Geophysical Research: Atmospheres*, *119*(23), 13,305–13,326, doi:10.1002/2014JD022511.

Bodman, R. W., P. J. Rayner, and D. J. Karoly (2013), Uncertainty in temperature projections reduced using carbon cycle and climate observations, *Nature Climate Change*, *3*, 725–729, doi:10.1038/nclimate1903.

Bony, S., R. Colman, V. M. Kattsov, R. P. Allan, C. S. Bretherton, J.-L. Dufresne, A. Hall, S. Hallegatte, M. M. Holland, W. Ingram, D. A. Randall, B. J. Soden, G. Tselioudis, and M. J. Webb (2006), How well do we understand and evaluate climate change feedback processes?, *Journal of Climate*, *19*(15), 3445–3482, doi:10.1175/JCLI3819.1.

Booth, B. B. B., C. D. Jones, M. Collins, I. J. Totterdell, P. M. Cox, S. Sitch, C. Huntingford, R. A. Betts, G. R. Harris, and J. Lloyd (2012), High sensitivity of future global warming to land carbon cycle processes, *Environmental Research Letters*, *7*(2), 024002, doi:10.1088/1748-9326/7/2/024002.

Bowman, K. W., J. Liu, A. A. Bloom, N. C. Parazoo, M. Lee, Z. Jiang, D. Menemenlis, M. M. Gierach, G. J. Collatz, K. R. Gurney, and D. Wunch (2017), Global and Brazilian carbon response to El Niño Modoki 2011–2010, *Earth and Space Science*, *4*(10), 637–660, doi:10.1002/2016EA000204.

Bradford, M. A., W. R. Wieder, G. B. Bonan, N. Fierer, P. A. Raymond, and T. W. Crowther (2016), Managing uncertainty in soil carbon feedbacks to climate change, *Nature Climate Change*, *6*, 751–758, doi:10.1038/nclimate3071.

Camino-Serrano, M., B. Guenet, S. Luyssaert, P. Ciais, V. Bastrikov, B. De Vos, B. Gielen, G. Gleixner, A. Jornet-Puig, K. Kaiser, D. Kothawala, R. Lauerwald, J. Peñuelas, M. Schrumpf, S. Vicca, N. Vuichard, D. Walmsley, and I. A. Janssens (2018), ORCHIDEE-SOM: modeling soil organic carbon (SOC) and dissolved organic carbon (DOC) dynamics along vertical soil profiles in Europe, *Geoscientific Model Development*, *11*(3), 937–957, doi:10.5194/gmd-11-937-2018.

Campbell, E. E., and K. Paustian (2015), Current developments in soil organic matter modeling and the expansion of model applications: a review, *Environmental Research Letters*, *10*(12), 123,004.

Castellano, M. J., K. E. Mueller, D. C. Olk, J. E. Sawyer, and J. Six (2015), Integrating plant litter quality, soil organic matter stabilization, and the carbon saturation concept, *Global Change Biology*, *21*(9), 3200–3209, doi:10.1111/gcb.12982.

Chen, J., Q. Zhu, W. J. Riley, Y. He, J. T. Randerson, and S. Trumbore (2019), Comparison with global soil radiocarbon observations indicates needed carbon cycle improvements in the E3SM land model, *Journal of Geophysical Research: Biogeosciences*, *124*, 1098–1114, doi: 10.1029/2018JG004795.

Chen, Y., J. T. Randerson, D. C. Morton, R. S. DeFries, G. J. Collatz, P. S. Kasibhatla, L. Giglio, Y. Jin, and M. E. Marlier (2011), Forecasting fire season severity in South America using sea surface temperature anomalies, *Science*, *334*(6057), 787–791, doi:10.1126/science.1209472.

Chen, Y., I. Velicogna, J. S. Famiglietti, and J. T. Randerson (2013), Satellite observations of terrestrial water storage provide early warning information about drought and fire season severity in the Amazon, *Journal of Geophysical Research: Biogeosciences*, *118*(2), 495–504, doi:10.1002/jgrg.20046.

Chen, Y., D. C. Morton, N. Andela, L. Giglio, and J. T. Randerson (2016), How much global burned area can be forecast on seasonal time scales using sea surface temperatures?, *Environmental Research Letters*, *11*(4), 045001, doi:10.1088/1748-9326/11/4/045001.

Chen, Y., D. C. Morton, N. Andela, G. R. van der Werf, L. Giglio, and J. T. Randerson (2017), A pan-tropical cascade of fire driven by El Niño/Southern Oscillation, *Nature Climate Change*, *7*(12), 906–911, doi:10.1038/s41558-017-0014-8.

Cheruy, F., A. Campoy, J.-C. Dupont, A. Ducharne, F. Hourdin, M. Haeffelin, M. Chiriaco, and A. Idelkadi (2013), Combined influence of atmospheric physics and soil hydrology on the simulated meteorology at the SIRTA atmospheric observatory, *Climate Dynamics*, *40*(9), 2251–2269, doi:10.1007/s00382-012-1469-y.

Cheruy, F., J. L. Dufresne, F. Hourdin, and A. Ducharne (2014), Role of clouds and land-atmosphere coupling in midlatitude continental summer warm biases and climate change amplification in cmip5 simulations, *Geophysical Research Letters*, *41*(18), 6493–6500, doi:10.1002/2014GL061145.

Cohen, J., and M. Barlow (2005), The NAO, the AO, and global warming: How closely related?, *Journal of Climate*, *18*(21), 4498–4513, doi:10.1175/JCLI3530.1.

Collier, N., F. M. Hoffman, D. M. Lawrence, G. Keppel-Aleks, C. D. Koven, W. J. Riley, M. Mu, and J. T. Randerson (2018), The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation, *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731–2754, doi:10.1029/2018MS001354.

Cox, P. M., D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke (2013), Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability, *Nature*, *494*(7437), 341–344, doi:10.1038/nature11882.

Crowther, T. W., S. M. Thomas, D. S. Maynard, P. Baldrian, K. Covey, S. D. Frey, L. T. A. van Diepen, and M. A. Bradford (2015), Biotic interactions mediate soil microbial feedbacks to climate change, *Proceedings of the National Academy of Sciences*, *112*(22), 7033–7038, doi:10.1073/pnas.1502956112.

Crowther, T. W., K. E. O. Todd-Brown, C. W. Rowe, W. R. Wieder, J. C. Carey, M. B. Machmuller, B. L. Snoek, S. Fang, G. Zhou, S. D. Allison, J. M. Blair, S. D. Bridgham, A. J. Burton, Y. Carrillo, P. B. Reich, J. S. Clark, A. T. Classen, F. A. Dijkstra, B. Elberling, B. A. Emmett, M. Estiarte, S. D. Frey, J. Guo, J. Harte, L. Jiang, B. R. Johnson, G. Kröel-Dulay, K. S. Larsen, H. Laudon, J. M. Lavallee, Y. Luo, M. Lupascu, L. N. Ma, S. Marhan, A. Michelsen, J. Mohan, S. Niu, E. Pendall, J. Peñuelas, L. Pfeifer-Meister, C. Poll, S. Reinsch, L. L. Reynolds, I. K. Schmidt, S. Sistla, N. W. Sokol, P. H. Templer, K. K. Treseder, J. M. Welker, and M. A. Bradford (2016), Quantifying global soil carbon losses in response to warming, *Nature*, *540*, 104–108, doi:10.1038/nature20150.

Dai, A. (2006), Precipitation characteristics in eighteen coupled climate models, *Journal of Climate*, *19*(18), 4605–4630, doi:10.1175/JCLI3884.1.

Davidson, E. A., and I. A. Janssens (2006), Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, *Nature*, *440*(7081), 165–173, doi:10.1038/nature04514.

de Linage, C., H. Kim, J. S. Famiglietti, and J.-Y. Yu (2013), Impact of Pacific and Atlantic sea surface temperatures on interannual and decadal variations of GRACE land water storage in tropical South America, *Journal of Geophysical Research: Atmospheres*, *118*(19), 10,811–10,829, doi:10.1002/jgrd.50820.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597, doi:10.1002/qj.828.

Dennis, J. M., J. Edwards, K. J. Evans, O. Guba, P. H. Lauritzen, A. A. Mirin, A. St-Cyr, M. A. Taylor, and P. H. Worley (2012), CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model, *The International Journal of High Performance Computing Applications*, *26*(1), 74–89, doi:10.1177/1094342011428142.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive canopies for a climate model, *Journal of Climate*, *11*(11), 2823–2836, doi:10.1175/1520-0442(1998)011<2823:ICFACM>2.0.CO;2.

Dirmeyer, P. A. (2006), The hydrologic feedback pathway for land–climate coupling, *Journal of Hydrometeorology*, *7*(5), 857–867, doi:10.1175/JHM526.1.

Dirmeyer, P. A. (2011), The terrestrial segment of soil moisture–climate coupling, *Geophysical Research Letters*, *38*(16), L16702, doi:10.1029/2011GL048268.

Dirmeyer, P. A., R. D. Koster, and Z. Guo (2006a), Do global models properly represent the feedback between land and atmosphere?, *Journal of Hydrometeorology*, *7*(6), 1177–1198, doi:10.1175/JHM532.1.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006b), GSWP-2: Multimodel analysis and implications for our perception of the land surface, *Bulletin of the American Meteorological Society*, *87*(10), 1381–1397, doi:10.1175/BAMS-87-10-1381.

Dirmeyer, P. A., Y. Jin, B. Singh, and X. Yan (2013), Trends in land–atmosphere interactions from CMIP5 simulations, *Journal of Hydrometeorology*, *14*(3), 829–849, doi:10.1175/JHM-D-12-0107.1.

Don, A., I. H. Böhme, A. B. Dohrmann, C. Poeplau, and C. C. Tebbe (2017), Microbial community composition affects soil organic carbon turnover in mineral soils, *Biology and Fertility of Soils*, *53*(4), 445–456, doi:10.1007/s00374-017-1198-9.

Dufresne, J.-L., M.-A. Foujols, S. Denvil, A. Caubel, O. Marti, O. Aumont, Y. Balkanski, S. Bekki, H. Bellenger, R. Benshila, S. Bony, L. Bopp, P. Braconnot, P. Brockmann, P. Cadule, F. Cheruy, F. Codron, A. Cozic, D. Cugnet, N. Noblet, J.-P. Duvel, C. Ethé, L. Fairhead, T. Fichefet, S. Flavoni, P. Friedlingstein, J.-Y. Grandpeix, L. Guez, E. Guilyardi, D. Hauglustaine, F. Hourdin, A. Idelkadi, J. Ghattas, S. Joussaume, M. Kageyama, G. Krinner, S. Labetoulle, A. Lahellec, M.-P. Lefebvre, F. Lefevre, C. Levy, Z. X. Li, J. Lloyd, F. Lott, G. Madec, M. Mancip, M. Marchand, S. Masson, Y. Meurdesoif, J. Mignot, I. Musat, S. Parouty, J. Polcher, C. Rio, M. Schulz, D. Swingedouw, S. Szopa, C. Talandier, P. Terray, N. Viovy, and N. Vuichard (2013), Climate change projections using the IPSL-CM5 Earth system model: from CMIP3 to CMIP5, *Climate Dynamics*, *40*(9), 2123–2165, doi:10.1007/s00382-012-1636-1.

Dunne, J. P., J. G. John, A. J. Adcroft, S. M. Griffies, R. W. Hallberg, E. Shevliakova, R. J. Stouffer, W. Cooke, K. A. Dunne, M. J. Harrison, J. P. Krasting, S. L. Malyshev, P. C. D. Milly, P. J. Phillipps, L. T. Sentman, B. L. Samuels, M. J. Spelman, M. Winton, A. T. Wittenberg, and N. Zadeh (2012), GFDL's ESM2 global coupled climate–carbon earth system models. Part i: Physical formulation and baseline simulation characteristics, *Journal of Climate*, *25*(19), 6646–6665, doi:10.1175/JCLI-D-11-00560.1.

Dwivedi, D., W. Riley, M. Torn, N. Spycher, F. Maggi, and J. Tang (2017), Mineral properties, microbes, transport, and plant-input profiles control vertical distribution and age of soil carbon stocks, *Soil Biology and Biochemistry*, *107*, 244–259, doi:https://doi.org/10.1016/j.soilbio.2016.12.019.

Eltahir, E. A. B. (1998), A soil moisture–rainfall feedback mechanism: 1. Theory and observations, *Water Resources Research*, *34*(4), 765–776, doi:10.1029/97WR03499.

Ferguson, C. R., E. F. Wood, and R. K. Vinukollu (2012), A global intercomparison of modeled and observed land–atmosphere coupling, *Journal of Hydrometeorology*, *13*(3), 749–784, doi: 10.1175/JHM-D-11-0119.1.

Findell, K. L., and E. A. B. Eltahir (1997), An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois, *Water Resources Research*, *33*(4), 725–735, doi: 10.1029/96WR03756.

Findell, K. L., and E. A. B. Eltahir (2003a), Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development, *Journal of Hydrometeorology*, *4*(3), 552–569, doi:10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2.

Findell, K. L., and E. A. B. Eltahir (2003b), Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the Continental United States, *Journal of Hydrometeorology*, *4*(3), 570–583, doi:10.1175/1525-7541(2003)004<0570:ACOSML>2.0.CO;2.

Findell, K. L., P. Gentine, B. R. Lintner, and C. Kerr (2011), Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation, *Nature Geosci*, *4*(7), 434–439, doi:10.1038/ngeo1174.

Findell, K. L., P. Gentine, B. R. Lintner, and B. P. Guillod (2015), Data length requirements for observational estimates of land–atmosphere coupling strength, *Journal of Hydrometeorology*, *16*(4), 1615–1635, doi:10.1175/JHM-D-14-0131.1.

Flannigan, M., A. S. Cantin, W. J. de Groot, M. Wotton, A. Newbery, and L. M. Gowman (2013), Global wildland fire season severity in the 21st century, *Forest Ecology and Management*, *294*, 54–61, doi:10.1016/j.foreco.2012.10.022, the Mega-fire reality.

Flechtner, F., P. Morton, M. Watkins, and F. Webb (2014), Status of the grace follow-on mission, in *Gravity, Geoid and Height Systems: Proceedings of the IAG Symposium GGHS2012, October 9-12, 2012, Venice, Italy*, edited by U. Marti, pp. 117–121, Springer International Publishing, Switzerland, doi:10.1007/978-3-319-10837-7_15.

Foley, J. A., A. Botta, M. T. Coe, and M. H. Costa (2002), El Niño–Southern Oscillation and the climate, ecosystems and rivers of Amazonia, *Global Biogeochemical Cycles*, *16*(4), 1132, doi:10.1029/2002GB001872.

Friedlingstein, P., P. Cox, R. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K.-G. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng (2006), Climate–carbon cycle feedback analysis: Results from the C4MIP model intercomparison, *Journal of Climate*, *19*(14), 3337–3353, doi:10.1175/JCLI3800.1.

Friedlingstein, P., M. Meinshausen, V. K. Arora, C. D. Jones, A. Anav, S. K. Liddicoat, and R. Knutti (2014), Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, *Journal of Climate*, *27*(2), 511–526, doi:10.1175/JCLI-D-12-00579.1.

Fu, R., R. E. Dickinson, M. Chen, and H. Wang (2001), How do tropical sea surface temperatures influence the seasonal distribution of precipitation in the equatorial Amazon?, *Journal of Climate*, *14*(20), 4003–4026, doi:{10.1175/1520-0442(2001)014<4003:HDTSST>2.0.CO;2}.

Gates, W. L., J. S. Boyle, C. Covey, C. G. Dease, C. M. Doutriaux, R. S. Drach, M. Fiorino, P. J. Gleckler, J. J. Hnilo, S. M. Marlais, T. J. Phillips, G. L. Potter, B. D. Santer, K. R. Sperber, K. E. Taylor, and D. N. Williams (1999), An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I), *Bulletin of the American Meteorological Society*, *80*(1), 29–55, doi:10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2.

Gelaro, R., W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S. D. Schubert, M. Sienkiewicz, and B. Zhao (2017), The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), *Journal of Climate*, *30*(14), 5419–5454, doi:10.1175/JCLI-D-16-0758.1.

Guillod, B. P., B. Orlowsky, D. Miralles, A. J. Teuling, P. D. Blanken, N. Buchmann, P. Ciais, M. Ek, K. L. Findell, P. Gentine, B. R. Lintner, R. L. Scott, B. Van den Hurk, and S. I. Seneviratne (2014), Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors, *Atmospheric Chemistry and Physics*, *14*(16), 8343–8367, doi:10.5194/acp-14-8343-2014.

Guillod, B. P., B. Orlowsky, D. G. Miralles, A. J. Teuling, and S. I. Seneviratne (2015), Reconciling spatial and temporal soil moisture effects on afternoon rainfall, *Nature Communications*, *6*(6443), doi:10.1038/ncomms7443.

Guo, Z., and P. A. Dirmeyer (2013), Interannual variability of land–atmosphere coupling strength, *Journal of Hydrometeorology*, *14*(5), 1636–1646, doi:10.1175/JHM-D-12-0171.1.

Guo, Z., P. A. Dirmeyer, R. D. Koster, Y. C. Sud, G. Bonan, K. W. Oleson, E. Chan, D. Verseghy, P. Cox, C. T. Gordon, J. L. McGregor, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, D. Mocko, C.-H. Lu, K. Mitchell, S. Malyshev, B. McAvaney, T. Oki, T. Yamada, A. Pitman, C. M. Taylor, R. Vasic, and Y. Xue (2006), GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis, *Journal of Hydrometeorology*, *7*(4), 611–625, doi:10.1175/JHM511.1.

Hajima, T., K. Tachiiri, A. Ito, and M. Kawamiya (2014), Uncertainty of concentration–terrestrial carbon feedback in Earth system models, *Journal of Climate*, *27*(9), 3425–3445, doi:10.1175/JCLI-D-13-00177.1.

Hall, A., and X. Qu (2006), Using the current seasonal cycle to constrain snow albedo feedback in future climate change, *Geophysical Research Letters*, *33*(3), L03502, doi:10.1029/2005GL025127.

Hansen, J., A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lerner (1984), Climate sensitivity: Analysis of feedback mechanisms, in *Climate Processes and Climate Sensitivity*, edited by J. Hansen and T. Takahashi, pp. 130–163, American Geophysical Union (AGU), doi:10.1029/GM029p0130.

Harper, A., I. T. Baker, A. S. Denning, D. A. Randall, D. Dazlich, and M. Branson (2013), Impact of evapotranspiration on dry season climate in the Amazon forest, *Journal of Climate*, *27*(2), 574–591, doi:10.1175/JCLI-D-13-00074.1.

Harper, R. J., and M. Tibbett (2013), The hidden organic carbon in deep mineral soils, *Plant and Soil*, *368*(1), 641–648, doi:10.1007/s11104-013-1600-9.

Harris, I., P. Jones, T. Osborn, and D. Lister (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset, *International Journal of Climatology*, *34*(3), 623–642, doi:10.1002/joc.3711.

Hashimoto, H., R. R. Nemani, M. A. White, W. M. Jolly, S. C. Piper, C. D. Keeling, R. B. Myneni, and S. W. Running (2004), El Niño–Southern Oscillation–induced variability in terrestrial carbon cycling, *Journal of Geophysical Research: Atmospheres*, *109*(D23), D23110, doi:10.1029/2004JD004959, d23110.

He, Y., S. E. Trumbore, M. S. Torn, J. W. Harden, L. J. S. Vaughn, S. D. Allison, and J. T. Randerson (2016), Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century, *Science*, *353*(6306), 1419–1424, doi:10.1126/science.aad4273.

Heimann, M., and M. Reichstein (2008), Terrestrial ecosystem carbon dynamics and climate feedbacks, *Nature*, *451*, 289–292.

Hengl, T., J. Mendes de Jesus, G. B. M. Heuvelink, M. Ruiperez Gonzalez, M. Kilibarda, A. Blagotić, W. Shangguan, M. N. Wright, X. Geng, B. Bauer-Marschallinger, M. A. Guevara, R. Vargas, R. A. MacMillan, N. H. Batjes, J. G. B. Leenaars, E. Ribeiro, I. Wheeler, S. Mantel, and B. Kempen (2017), Soilgrids250m: Global gridded soil information based on machine learning, *PLOS ONE*, *12*(2), 1–40, doi:10.1371/journal.pone.0169748.

Hilker, T., A. I. Lyapustin, C. J. Tucker, F. G. Hall, R. B. Myneni, Y. Wang, J. Bi, Y. Mendes de Moura, and P. J. Sellers (2014), Vegetation dynamics and rainfall sensitivity of the Amazon, *Proceedings of the National Academy of Sciences*, *111*(45), 16,041–16,046, doi:10.1073/pnas.1404870111.

Hirschi, M., S. I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O. B. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek (2011), Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, *Nature Geoscience*, *4*(1), 17–21, doi:10.1038/ngeo1032.

Hirschi, M., B. Mueller, W. Dorigo, and S. Seneviratne (2014), Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability, *Remote Sensing of Environment*, *154*, 246–252, doi:10.1016/j.rse.2014.08.030.

Hoffman, F. M., J. T. Randerson, V. K. Arora, Q. Bao, P. Cadule, D. Ji, C. D. Jones, M. Kawamiya, S. Khatiwala, K. Lindsay, A. Obata, E. Shevliakova, K. D. Six, J. F. Tjiputra, E. M. Volodin,

and T. Wu (2014), Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, *Journal of Geophysical Research: Biogeosciences*, *119*(2), 141–162, doi:10.1002/2013JG002381.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär (2009), The soil moisture–precipitation feedback in simulations with explicit and parameterized convection, *Journal of Climate*, *22*(19), 5003–5020, doi:10.1175/2009JCLI2604.1.

Hourdin, F., M.-A. Foujols, F. Codron, V. Guemas, J.-L. Dufresne, S. Bony, S. Denvil, L. Guez, F. Lott, J. Ghattas, P. Braconnot, O. Marti, Y. Meurdesoif, and L. Bopp (2013), Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model, *Climate Dynamics*, *40*(9), 2167–2192, doi:10.1007/s00382-012-1411-3.

Huber, M., I. Mahlstein, M. Wild, J. Fasullo, and R. Knutti (2011), Constraints on climate sensitivity from radiation patterns in climate models, *Journal of Climate*, *24*(4), 1034–1052, doi:10.1175/2010JCLI3403.1.

Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu (2009), Improving the global precipitation record: GPCP version 2.1, *Geophysical Research Letters*, *36*(17), L17808, doi:10.1029/2009GL040000.

Humphrey, V., J. Zscheischler, P. Ciais, L. Gudmundsson, S. Sitch, and S. I. Seneviratne (2018), Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage, *Nature*, *560*(7720), 628–631, doi:10.1038/s41586-018-0424-4.

Hung, M.-P., J.-L. Lin, W. Wang, D. Kim, T. Shinoda, and S. J. Weaver (2013), MJO and convectively coupled equatorial waves simulated by CMIP5 climate models, *Journal of Climate*, *26*(17), 6185–6214, doi:10.1175/JCLI-D-12-00541.1.

Jackson, R. B., K. Lajtha, S. E. Crow, G. Hugelius, M. G. Kramer, and G. Piñeiro (2017), The ecology of soil carbon: Pools, vulnerabilities, and biotic and abiotic controls, *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 419–445, doi:10.1146/annurev-ecolsys-112414-054234.

Jenkinson, D. (1990), The turnover of organic carbon and nitrogen in soil, *Philosophical Transactions of the Royal Society B: Biological Sciences*, *329*(1255), 361–368, doi:10.1098/rstb.1990.0177.

Jiménez-Muñoz, J. C., C. Mattar, J. Barichivich, A. Santamaría-Artigas, K. Takahashi, Y. Malhi, J. Sobrino, and G. v. d. Schrier (2016), Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016, *Scientific Reports*, *6*, 33,130, doi:10.1038/srep33130.

Joetzjer, E., H. Douville, C. Delire, and P. Ciais (2013), Present-day and future Amazonian precipitation in global climate models: CMIP5 versus CMIP3, *Climate Dynamics*, *41*(11), 2921–2936, doi:10.1007/s00382-012-1644-1.

Jones, C., P. Cox, and C. Huntingford (2011), Uncertainty in climate–carbon-cycle projections associated with the sensitivity of soil respiration to temperature, *Tellus B*, *55*(2), 642–648, doi:10.1034/j.1600-0889.2003.01440.x.

Jones, C. D., M. Collins, P. M. Cox, and S. A. Spall (2001), The carbon cycle response to ENSO: A coupled climate–carbon cycle model study, *Journal of Climate*, *14*(21), 4113–4129, doi:10.1175/1520-0442(2001)014<4113:TCCRTE>2.0.CO;2.

Juárez, R. I. N., M. G. Hodnett, R. Fu, M. L. Goulden, and C. von Randow (2007), Control of dry season evapotranspiration over the Amazonian forest as inferred from observations at a southern Amazon forest site, *Journal of Climate*, *20*(12), 2827–2839, doi:10.1175/JCLI4184.1.

Jung, M., M. Reichstein, C. R. Schwalm, C. Huntingford, S. Sitch, A. Ahlström, A. Arneth, G. Camps-Valls, P. Ciais, P. Friedlingstein, F. Gans, K. Ichii, A. K. Jain, E. Kato, D. Papale, B. Poulter, B. Raduly, C. Rödenbeck, G. Tramontana, N. Viovy, Y.-P. Wang, U. Weber, S. Zaehle, and N. Zeng (2017), Compensatory water effects link yearly global land CO_2 sink changes to temperature, *Nature*, *541*, 516–520.

Kataoka, F., D. Crisp, T. E. Taylor, C. W. O'Dell, A. Kuze, K. Shiomi, H. Suto, C. Bruegge, F. M. Schwandner, R. Rosenberg, L. Chapsky, and R. A. M. Lee (2017), The cross-calibration of spectral radiances and cross-validation of CO_2 estimates from GOSAT and OCO-2, *Remote Sensing*, *9*(11), 1158, doi:10.3390/rs9111158.

Kay, J. E., C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. M. Arblaster, S. C. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J. F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein (2014), The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability, *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349, doi:10.1175/BAMS-D-13-00255.1.

Keeling, C. D., and R. Revelle (1985), Effects of El Nino/Southern Oscillation on the atmospheric content of carbon dioxide, *Meteoritics*, *20*(2), 437–450.

Keenan, T. F., I. C. Prentice, J. G. Canadell, C. A. Williams, H. Wang, M. Raupach, and G. J. Collatz (2016), Recent pause in the growth rate of atmospheric CO_2 due to enhanced terrestrial carbon uptake, *Nature Communications*, *7*, 13,428, doi:10.1038/ncomms13428.

Keppel-Aleks, G., A. S. Wolf, M. Mu, S. C. Doney, D. C. Morton, P. S. Kasibhatla, J. B. Miller, E. J. Dlugokencky, and J. T. Randerson (2014), Separating the influence of temperature, drought, and fire on interannual variability in atmospheric CO_2, *Global Biogeochemical Cycles*, *28*(11), 1295–1310, doi:10.1002/2014GB004890, 2014GB004890.

Keppel-Aleks, G., S. J. Basile, and F. M. Hoffman (2018), A functional response metric for the temperature sensitivity of tropical ecosystems, *Earth Interactions*, *22*(7), 1–20, doi:10.1175/EI-D-17-0017.1.

Kim, H., P. J.-F. Yeh, T. Oki, and S. Kanae (2009), Role of rivers in the seasonal variations of terrestrial water storage over global basins, *Geophysical Research Letters*, *36*(17), L17402, doi:10.1029/2009GL039006.

Klein, S. A., and A. Hall (2015), Emergent constraints for cloud feedbacks, *Current Climate Change Reports*, *1*(4), 276–287, doi:10.1007/s40641-015-0027-1.

Koarashi, J., W. C. Hockaday, C. A. Masiello, and S. E. Trumbore (2012), Dynamics of decadally cycling carbon in subsurface soils, *Journal of Geophysical Research: Biogeosciences*, *117*(G3), G03033, doi:10.1029/2012JG002034.

Kooperman, G. J., M. S. Pritchard, M. A. Burt, M. D. Branson, and D. A. Randall (2016), Robust effects of cloud superparameterization on simulated daily rainfall intensity statistics across multiple versions of the Community Earth System Model, *Journal of Advances in Modeling Earth Systems*, *8*, 140–165, doi:10.1002/2015MS000574.

Kooperman, G. J., Y. Chen, F. M. Hoffman, C. D. Koven, K. Lindsay, M. S. Pritchard, A. L. S. Swann, and J. T. Randerson (2018), Forest response to rising CO_2 drives zonally asymmetric rainfall change over tropical land, *Nature Climate Change*, *8*(5), 434–440, doi:10.1038/s41558-018-0144-7.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*(5687), 1138–1140, doi:10.1126/science.1100217.

Koster, R. D., Y. C. Sud, Z. Guo, P. A. Dirmeyer, G. Bonan, K. W. Oleson, E. Chan, D. Verseghy, P. Cox, H. Davies, E. Kowalczyk, C. T. Gordon, S. Kanae, D. Lawrence, P. Liu, D. Mocko, C.-H. Lu, K. Mitchell, S. Malyshev, B. McAvaney, T. Oki, T. Yamada, A. Pitman, C. M. Taylor, R. Vasic, and Y. Xue (2006), GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview, *Journal of Hydrometeorology*, *7*(4), 590–610, doi:10.1175/JHM510.1.

Koster, R. D., S. P. P. Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, P. A. Dirmeyer, F. J. Doblas-Reyes, G. Drewitt, C. T. Gordon, Z. Guo, J.-H. Jeong, D. M. Lawrence, W.-S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B. J. J. M. van den Hurk, F. Vitart, and E. F. Wood (2010), Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment, *Geophysical Research Letters*, *37*(2), L02402, doi:10.1029/2009GL041677.

Koster, R. D., S. P. P. Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, P. A. Dirmeyer, F. J. Doblas-Reyes, G. Drewitt, C. T. Gordon, Z. Guo, J. H. Jeong, W. S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B. J. J. M. van den Hurk, F. Vitart, and E. F. Wood (2011), The second phase of the Global Land–Atmosphere Coupling Experiment: Soil moisture contributions to subseasonal forecast skill, *Journal of Hydrometeorology*, *12*(5), 805–822, doi:10.1175/2011JHM1365.1.

Koven, C. D., W. J. Riley, Z. M. Subin, J. Y. Tang, M. S. Torn, W. D. Collins, G. B. Bonan, D. M. Lawrence, and S. C. Swenson (2013), The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, *Biogeosciences*, *10*(11), 7109–7131, doi:10.5194/bg-10-7109-2013.

Kramer, M. G., and O. A. Chadwick (2018), Climate-driven thresholds in reactive mineral retention of soil carbon at the global scale, *Nature Climate Change*, *8*(12), 1104–1108, doi:10.1038/s41558-018-0341-4.

Kramer, M. G., K. Lajtha, and A. K. Aufdenkampe (2017), Depth trends of soil organic matter C:N and ^{15}N natural abundance controlled by association with minerals, *Biogeochemistry*, *136*(3), 237–248, doi:10.1007/s10533-017-0378-x.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled grace terrestrial water storage estimates, *Water Resources Research*, *48*(4), W04/,/!531, doi:10.1029/2011WR011453.

Lawrence, C., J. Beem-Miller, A. M. Hoyt, G. Monroe, C. A. Sierra, S. Stoner, K. Heckman, J. Blankinship, S. E. Crow, G. McNicol, S. E. Trumbore, P. A. Levine, O. Vindušková, K. E. O. Todd-Brown, C. Rasmussen, C. H. Pries, C. Schädel, K. McFarlane, S. Doetterl, C. Hatté, Y. He, C. Treat, J. Harden, M. Torn, C. Estop-Aragonés, A. A. Berhe, M. Keiluweit, E. Marin-Spiotta, A. Plante, A. Thomson, J. Schimel, L. Vaughn, and R. Wagai (2019), An open source database for the synthesis of soil radiocarbon data: ISRaD version 1.0, *Earth System Science Data Discussions*, *in review*, doi:10.5194/essd-2019-55.

Lawrence, D. M., K. W. Oleson, M. G. Flanner, C. G. Fletcher, P. J. Lawrence, S. Levis, S. C. Swenson, and G. B. Bonan (2011), The CCSM4 land simulation, 1850–2005: Assessment of surface climate and new capabilities, *Journal of Climate*, *25*(7), 2240–2260, doi:10.1175/JCLI-D-11-00103.1.

Lawrence, M. G. (2005), The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications, *Bulletin of the American Meteorological Society*, *86*(2), 225–233, doi:10.1175/BAMS-86-2-225.

Le Quéré, C., R. M. Andrew, P. Friedlingstein, S. Sitch, J. Hauck, J. Pongratz, P. A. Pickers, J. I. Korsbakken, G. P. Peters, J. G. Canadell, A. Arneth, V. K. Arora, L. Barbero, A. Bastos, L. Bopp, F. Chevallier, L. P. Chini, P. Ciais, S. C. Doney, T. Gkritzalis, D. S. Goll, I. Harris, V. Haverd, F. M. Hoffman, M. Hoppema, R. A. Houghton, G. Hurtt, T. Ilyina, A. K. Jain, T. Johannessen, C. D. Jones, E. Kato, R. F. Keeling, K. K. Goldewijk, P. Landschützer, N. Lefèvre, S. Lienert, Z. Liu, D. Lombardozzi, N. Metzl, D. R. Munro, J. E. M. S. Nabel, S.-I. Nakaoka, C. Neill, A. Olsen, T. Ono, P. Patra, A. Peregon, W. Peters, P. Peylin, B. Pfeil, D. Pierrot, B. Poulter, G. Rehder, L. Resplandy, E. Robertson, M. Rocher, C. Rödenbeck, U. Schuster, J. Schwinger, R. Séférian, I. Skjelvan, T. Steinhoff, A. Sutton, P. P. Tans, H. Tian, B. Tilbrook, F. N. Tubiello, I. T. van der Laan-Luijkx, G. R. van der Werf, N. Viovy, A. P. Walker, A. J. Wiltshire, R. Wright, S. Zaehle, and B. Zheng (2018), Global carbon budget 2018, *Earth System Science Data*, *10*(4), 2141–2194, doi:10.5194/essd-10-2141-2018.

Lee, J.-E., R. S. Oliveira, T. E. Dawson, and I. Fung (2005), Root functioning modifies seasonal climate, *Proceedings of the National Academy of Sciences of the United States of America*, *102*(49), 17,576–17,581, doi:10.1073/pnas.0508785102.

Lee, J.-E., B. R. Lintner, C. K. Boyce, and P. J. Lawrence (2011), Land use change exacerbates tropical South American drought by sea surface temperature variability, *Geophysical Research Letters*, *38*(19), L19706, doi:10.1029/2011GL049066, l19706.

Levine, P. A., J. T. Randerson, S. C. Swenson, and D. M. Lawrence (2016), Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations, *Hydrology and Earth System Sciences*, *20*, 4837–4856, doi:10.5194/hess-20-4837-2016.

Levine, P. A., J. T. Randerson, Y. Chen, M. S. Pritchard, M. Xu, and F. M. Hoffman (2019), Soil moisture variability intensifies and prolongs eastern Amazon temperature and carbon cycle response to El Niño–Southern Oscillation, *Journal of Climate*, *32*(4), 1273–1292, doi: 10.1175/JCLI-D-18-0150.1.

Lindsay, K., G. B. Bonan, S. C. Doney, F. M. Hoffman, D. M. Lawrence, M. C. Long, N. M. Mahowald, J. Keith Moore, J. T. Randerson, and P. E. Thornton (2014), Preindustrial-control and twentieth-century carbon cycle experiments with the Earth system model CESM1(BGC), *Journal of Climate*, *27*(24), 8981–9005, doi:10.1175/JCLI-D-12-00565.1.

Liu, J., K. W. Bowman, M. Lee, D. K. Henze, N. Bousserez, H. Brix, G. J. Collatz, D. Menemenlis, L. Ott, S. Pawson, D. Jones, and R. Nassar (2014), Carbon monitoring system flux estimation and attribution: impact of ACOS-GOSAT X_{CO_2} sampling on the inference of terrestrial biospheric sources and sinks, *Tellus B: Chemical and Physical Meteorology*, *66*(1), 22486, doi:10.3402/tellusb.v66.22486.

Liu, J., K. W. Bowman, D. S. Schimel, N. C. Parazoo, Z. Jiang, M. Lee, A. A. Bloom, D. Wunch, C. Frankenberg, Y. Sun, C. W. O'Dell, K. R. Gurney, D. Menemenlis, M. Gierach, D. Crisp, and A. Eldering (2017), Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño, *Science*, *358*(6360), eaam5690, doi:10.1126/science.aam5690.

Llovel, W., M. Becker, A. Cazenave, S. Jevrejeva, R. Alkama, B. Decharme, H. Douville, M. Ablain, and B. Beckley (2011), Terrestrial waters and sea level variations on interannual time scale, *Global and Planetary Change*, *75*(1–2), 76 – 82, doi:10.1016/j.gloplacha.2010.10.008.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong (2009), Toward optimal closure of the Earth's top-of-atmosphere radiation budget, *Journal of Climate*, *22*(3), 748–766, doi:10.1175/2008JCLI2637.1.

Lorenz, R., and A. J. Pitman (2014), Effect of land-atmosphere coupling strength on impacts from Amazonian deforestation, *Geophysical Research Letters*, *41*(16), 5987–5995, doi:10.1002/2014GL061017.

Luo, Y., A. Ahlström, S. D. Allison, N. H. Batjes, V. Brovkin, N. Carvalhais, A. Chappell, P. Ciais, E. A. Davidson, A. Finzi, K. Georgiou, B. Guenet, O. Hararuk, J. W. Harden, Y. He, F. Hopkins, L. Jiang, C. Koven, R. B. Jackson, C. D. Jones, M. J. Lara, J. Liang, A. D. McGuire, W. Parton, C. Peng, J. T. Randerson, A. Salazar, C. A. Sierra, M. J. Smith, H. Tian, K. E. O. Todd-Brown, M. Torn, K. J. Groenigen, Y. P. Wang, T. O. West, Y. Wei, W. R. Wieder, J. Xia, X. Xu, X. Xu, and T. Zhou (2016), Toward more realistic projections of soil carbon dynamics by Earth system models, *Global Biogeochemical Cycles*, *30*(1), 40–56, doi:10.1002/2015GB005239.

Ma, H.-Y., C. R. Mechoso, Y. Xue, H. Xiao, C.-M. Wu, J.-L. Li, and F. Sales (2011), Impact of land surface processes on the South American warm season climate, *Climate Dynamics*, *37*(1), 187–203, doi:10.1007/s00382-010-0813-3.

Manzoni, S., and A. Porporato (2009), Soil carbon and nitrogen mineralization: Theory and models across scales, *Soil Biology and Biochemistry*, *41*(7), 1355–1379, doi:https://doi.org/10.1016/j.soilbio.2009.02.031.

Marengo, J. A. (1992), Interannual variability of surface climate in the Amazon Basin, *International Journal of Climatology*, *12*(8), 853–863, doi:10.1002/joc.3370120808.

Mathieu, J. A., C. Hatté, J. Balesdent, and É. Parent (2015), Deep soil carbon dynamics are driven more by soil type than by climate: a worldwide meta-analysis of radiocarbon profiles, *Global Change Biology*, *21*(11), 4278–4292, doi:10.1111/gcb.13012.

May, W., A. Meier, M. Rummukainen, A. Berg, F. Chéruy, and S. Hagemann (2015), Contributions of soil moisture interactions to climate change in the tropics in the GLACE–CMIP5 experiment, *Climate Dynamics*, *45*(11), 1–23, doi:10.1007/s00382-015-2538-9.

Mecklenburg, S., M. Drusch, Y. Kerr, J. Font, M. Martin-Neira, S. Delwart, G. Buenadicha, N. Reul, E. Daganzo-Eusebio, R. Oliva, and R. Crapolicchio (2012), ESA's Soil Moisture and Ocean Salinity mission: Mission performance and operations, *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1354–1366, doi:10.1109/TGRS.2012.2187666.

Medlyn, B. E., R. A. Duursma, D. Eamus, D. S. Ellsworth, I. C. Prentice, C. V. M. Barton, K. Y. Crous, P. De Angelis, M. Freeman, and L. Wingate (2011), Reconciling the optimal and empirical approaches to modelling stomatal conductance, *Global Change Biology*, *17*(6), 2134–2144, doi:10.1111/j.1365-2486.2010.02375.x.

Meehl, G. A., W. M. Washington, J. M. Arblaster, A. Hu, H. Teng, J. E. Kay, A. Gettelman, D. M. Lawrence, B. M. Sanderson, and W. G. Strand (2013), Climate change projections in CESM1(CAM5) compared to CCSM4, *Journal of Climate*, *26*(17), 6287–6308, doi:10.1175/JCLI-D-12-00572.1.

Mei, R., and G. Wang (2011), Impact of sea surface temperature and soil moisture on summer precipitation in the United States based on observational data, *Journal of Hydrometeorology*, *12*(5), 1086–1099, doi:10.1175/2011JHM1312.1.

Mei, R., and G. Wang (2012), Summer land–atmosphere coupling strength in the United States: Comparison among observations, reanalysis data, and numerical models, *Journal of Hydrometeorology*, *13*(3), 1010–1022, doi:10.1175/JHM-D-11-075.1.

Melillo, J. M., S. D. Frey, K. M. DeAngelis, W. J. Werner, M. J. Bernard, F. P. Bowles, G. Pold, M. A. Knorr, and A. S. Grandy (2017), Long-term pattern and magnitude of soil carbon feedback to the climate system in a warming world, *Science*, *358*(6359), 101–105, doi:10.1126/science.aan2874.

Miguez-Macho, G., and Y. Fan (2012), The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration, *Journal of Geophysical Research: Atmospheres*, *117*(D15), D15114, doi:10.1029/2012JD017540.

Minasny, B., B. P. Malone, A. B. McBratney, D. A. Angers, D. Arrouays, A. Chambers, V. Chaplot, Z.-S. Chen, K. Cheng, B. S. Das, D. J. Field, A. Gimona, C. B. Hedley, S. Y. Hong, B. Mandal, B. P. Marchant, M. Martin, B. G. McConkey, V. L. Mulder, S. O'Rourke, A. C. R. de Forges, I. Odeh, J. Padarian, K. Paustian, G. Pan, L. Poggio, I. Savin, V. Stolbovoy, U. Stockmann, Y. Sulaeman, C.-C. Tsui, T.-G. Vågen, B. van Wesemael, and L. Winowiecki (2017), Soil carbon 4 per mille, *Geoderma*, *292*, 59–86, doi:https://doi.org/10.1016/j.geoderma.2017.01.002.

Miralles, D. G., M. J. van den Berg, A. J. Teuling, and R. A. M. de Jeu (2012), Soil moisture-temperature coupling: A multiscale observational analysis, *Geophysical Research Letters*, *39*(21), L21707, doi:10.1029/2012GL053703.

Miralles, D. G., A. J. Teuling, C. C. van Heerwaarden, and J. Vila-Guerau de Arellano (2014), Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, *Nature Geoscience*, *7*(5), 345–349, doi:10.1038/ngeo2141.

Mueller, B., and S. I. Seneviratne (2014), Systematic land climate and evapotranspiration biases in CMIP5 simulations, *Geophysical Research Letters*, *41*(1), 128–134, doi:10.1002/2013GL058055.

Neale, R. B., C.-C. Chen, A. Gettelman, P. H. Lauritzen, S. Park, D. L. Williamson, A. J. Conley, R. Garcia, D. Kinnison, J.-F. Lamarque, et al. (2012), Description of the NCAR Community Atmosphere Model (CAM 5.0), *Technical Note NCAR/TN-486+ STR*, National Center for Atmospheric Research.

Neale, R. B., J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang (2013), The mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully coupled experiments, *Journal of Climate*, *26*(14), 5150–5168, doi:10.1175/JCLI-D-12-00236.1.

Nepstad, D. C., C. R. de Carvalho, E. A. Davidson, P. H. Jipp, P. A. Lefebvre, G. H. Negreiros, E. D. da Silva, T. A. Stone, S. E. Trumbore, and S. Vieira (1994), The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures, *Nature*, *372*, 666–669, doi:10.1038/372666a0.

Norby, R. J., and D. R. Zak (2011), Ecological lessons from Free-Air CO_2 Enrichment (FACE) experiments, *Annual Review of Ecology, Evolution, and Systematics*, *42*(1), 181–203, doi:10.1146/annurev-ecolsys-102209-144647.

Notaro, M. (2008), Statistical identification of global hot spots in soil moisture feedbacks among IPCC AR4 models, *Journal of Geophysical Research: Atmospheres*, *113*(D9), D09101, doi:10.1029/2007JD009199.

Oleson, K. W., D. M. Lawrence, G. B. Bonan, B. Drewniak, M. Huang, C. Koven, S. Levis, F. Li, W. Riley, Z. Subin, S. Swenson, P. E. Thornton, A. Bozbiyik, R. Fisher, C. Heald, E. Kluzek, J.-F. Lamarque, P. Lawrence, L. Leung, W. Lipscomb, S. Muszala, D. Ricciuto, W. Sacks, Y. Sun, J. Tang, and Z. L. Yang (2013), Technical description of version 4.5 of the Community Land Model (CLM), *Technical Note NCAR/TN-503+STR*, National Center for Atmospheric Research, doi:10.5065/D6RR1W7M.

Oliveira, R. S., T. E. Dawson, S. S. O. Burgess, and D. C. Nepstad (2005), Hydraulic redistribution in three Amazonian trees, *Oecologia*, *145*(3), 354–363, doi:10.1007/s00442-005-0108-2.

Orlowsky, B., and S. I. Seneviratne (2010), Statistical analyses of land–atmosphere feedbacks and their possible pitfalls, *Journal of Climate*, *23*(14), 3918–3932, doi:10.1175/2010JCLI3366.1.

Orth, R., and S. I. Seneviratne (2017), Variability of soil moisture and sea surface temperatures similarly important for warm-season land climate in the Community Earth System Model, *Journal of Climate*, *30*(6), 2141–2162, doi:10.1175/JCLI-D-15-0567.1.

Panciera, R., J. Walker, T. Jackson, D. Gray, M. Tanase, D. Ryu, A. Monerris, H. Yardley, C. Rudiger, X. Wu, Y. Gao, and J. Hacker (2014), The Soil Moisture Active Passive Experiments (SMAPEx): Toward soil moisture retrieval from the SMAP mission, *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 490–507, doi:10.1109/TGRS.2013.2241774.

Parazoo, N. C., K. Bowman, J. B. Fisher, C. Frankenberg, D. B. A. Jones, A. Cescatti, Ó. Pérez-Priego, G. Wohlfahrt, and L. Montagnani (2014), Terrestrial gross primary production inferred from satellite fluorescence and vegetation models, *Global Change Biology*, *20*(10), 3103–3121, doi:10.1111/gcb.12652.

Parton, W. J., J. M. O. Scurlock, D. S. Ojima, T. G. Gilmanov, R. J. Scholes, D. S. Schimel, T. Kirchner, J. Menaut, T. Seastedt, E. G. Moya, A. Kamnalrut, and J. I. Kinyamario (1993), Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, *Global Biogeochemical Cycles*, *7*(4), 785–809, doi:10.1029/93GB02042.

Paul, E. A. (2016), The nature and dynamics of soil organic matter: Plant inputs, microbial transformations, and organic matter stabilization, *Soil Biology and Biochemistry*, *98*, 109–126, doi:https://doi.org/10.1016/j.soilbio.2016.04.001.

Phillips, T. J., and S. A. Klein (2014), Land-atmosphere coupling manifested in warm-season observations on the U.S. southern great plains, *Journal of Geophysical Research: Atmospheres*, *119*(2), 509–528, doi:10.1002/2013JD020492.

Piao, S., S. Sitch, P. Ciais, P. Friedlingstein, P. Peylin, X. Wang, A. Ahlström, A. Anav, J. G. Canadell, N. Cong, C. Huntingford, M. Jung, S. Levis, P. E. Levy, J. Li, X. Lin, M. R. Lomas, M. Lu, Y. Luo, Y. Ma, R. B. Myneni, B. Poulter, Z. Sun, T. Wang, N. Viovy, S. Zaehle, and N. Zeng (2013), Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends, *Global Change Biology*, *19*(7), 2117–2132, doi:10.1111/gcb.12187.

Qian, H., R. Joseph, and N. Zeng (2008), Response of the terrestrial carbon cycle to the El Niño-Southern Oscillation, *Tellus B*, *60*(4), 537–550, doi:10.1111/j.1600-0889.2008.00360.x.

Qu, X., and A. Hall (2014), On the persistent spread in snow-albedo feedback, *Climate Dynamics*, *42*(1-2), 69–81, doi:10.1007/s00382-013-1774-0.

Randerson, J. T., F. M. Hoffman, P. E. Thornton, N. M. Mahowald, K. Lindsay, Y.-H. Lee, C. D. Nevison, S. C. Doney, G. Bonan, R. Stöckli, C. Covey, S. W. Running, and I. Y. Fung (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate–carbon models, *Global Change Biology*, *15*(10), 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x.

Rasmussen, M., A. Hastings, M. J. Smith, F. B. Agusto, B. M. Chen-Charpentier, F. M. Hoffman, J. Jiang, K. E. O. Todd-Brown, Y. Wang, Y.-P. Wang, and Y. Luo (2016), Transit times and mean ages for nonautonomous and autonomous compartmental systems, *Journal of Mathematical Biology*, *73*(6), 1379–1398, doi:10.1007/s00285-016-0990-8.

Rayner, P. J., R. M. Law, C. E. Allison, R. J. Francey, C. M. Trudinger, and C. Pickett-Heaps (2008), Interannual variability of the global carbon cycle (1992–2005) inferred by inversion of atmospheric CO_2 and $\delta^{13}CO_2$ measurements, *Global Biogeochemical Cycles*, *22*(3), doi:10.1029/2007GB003068.

Reager, J. T., A. S. Gardner, J. S. Famiglietti, D. N. Wiese, A. Eicker, and M.-H. Lo (2016), A decade of sea level rise slowed by climate-driven hydrology, *Science*, *351*(6274), 699–703, doi:10.1126/science.aad8386.

Reichstein, M., and C. Beer (2008), Soil respiration across scales: The importance of a model–data integration framework for data interpretation, *Journal of Plant Nutrition and Soil Science*, *171*(3), 344–354, doi:10.1002/jpln.200700075.

Riley, W. J., F. Maggi, M. Kleber, M. S. Torn, J. Y. Tang, D. Dwivedi, and N. Guerry (2014), Long residence times of rapidly decomposable soil organic matter: application of a multi-phase, multi-component, and vertically resolved model (BAMS1) to soil carbon dynamics, *Geoscientific Model Development*, *7*(4), 1335–1355, doi:10.5194/gmd-7-1335-2014.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin*, J. P. Walker, D. Lohmann, and D. Toll (2004), The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, *85*(3), 381–394, doi:10.1175/BAMS-85-3-381.

Sanderman, J., T. Hengl, and G. J. Fiske (2017), Soil carbon debt of 12,000 years of human land use, *Proceedings of the National Academy of Sciences*, *114*(36), 9575–9580, doi:10.1073/pnas.1706103114.

Santanello, J. A., C. D. Peters-Lidard, S. V. Kumar, C. Alonge, and W.-K. Tao (2009), A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales, *Journal of Hydrometeorology*, *10*(3), 577–599, doi:10.1175/2009JHM1066.1.

Schimel, D., B. B. Stephens, and J. B. Fisher (2015), Effect of increasing CO_2 on the terrestrial carbon cycle, *Proceedings of the National Academy of Sciences*, *112*(2), 436–441, doi:10.1073/pnas.1407302112.

Schmidt, M. W. I., M. S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I. A. Janssens, M. Kleber, I. Kogel-Knabner, J. Lehmann, D. A. C. Manning, P. Nannipieri, D. P. Rasse, S. Weiner, and S. E. Trumbore (2011), Persistence of soil organic matter as an ecosystem property, *Nature*, *478*(7367), 49–56, doi:10.1038/nature10386.

Schrumpf, M., and K. Kaiser (2015), Large differences in estimates of soil organic carbon turnover in density fractions by using single and repeated radiocarbon inventories, *Geoderma*, *239-240*, 168–178, doi:https://doi.org/10.1016/j.geoderma.2014.09.025.

Seneviratne, S. I., R. D. Koster, Z. Guo, P. A. Dirmeyer, E. Kowalczyk, D. Lawrence, P. Liu, D. Mocko, C.-H. Lu, K. W. Oleson, and D. Verseghy (2006), Soil moisture memory in AGCM simulations: Analysis of Global Land–Atmosphere Coupling Experiment (GLACE) data, *Journal of Hydrometeorology*, *7*(5), 1090–1112, doi:10.1175/JHM533.1.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture–climate interactions in a changing climate: A review, *Earth-Science Reviews*, *99*(3-4), 125–161, doi:10.1016/j.earscirev.2010.02.004.

Seneviratne, S. I., M. Wilhelm, T. Stanelle, B. van den Hurk, S. Hagemann, A. Berg, F. Cheruy, M. E. Higgins, A. Meier, V. Brovkin, M. Claussen, A. Ducharne, J.-L. Dufresne, K. L. Findell, J. Ghattas, D. M. Lawrence, S. Malyshev, M. Rummukainen, and B. Smith (2013), Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, *Geophysical Research Letters*, *40*(19), 5212–5217, doi:10.1002/grl.50956.

Sherwood, S. C., S. Bony, O. Boucher, C. Bretherton, P. M. Forster, J. M. Gregory, and B. Stevens (2015), Adjustments in the forcing-feedback framework for understanding climate change, *Bulletin of the American Meteorological Society*, *96*(2), 217–228, doi:10.1175/BAMS-D-13-00167.1.

Shevliakova, E., S. W. Pacala, S. Malyshev, G. C. Hurtt, P. C. D. Milly, J. P. Caspersen, L. T. Sentman, J. P. Fisk, C. Wirth, and C. Crevoisier (2009), Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, *Global Biogeochemical Cycles*, *23*(2), GB2022, doi:10.1029/2007GB003176.

Shi, Z., J. T. Randerson, S. D. Allison, P. A. Levine, and S. E. Trumbore (2019), Limited carbon sequestration potential of global soils implied by radiocarbon observations, *In prep.*

Sierra, C. A., A. M. Hoyt, Y. He, and S. E. Trumbore (2018), Soil organic matter persistence as a stochastic process: Age and transit time distributions of carbon in soils, *Global Biogeochemical Cycles*, *32*(10), 1574–1588, doi:10.1029/2018GB005950.

Skjemstad, J., L. Spouncer, B. Cowie, and R. Swift (2004), Calibration of the Rothamsted organic carbon turnover model (RothC ver. 26.3), using measurable soil organic carbon pools, *Australian Journal of Soil Research*, *42*(1), 79–88, doi:10.1071/SR03013.

Soden, B. J., and I. M. Held (2006), An assessment of climate feedbacks in coupled ocean–atmosphere models, *Journal of Climate*, *19*(14), 3354–3360, doi:10.1175/JCLI3799.1.

Spennemann, P. C., and A. C. Saulo (2015), An estimation of the land-atmosphere coupling strength in South America using the Global Land Data Assimilation System, *International Journal of Climatology*, *35*(14), 4151–4166, doi:10.1002/joc.4274.

Stockmann, U., M. A. Adams, J. W. Crawford, D. J. Field, N. Henakaarchchi, M. Jenkins, B. Minasny, A. B. McBratney, V. de Remy de Courcelles, K. Singh, I. Wheeler, L. Abbott, D. A. Angers, J. Baldock, M. Bird, P. C. Brookes, C. Chenu, J. D. Jastrow, R. Lal, J. Lehmann, A. G. O'Donnell, W. J. Parton, D. Whitehead, and M. Zimmermann (2013), The knowns, known unknowns and unknowns of sequestration of soil organic carbon, *Agriculture, Ecosystems & Environment*, *164*, 80–99, doi:https://doi.org/10.1016/j.agee.2012.10.001.

Sulman, B. N., R. P. Phillips, A. C. Oishi, E. Shevliakova, and S. W. Pacala (2014), Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO_2, *Nature Climate Change*, *4*, 1099–1102.

Sun, S., and G. Wang (2012), The complexity of using a feedback parameter to quantify the soil moisture-precipitation relationship, *Journal of Geophysical Research: Atmospheres*, *117*(D11), D11113, doi:10.1029/2011JD017173.

Sun, S., and G. Wang (2013), Climate variability attributable to terrestrial and oceanic forcing in the NCAR CAM3-CLM3 models, *Climate Dynamics*, *42*(7), 2067–2078, doi: 10.1007/s00382-013-1913-7.

Sun, Y., L. Gu, R. E. Dickinson, R. J. Norby, S. G. Pallardy, and F. M. Hoffman (2014), Impact of mesophyll diffusion on estimated global land CO_2 fertilization, *Proceedings of the National Academy of Sciences*, *111*(44), 15,774–15,779, doi:10.1073/pnas.1418075111.

Sun, Y., C. Frankenberg, M. Jung, J. Joiner, L. Guanter, P. Köhler, and T. Magney (2018), Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP, *Remote Sensing of Environment*, *209*, 808–823, doi:10.1016/j.rse.2018.02.016.

Susskind, J., J. M. Blaisdell, and L. Iredell (2014), Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm, *Journal of Applied Remote Sensing*, *8*(1), 084994, doi:10.1117/1.JRS.8.084994.

Swann, A. L. S., and C. D. Koven (2017), A direct estimate of the seasonal cycle of evapotranspiration over the Amazon basin, *Journal of Hydrometeorology*, *18*(8), 2173–2185, doi: 10.1175/JHM-D-17-0004.1.

Swann, A. L. S., F. M. Hoffman, C. D. Koven, and J. T. Randerson (2016), Plant responses to increasing CO_2 reduce estimates of climate impacts on drought severity, *Proceedings of the National Academy of Sciences*, *113*(36), 10,019–10,024, doi:10.1073/pnas.1604581113.

Swenson, S., and J. Wahr (2006), Estimating large-scale precipitation minus evapotranspiration from grace satellite gravity measurements, *Journal of Hydrometeorology*, *7*(2), 252–270, doi: 10.1175/JHM478.1.

Swenson, S. C. (2012), GRACE monthly land water mass grids NetCDF release 5.0, Ver. 5.0, PO.DAAC, CA, USA, doi:10.5067/TELND-NC005.

Swenson, S. C., and D. M. Lawrence (2014), Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data, *Journal of Geophysical Research: Atmospheres*, *119*(17), 10,299–10,312, doi:10.1002/2014JD022314.

Swenson, S. C., and D. M. Lawrence (2015), A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model, *Water Resources Research*, *51*(11), 8817–8833, doi: 10.1002/2015WR017582.

Tang, J., and W. J. Riley (2014), Weaker soil carbon-climate feedbacks resulting from microbial and abiotic interactions, *Nature Clim. Change*, *5*, 56–60, doi:10.1038/nclimate2438.

Tang, J., W. J. Riley, and J. Niu (2015), Incorporating root hydraulic redistribution in CLM4.5: Effects on predicted site and global evapotranspiration, soil moisture, and water storage, *Journal of Advances in Modeling Earth Systems*, *7*(4), 1828–1848, doi:10.1002/2015MS000484.

Taylor, C. M., R. A. M. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo (2012), Afternoon rain more likely over drier soils, *Nature*, *489*(7416), 423–426, doi:10.1038/nature11377.

Taylor, C. M., C. E. Birch, D. J. Parker, N. Dixon, F. Guichard, G. Nikulin, and G. M. S. Lister (2013), Modeling soil moisture-precipitation feedback in the Sahel: Importance of spatial scale versus convective parameterization, *Geophysical Research Letters*, *40*(23), 6213–6218, doi:10.1002/2013GL058511.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192, doi:10.1029/2000JD900719.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2011), An overview of CMIP5 and the experiment design, *Bulletin of the American Meteorological Society*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

Teixeira, J., D. Waliser, R. Ferraro, P. Gleckler, T. Lee, and G. Potter (2014), Satellite observations for CMIP5: The genesis of Obs4MIPs, *Bulletin of the American Meteorological Society*, *95*(9), 1329–1334, doi:10.1175/BAMS-D-12-00204.1.

Terai, C. R., P. M. Caldwell, S. A. Klein, Q. Tang, and M. L. Branstetter (2017), The atmospheric hydrologic cycle in the ACME v0.3 model, *Climate Dynamics*, *50*(9-10), 3251–3279, doi:10.1007/s00382-017-3803-x.

Tifafi, M., M. Camino-Serrano, C. Hatté, H. Morras, L. Moretti, S. Barbaro, S. Cornu, and B. Guenet (2018a), The use of radiocarbon ^{14}C to constrain carbon dynamics in the soil module of the land surface model ORCHIDEE (SVN r5165), *Geoscientific Model Development*, *11*(12), 4711–4726, doi:10.5194/gmd-11-4711-2018.

Tifafi, M., B. Guenet, and C. Hatté (2018b), Large differences in global and regional total soil carbon stock estimates based on SoilGrids, HWSD, and NCSCD: Intercomparison and evaluation based on field data from USA, England, Wales, and France, *Global Biogeochemical Cycles*, *32*(1), 42–56, doi:10.1002/2017GB005678.

Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D. Allison (2013), Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, *Biogeosciences*, *10*(3), 1717–1736, doi:10.5194/bg-10-1717-2013.

Todd-Brown, K. E. O., J. T. Randerson, F. Hopkins, V. Arora, T. Hajima, C. Jones, E. Shevliakova, J. Tjiputra, E. Volodin, T. Wu, Q. Zhang, and S. D. Allison (2014), Changes in soil organic carbon storage predicted by Earth system models during the 21st century, *Biogeosciences*, *11*(8), 2341–2356, doi:10.5194/bg-11-2341-2014.

Todd-Brown, K. E. O., B. Zheng, and T. W. Crowther (2018), Field-warmed soil carbon changes imply high 21st-century modeling uncertainty, *Biogeosciences*, *15*(12), 3659–3671, doi:10.5194/bg-15-3659-2018.

Trenberth, K. E., J. M. Caron, D. P. Stepaniak, and S. Worley (2002), Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures, *Journal of Geophysical Research: Atmospheres*, *107*(D8), AAC 5–1–AAC 5–17, doi:10.1029/2000JD000298.

Trumbore, S. (2009), Radiocarbon and soil carbon dynamics, *Annual Review of Earth and Planetary Sciences*, *37*(1), 47–66, doi:10.1146/annurev.earth.36.031207.124300.

Tyrrell, N. L., D. Dommenget, C. Frauen, S. Wales, and M. Rezny (2014), The influence of global sea surface temperature variability on the large-scale land surface temperature, *Climate Dynamics*, *44*(7), 2159–2176, doi:10.1007/s00382-014-2332-0.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophysical Research Letters*, *31*(11), L11501, doi:10.1029/2004GL019779, l11501.

Wang, J., N. Zeng, and M. Wang (2016a), Interannual variability of the atmospheric CO_2 growth rate: roles of precipitation and temperature, *Biogeosciences*, *13*(8), 2339–2352, doi:10.5194/bg-13-2339-2016.

Wang, W., P. Ciais, R. R. Nemani, J. G. Canadell, S. Piao, S. Sitch, M. A. White, H. Hashimoto, C. Milesi, and R. B. Myneni (2013), Variations in atmospheric CO_2 growth rates coupled with tropical temperature, *Proceedings of the National Academy of Sciences*, *110*(32), 13,061–13,066, doi:10.1073/pnas.1219683110.

Wang, X., S. Piao, P. Ciais, P. Friedlingstein, R. B. Myneni, P. Cox, M. Heimann, J. Miller, S. Peng, T. Wang, H. Yang, and A. Chen (2014), A two-fold increase of carbon cycle sensitivity to tropical temperature variations, *Nature*, *506*(7487), 212–215, doi:10.1038/nature12915.

Wang, Y., Z. Xie, and B. Jia (2016b), Incorporation of a dynamic root distribution into CLM4.5: Evaluation of carbon and water fluxes over the Amazon, *Advances in Atmospheric Sciences*, *33*(9), 1047–1060, doi:10.1007/s00376-016-5226-8.

Watkins, M. M., D. N. Wiese, D.-N. Yuan, C. Boening, and F. W. Landerer (2015), Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648–2671, doi:10.1002/2014JB011547.

Wehrli, K., B. P. Guillod, M. Hauser, M. Leclair, and S. I. Seneviratne (2018), Assessing the dynamic versus thermodynamic origin of climate model biases, *Geophysical Research Letters*, *45*(16), 8471–8479, doi:10.1029/2018GL079220.

Wei, J., R. E. Dickinson, and H. Chen (2008), A negative soil moisture–precipitation relationship and its causes, *Journal of Hydrometeorology*, *9*(6), 1364–1376, doi:10.1175/2008JHM955.1.

Wenzel, S., P. M. Cox, V. Eyring, and P. Friedlingstein (2014), Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth system models, *Journal of Geophysical Research: Biogeosciences*, *119*(5), 794–807, doi:10.1002/2013JG002591.

Whan, K., J. Zscheischler, R. Orth, M. Shongwe, M. Rahimi, E. O. Asare, and S. I. Seneviratne (2015), Impact of soil moisture on extreme maximum temperatures in Europe, *Weather and Climate Extremes*, *9*, 57–67, doi:10.1016/j.wace.2015.05.001.

Wiese, D. N., D.-N. Yuan, C. Boening, F. W. Landerer, and M. M. Watkins (2015), JPL GRACE Mascon Ocean, Ice, and Hydrology Equivalent Water Height JPL RL05M.1., Ver. 1., PO.DAAC,CA,USA, doi:10.5067/TEMSC-OCL05.

Winkler, A. J., R. B. Myneni, G. A. Alexandrov, and V. Brovkin (2019), Earth system models underestimate carbon fixation by plants in the high latitudes, *Nature Communications*, *10*(1), 885, doi:10.1038/s41467-019-08633-z.

Woodwell, G. M., F. T. Mackenzie, R. A. Houghton, M. Apps, E. Gorham, and E. Davidson (1998), Biotic feedbacks in the warming of the Earth, *Climatic Change*, *40*(3), 495–518, doi:10.1023/A:1005345429236.

Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bulletin of the American Meteorological Society*, *78*(11), 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Yin, L., R. Fu, E. Shevliakova, and R. E. Dickinson (2013), How well can CMIP5 simulate precipitation and its controlling processes over tropical South America?, *Climate Dynamics*, *41*(11-12), 3127–3143, doi:10.1007/s00382-012-1582-y.

Zeng, N., A. Mariotti, and P. Wetzel (2005), Terrestrial mechanisms of interannual CO_2 variability, *Global Biogeochemical Cycles*, *19*(1), GB1016, doi:10.1029/2004GB002273.

Zeng, X., M. Barlage, C. Castro, and K. Fling (2010), Comparison of land–precipitation coupling strength using observations and models, *Journal of Hydrometeorology*, *11*(4), 979–994, doi:10.1175/2010JHM1226.1.

Zhu, Q., W. J. Riley, and J. Tang (2017), A new theory of plant–microbe nutrient competition resolves inconsistencies between observations and model predictions, *Ecological Applications*, *27*(3), 875–886, doi:10.1002/eap.1490.

Zimmermann, M., J. Leifeld, M. W. I. Schmidt, P. Smith, and J. Fuhrer (2007), Measured soil organic matter fractions can be related to pools in the RothC model, *European Journal of Soil Science*, *58*(3), 658–667, doi:10.1111/j.1365-2389.2006.00855.x.