Investigating the impact of agricultural fire management practices on the terrestrial carbon cycle

Sam Sidwell Rabin

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Ecology and Evolutionary Biology
Adviser: Stephen W. Pacala

May 2016

Abstract

In this dissertation, I explore the role that agricultural fire management plays in global patterns of vegetation fire and thus carbon cycling between ecosystems and the atmosphere. No estimates previously existed of the amount of fire associated with pasture and rangelands at the global scale, and so my first chapter details the development of a dataset separating the influences of cropland, pasture, and non-agricultural land on fire activity. Pasture turns out to be responsible for nearly half of all the area that burns every year, and often differs in frequency and seasonal timing from adjacent non-agricultural ecosystems.

These new observational data allowed me to design and parameterize a global fire model that, for the first time, explicitly simulates the way that modern land managers use fire on cropland and pasture. The information about non-agricultural fire also allowed me to construct the first model for burning on non-agricultural land based on observations with the influence of pasture burning – which is governed by wholly different processes – removed. Chapter 2 details the structure and performance of this new fire and vegetation model. After calibration against observations using a new, automated method, the model is shown to successfully reproduce the general global patterns of fire activity.

In my third chapter, I use the model to test whether pasture burning practices have any effect on terrestrial carbon cycling. Significantly more carbon emissions are associated with pasture fire than would be expected if left to burn according to the same mechanisms that govern non-agricultural fire. The mean difference is even larger than the mean annual modeled net land carbon flux. Although there is a fair amount of interannual variation, and there are some areas for improvement in the model's simulation of both vegetation dynamics and fire activity, these results highlight the importance to Earth system modeling of including realistic representations of how people manage fire on agricultural lands.

Acknowledgements

No part of the work presented here could have been accomplished without the contributions – intellectual and emotional, professional and personal – of dozens of some of the best people in the world.

Briefly, I would like to thank Paul Laris for providing perspective on and Kees Klein Goldewijk for providing data for Chapter 1. I also appreciate the helpful advice Jim Randerson and Yang Chen gave me about the GFED3s data.

The Princeton EEB department has been an incredible place to develop as a scientist. I am eternally grateful to my adviser, Steve Pacala, for continually pushing me forward as a researcher and supporting me throughout my career here. I like to think that at least some of your mathematical intuition and intellectual curiosity have rubbed off on me, and I'm sure those will continue to be a critical part of my research style for the rest of my career. The mix of independence and support that you have cultivated in PacLab is exactly what I needed as a young scientist. My other committee members – Lars Hedin, Elena Shevliakova, and David Wilcove – also provided valuable input and perspectives on my research as it evolved from a general desire to study "carbon cycling and maybe biodiversity" to the focused package presented in this dissertation.

Elena Shevliakova and Sergey Malyshev, at the NOAA Geophysical Fluid Dynamics Laboratory (NOAA-GFDL), provided extensive amounts of their time and expertise towards my work on the fire model. Sergey has been especially generous, often replying to my emails late at night or on weekends in the service of helping me squash bugs or understand how LM3 actually works. I am convinced that, without him, it would have taken me six more years to finish up. Many thanks to the GFDL "B group" in general for allowing me to use so much processing time and trusting me to work on such an important piece of code; the collaboration between GFDL and Princeton is an amazing resource and one that I know will continue to produce some of the best Earth system science – and scientists – in the world. That collaboration also brought me into contact with Brian Magi, now at the University of North Carolina at Charlotte, whose initial work towards a new fire model for LM3 got me interested in working on this problem and laid the intellectual framework that eventually resulted in this dissertation. His continual collaboration throughout my graduate career, especially his thoughtful help on the paper that is now Chapter 1, have been invaluable. My funding sources have included the National Science Foundation Graduate Research Fellowship Program, the Carbon Mitigation Initiative, and the Princeton EEB department.

While the professional environment here in Princeton is top-notch, my fellow grad students in EEB and other departments have made it into simply a great place to *be*. So many of you have helped me develop my skills and ideas during my time here, but to highlight a few: Sarah Batterman, Cleo Chou, Caroline Farrior, Adam Pellegrini, Daniel Stanton, Carla Staver, Annette Trierweiler, and other current and former members of the Pacala and Hedin lab groups all stand out in my mind. The grad students of EEB are a very social bunch, and it's been so much fun to spend time with all of you – whether playing sports (especially flag football, Ultimate, Friday soccer, and of course Coprolites softball), losing at pub trivia (Hand Turkeys forever!), birdwatching, quenching our thirst at beer hour, or just generally hanging out. I have made some of my best friends in the world during my time here. Each and every member of the 2010 cohort – Matthew Aardema, Cleo Chou, Paul Elsen, Christina Faust, Matt Lutz, Lena Senko, Jenny Schieltz, and Annette Trierweiler – have been awesome since the beginning, and I still daydream about being back on the Osa Peninsula with y'all. And last but not least, Blair and Mike Costelloe: You two have made Colonnade Pointe into a hotspot of awesomeness, and it has been my privilege to call myself your neighbor and friend.

Finally, I couldn't have gotten here or made it through without the support of my family. Mom and Dad, you always made sure I got my homework done but also encouraged me to go outside and play – in retrospect, the perfect upbringing to raise an ecologist with the work ethic required to get through a PhD. Gene, you have a job doing what you love and you're the coach of a damn roller derby team – you remind me that life's too rich to get stuck on a boring path. I have been so lucky to have the ear of my aunt Susan Solomon, a bona fide science rockstar, whose example, experience, and advice have been (and will continue to be, I'm sure) invaluable as I

have navigated along my career path. The example of my late grandfather – Col. Lowell Gene Sidwell, USAF (Ret.) – has also been a continual source of inspiration over these six years. Whenever I felt like the workload was crushing, I'd remember that I wasn't having to work on a ranch and train in ROTC to put myself through school; whenever I got nervous about a big committee meeting or presentation, I'd remember that I wasn't piloting a bomber over enemy territory through clouds of flak – let alone doing it 49 times. Finally, the Maslankas and Hendricksons have given me nothing but love since I first came down to visit Pensacola. Donna, Marie, Lance, and Jim – I'm so glad to be a part of your family, and so fortunate that you raised such a great daughter.

But the most important person throughout my graduate career – the most important person in my life – has been my wife Laura. You have been unbelievably patient with me as grad school has brought me into my fully evolved state as absent-minded academic. You have understood when I needed to work during the brief times we had together on weeknights and even on the weekends, and coming home to see you every night has kept me sane. Don't think I've forgotten how lucky I am to be with you. I couldn't have done this without your infinite support, and I am so excited to see where we go from here.

Contents

List of Tables

List of Figures

Introduction

Agricultural ecosystems, including cropland and pasture, together cover more than a third of Earth's surface. To fully understand the impact of humans on the Earth system, and to predict how climate and ecosystems can be expected to change in the future, it is critical to understand how agriculture interacts with other parts of the Earth system. The scientific community has dedicated much effort to developing Earth system modeling approaches that allow such questions to be explored, but the way people manage fire on cropland and pasture has been neglected at this important scale. The work presented in this dissertation represents a step forward in understanding the role such fire management plays in determining regional and global fire-vegetation interactions.

The first chapter of this dissertation presents the results of a novel method to estimate the amount of burning associated with cropland, pasture, and non-agricultural lands. This work is novel in that it separates out for the first time the separate influences of these three land cover and use types on fire patterns at large scales, and also because it quantifies the suppressive effect that can sometimes result from intensive land use and management.

The next chapter builds off the results from the first, using them to drive a new global model of cropland and pasture fire that is embedded within a dynamic global vegetation model. Fire on non-agricultural land is also simulated, the model for which is fit using an automated algorithm. Chapter 2 describes the structure and performance of this new fire model. Chapter 3 then uses the model to quantify the effect of pasture management fire on fire patterns and carbon cycling in the terrestrial biosphere.

Chapter 1

Quantifying regional, time-varying effects of cropland and pasture on vegetation fire

1.1 Abstract

The global extent of agriculture demands a thorough understanding of the ways it impacts the Earth system through the modification of both the physical and biological characteristics of the landscape as well as through emissions of greenhouse gases and aerosols. People use fire to manage cropland and pasture in many parts of the world, impacting both the timing and amount of fire. So far, much previous research into how these land uses affect fire regimes has focused on either individual small regions or global patterns at annual or decadal scales. Moreover, because pasture is not mapped globally at high resolution, the amount of fire associated with pasture has never been quantified as it has for cropland. The work presented here resolves the effects of agriculture – including pasture – on fire on a monthly basis for regions across the world, using globally gridded data on fire activity and land use at 0.25° resolution. The first global estimate of pasture-associated fire reveals that it accounts for over 40 % of annual burned area. Cropland, generally assumed to reduce fire occurrence, is shown to enhance or suppress fire at different times of year within individual regions. These results bridge important gaps in the understanding of how agriculture and associated management practices influence vegetation fire, enabling the next generation of vegetation and Earth system models more realistically incorporate these anthropogenic effects.

1.2 Introduction

Vegetation fire is a worldwide phenomenon with consequences for the biosphere, atmosphere, climate, and human health. Annual emissions of carbon (in various chemical forms) from fire have been estimated at $2.5\,\mathrm{Pg\,yr^{-1}}$ (2001–2009; Randerson et al., 2012). The radiative forcing from the black carbon emitted from fires since 1750 has been estimated to be $0.2\,\mathrm{W\,m^{-2}}$, which is about equivalent to 12 % of radiative forcing due to the accumulated anthropogenic CO_2 over the same time period (Bond et al., 2013; Myhre et al., 2013). Other gas and aerosol emissions from biomass burning can have notable impacts on atmospheric composition and regional weather (Crutzen and Andreae, 1990; Cox et al., 2008). Many ecosystems are shaped by fire (or the lack thereof): The frequency and seasonal timing of burns are integral to what is known as a fire regime, changes to which can, over time, result in shifts to different ecosystem types (Pyne et al., 1996b; Archibald et al., 2013; Scott et al., 2014). Model simulations of an Earth without fire have resulted in about twice as much forest area (Bond et al., 2005) or nearly 30 % more carbon stored in land ecosystems (Ward et al., 2012), which illustrates the important role that fire plays in the global carbon cycle.

Humans have been manipulating fire regimes for at least several thousand years, with anthropogenic influence having grown considerably since the Industrial Revolution (Marlon et al., 2008; Bowman et al., 2011; Archibald et al., 2012). People have suppressed wildfire actively to protect lives and property, and passively by creating landscapes that inhibit large-scale fire spread. Humans have also induced burning both intentionally and unintentionally (Pyne et al., 1996a; Bowman et al., 2011). Such anthropogenic influences can result in fire regimes that differ in important ways from how ecosystems would burn in the absence of humans, such as in terms of frequency, severity, and seasonality. For example, evidence suggests that burning often does not occur during the period of the year with peak flammability, likely reflecting human fire practices at local to regional scales rather than natural or even accidental ignitions (Le Page et al., 2010; Magi et al., 2012). In order to understand the changes humanity has made to fire regimes and how patterns of vegetation fire will continue into the future, we must identify and interpret the signatures of different human activities on observed fire patterns.

One widespread example of humans' influence on fire regimes is prescribed burning for agricultural management. Farmers may use fire to prepare fields for planting or to dispose of waste after harvest (Yevich and Logan, 2003); pastoralists can burn to enhance forage nutrient content or prevent woody encroachment (Uhl and Buschbacher, 1985). The presence of cropland or heavily grazed pasture can also reduce fire in the surrounding landscape by limiting fire spread (Archibald et al., 2009; Andela and van der Werf, 2014). Land managers sometimes take advantage of a similar effect by burning small patches of land surrounding their property, reducing the chances that a burn could spread into their fields (Laris, 2002). Fire amplification can happen as well, with agricultural management fires spreading onto non-agricultural lands. The total worldwide influence of these and other effects of agriculture on vegetation fire is poorly understood, even though cropland and pasture respectively accounted for 11 and 24 % of the Earth's land area at the beginning of this century (Klein Goldewijk et al., 2010).

Dynamic global vegetation models and Earth system models often include process-based simulations of vegetation fires (e.g., Lenihan et al., 1998; Arora and Boer, 2005; Thonicke et al., 2010). Human influence is usually included as a function of population density (Venevsky et al., 2002; Pechony and Shindell, 2009), although some authors have noted that such relationships are too simplistic, with the effect of population density actually varying based on biome or amount and type of land use (Bistinas et al., 2013). Recent work has included the suppressive effect associated with cropland through landscape fragmentation (Pfeiffer et al., 2013; Le Page et al., 2015). These effects of humans in global models are based on analyses done at the scale of individual regions (e.g., Archibald et al., 2009) or the entire globe (e.g., Bistinas et al., 2014). Bistinas et al. (2014), for example, found that fire is negatively correlated with cropland but positively correlated with pasture, taking into account a number of other variables. Such findings, however, do not fully capture the complexity and multitude of effects that managed ecosystems can have on fire. It is possible, for instance, that farmers in some part of the world might burn cropland during an otherwise fire-free season, but that in drier parts of the year cropland could fragment the burnable landscape and thus have a suppressive effect on fire. Remote sensing data from satellites can partially fill in such gaps: Estimates of burning on different land cover types are generated by overlaying fire activity data with maps of land-use and vegetation type, including cropland, which are produced by some of the same satellites (Korontzi et al., 2006; Giglio et al., 2010). For example, such estimates were used by Li et al. (2013) to incorporate cropland burning into a global fire model. However, because pasture has not been mapped by satellite as cropland has, no global estimates of pasture burning have ever been produced. This means that estimates of pasture and non-agricultural fire are entangled in global data sets, and thus observations have not distinguished what may be important differences in fire regime. To understand the total effect of agricultural management on fire occurrence, then, the scientific community must go beyond estimates of cropland burned area and associated emissions.

The work presented here is an effort to bridge these gaps in our knowledge. We present a method that uses fire observations in conjunction with estimates of land-use distribution to statistically estimate the amount of fire associated with cropland, pasture, and other lands at global and regional scales. In addition to examining the total area of such burning, the same method is used to investigate patterns of associated carbon emissions.

1.3 Methods

1.3.1 Analytical technique

Magi et al. (2012) analyzed seasonal patterns of agricultural burning (i.e., combined cropland and pasture) from non-agricultural burning using estimates of land-use distributions and satellite-derived fire data. This study builds upon the methods presented by Magi et al. (2012), differentiating between cropland, pasture, and other burning and generating estimates of the amount of each type of fire in terms of both burned area and carbon emissions.

The total amount of burned area in some gridcell i (B_i) can be represented as the sum of the burned area on each land-use type k. This can in turn be represented as the product of the area of that land cover type in the gridcell ($A_{k,i}$) and the fraction of that land-use type that burned in that gridcell ($F_{k,i}$):

$$B_i = F_{c,i}A_{c,i} + F_{p,i}A_{p,i} + F_{o,i}A_{o,i}, \tag{1.1}$$

where the subscripts c, p, and o refer to cropland, pasture, and other land, respectively. The values of each $F_{k,i}$ are unknown, but a best-guess $\widehat{F_k}$ can be estimated across a group of N gridcells:

$$\begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_i \\ \vdots \\ B_N \end{bmatrix} = \begin{bmatrix} A_{\mathrm{c1}} & A_{\mathrm{p1}} & A_{\mathrm{o1}} \\ A_{\mathrm{c2}} & A_{\mathrm{p2}} & A_{\mathrm{o2}} \\ \vdots & \vdots & \vdots \\ A_{\mathrm{ci}} & A_{\mathrm{pi}} & A_{\mathrm{oi}} \\ \vdots & \vdots & \vdots \\ A_{cN} & A_{pN} & A_{oN} \end{bmatrix} \times \begin{bmatrix} \widehat{F_c} \\ \widehat{F_p} \\ \widehat{F_o} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_N \end{bmatrix} \tag{1.2}$$

$$\boldsymbol{B} = \mathbf{A}\widehat{\boldsymbol{F}} + \boldsymbol{\epsilon}, \tag{1.3}$$

where ϵ_i is the residual for gridcell i. The set of $\widehat{F_k}$ values that minimize the sum of squared errors across a large number of gridcells can be calculated using

$$\widehat{\boldsymbol{F}} = (\mathbf{A}^{\intercal}\mathbf{A})^{-1}\mathbf{A}^{\intercal}\boldsymbol{B}, \tag{1.4}$$

where $\mathbf{A}$ and $\boldsymbol{B}$ are observations of land-use distributions and burned area, respectively. We have observed that a number of $\widehat{F_k}$ values are found to be negative. This has two possible interpretations. One is that negative $\widehat{F_k}$ values are simply a statistical artifact of the analysis without physical meaning, and that such lands burn either very little or not at all. The other possibility is that negative $\widehat{F_k}$ values represent a real aspect of fire occurrence: namely, that the negative influence of such land covers on other land covers outweighs any fire happening on the land cover itself. This could be considered to represent either active suppression to protect high-value land such as crop fields, and/or to reflect the widely documented role of anthropogenic land covers (especially cropland) in fragmenting the burnable landscape (Archibald et al., 2009; Andela and van der Werf, 2014; Hantson et al., 2015).

For the purposes of illustration, consider a hypothetical gridcell for which the analysis estimates $5\,\mathrm{km}^2$ of burned area for cropland:

$$\widehat{F_c}A_{c,i} = 5, \tag{1.5}$$

$$B_i = \widehat{F_c} A_{c,i} + \widehat{F_p} A_{p,i} + \widehat{F_o} A_{o,i}. \tag{1.6}$$

A different gridcell with equal $\widehat{F_k}$ values and twice the area of cropland but the same amounts of pasture and other land would have 5 km^2 more burning estimated:

$$\widehat{F_c}(2A_{c,i}) + \widehat{F_p} A_{p,i} + \widehat{F_o} A_{o,i} = B_i + \widehat{F_c} A_{c,i} = B_i + 5. \tag{1.7}$$

The same logic shows that there would be less fire in the second gridcell if $\widehat{F_c}$ were negative.

Conversely, $\widehat{F_k}$ values could also incorporate positive effects of one land-use type on the others. For example, much of the fire observed in the frontier of the Amazon rainforest is associated with land management burning that escapes into surrounding forest (Uhl and Buschbacher, 1985; Cochrane and Schulze, 1998). The $\widehat{F_c}$ and $\widehat{F_p}$ values in that region could potentially account for this effect as well. In this conceptualization, then, $\widehat{F_k}$ values should be interpreted not as "the fraction of land use k that burns across the region" but rather as "the net effect of land use k on fire in the region, expressed as a fraction of the area of land use k in the region". That is, for every additional unit area of land use k, we expect $\widehat{F_k}$ more (if $\widehat{F_k} > 0$) or fewer (if $\widehat{F_k} < 0$) units of burning.

To clarify, imagine a region with 2000 km^2 of cropland, 3000 km^2 of pasture, and 5000 km^2 of other land. For some month, this region has $\widehat{F_c} = -0.1$, $\widehat{F_p} = 0.2$, and $\widehat{F_o} = 0.1$. The associated burned area values would be $2000 \times -0.1 = -200$ km^2 for cropland, $3000 \times 0.2 = 600$ km^2 for pasture, and $5000 \times 0.1 = 500$ km^2 for other land, for a total of 900 km^2 of burning across the region. Now imagine another region that is identical except that it contains an extra 1000 km^2 of cropland. This new region would have $3000 \times -0.1 = -300$ km^2 of burned area associated with cropland, for a total of 800 km^2 of burning across the region. The interpretation of negative cropland-associated burned area is not that some actual negative area is burning somehow but rather that, however much cropland is burning, it is preventing so much fire on pasture and/or other land that its net influence on fire in the region is negative.

The results presented in this study are explored in the main text with this latter interpretation of $\widehat{F_k}$ values in mind. Equivalent figures in Appendix A show results with $\widehat{F_k}$ restricted to positive values, essentially interpreting $\widehat{F_k}$ values as "the fraction of land use k that burns across the region".

To account for temporal variability in the total amount of fire and its distribution among different land-use types, the analysis is performed separately for each month and year. Fire patterns and practices also vary across space, so each of 132 regions is analyzed separately. This set of regions (Fig. 1.1) was created with the goal of minimizing within-region heterogeneity in terms of climate, biome, and fire extent and timing, while still including enough gridcells to ensure an adequate sample size for estimation. The final region set resulted from an iterative process whereby we performed the analysis for a candidate region set, noted areas of severe under- or overestimation, drew new region boundaries, and re-ran the analysis. The Terrestrial Ecosystems of the World map (Olson et al., 2001), agricultural distribution maps (Klein Goldewijk et al., 2010), and observations of fire extent and timing (Randerson et al., 2012) guided development of the regions map. For example, regions were designed to avoid containing multiple patches of high concentration of a land use that appeared to vary widely in seasonal timing or amount of fire. As in Magi et al. (2012), the 14 regions developed for the Global Fire Emissions Database (Giglio et al., 2006) are used to structure the discussion of the results presented here (Fig. 1.1, Table 1.1). For clarity, these will be referred to as the "GFED regions" to distinguish them from the 132 "analysis regions." In all, 4752 $\widehat{F_k}$ values are estimated per year (3 land-use types $\times$ 12 months $\times$ 132 analysis regions) from 2001 to 2009.

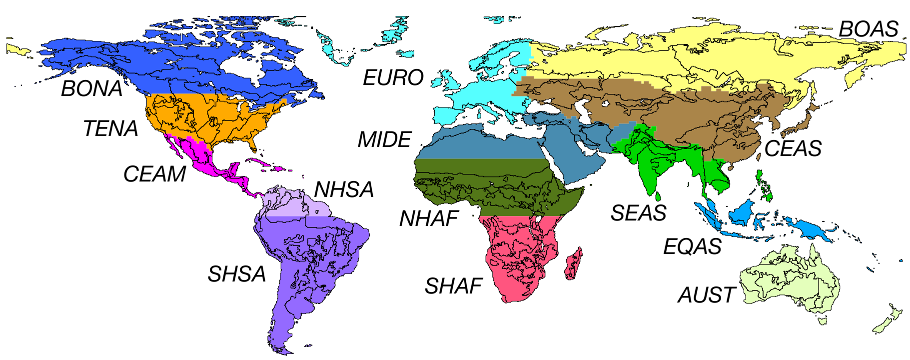

Figure 1.1: Regions used for analysis (outlines) overlaid on GFED regions (colors and labels; Giglio et al., 2006). See Table 1.1 for abbreviations.

Table 1.1: List of GFED regions and abbreviations (Giglio et al., 2006).

Abbreviation	Full name
BONA	Boreal North America
TENA	Temperate North America
CEAM	Central America
NHSA	Northern Hemisphere South America
SHSA	Southern Hemisphere South America
EURO	Europe
MIDE	Middle East
NHAF	Northern Hemisphere Africa
SHAF	Southern Hemisphere Africa
BOAS	Boreal Asia
CEAS	Central Asia
SEAS	Southeast Asia
EQAS	Equatorial Asia
AUST	Australia and New Zealand

Some restrictions were imposed on the analysis. Any land-use type whose prevalence across a region during a given year was on average less than 5 % was excluded, with the $\widehat{F_k}$ value for such land cover types taken to be zero, to avoid issues of near-singularity in the matrix calculations. Also, for region-months with no observed fire, all $\widehat{F_k}$ values were assumed to be zero.

1.3.2 Input data

Burned area and fire emissions

Observations of monthly burned area and carbon emissions at 0.25° resolution were obtained from the GFED3s data set (Randerson et al., 2012). Based on the Global Fire Emissions Database version 3 (GFED3; Giglio et al., 2010; van der Werf et al., 2010), GFED3s was designed to improve detection of small fires by incorporating an estimate of burned area based on detections of active fires outside observed fire scars. This algorithm produces an estimate of annual burned area 35 % higher than the Collection 5 MCD64A1 burned area product, which was produced using the same algorithm as most of the GFED3 data, across the time period of its coverage (2001–2010), with several large regions seeing their burned area estimates more than double (Randerson et al., 2012). Nearly a fifth of that increase occurred in croplands and cropland–natural vegetation mosaic, the estimated burned area of which increased by 123 and 79 %, respectively. Moreover, about a third occurred in savannas and grasslands, which could feasibly serve as pasture (Randerson et al., 2012). Results for cropland influence on burned area from this analysis are compared to GFED3s estimates of burned area on cropland as well as "cropland–natural mosaic," which is defined as land with "a mosaic of croplands, forests, shrubland, and grasslands in which no one component comprises more than 60 % of the landscape" (Friedl et al., 2002).

GFED3s estimates of fire-related emissions were generated, as for the original GFED3 data set (van der Werf et al., 2010), by coupling the burned area observations for each land-use type with a climate-driven vegetation model (Randerson et al., 2012). Biome-specific emissions factors combined with biomass estimates from the vegetation model then produced the amount of emissions per area burned. The analytical technique described in Sect. 2.1 can be as easily applied to emissions as it can to burned area, in which case the $\widehat{F_k}$ values represent the net effect per square kilometer of each land-use type on fire emissions. Here, an analysis of emissions of carbon-containing compounds was conducted in parallel with the analysis of burned area. a breakdown of GFED3s carbon emissions by land cover type, such as was provided for burned area, was not available.

Land use

Data on area of cropland and pasture were taken from an annualized version of the History Database of the Global Environment version 3.1 (HYDEv3.1), described by Klein Goldewijk et al. (2010). This public data set, available at 5 min spatial resolution, is the basis for the historical part of the standardized gridded land-use transitions reconstructions (Hurtt et al., 2011) used in the Coupled Model Intercomparison Project, phase 5 (Taylor et al., 2012). The publicly available data are only produced for every 5 years during the recent past, but K. Klein Goldewijk provided annual estimates for the period 2000–2009 (K. Klein Goldewijk, personal communication, 2012). Distributions are assumed to not change within years. The amount of "other" ("non-agricultural") land is calculated as the fraction of land not classified as cropland or pasture. Maps of the mean land cover distributions from HYDE for 2001–2009 can be found in Figure A.1.

Grazing land can take many different forms, including both planted forage species and naturally occurring species (often referred to as rangeland). Data from the Food and Agriculture Organization (FAO) were used in compiling maps of present-day land use in HYDE; HYDE's pasture data is based on the FAO's "permanent meadows and pastures" (Klein Goldewijk et al., 2007). These are defined as lands "used permanently (five years or more) to grow herbaceous forage crops, either cultivated or growing wild (wild prairie or grazing land)" (FAO, 2015). The term "pasture" is thus used throughout this chapter in this broad land-use sense. Note, however, that this is distinct from any given land cover type, such as grassland or savanna – that is, all pasture has herbaceous vegetation, but not all land with herbaceous vegetation is necessarily pasture.

Spatiotemporal coverage and resolution

All analyses were performed at the native resolution of GFED3s, 0.25°, with HYDE land-use data being downscaled to match. The analysis covered the period 2001–2009, as HYDE data for 2010 were not available.

1.4 Results

1.4.1 Fire extent

Every year, nearly half of all burned area is associated with agricultural lands (Fig. 1.2a): pasture contributes 203 Mha yr^{-1} of burned area, while cropland accounts for 21 Mha yr^{-1}. Non-agricultural lands are associated with 243 Mha yr^{-1} of burned area. Overall, the analysis slightly overestimated total global annual burned area, giving 467.6 Mha yr^{-1} instead of 466.9 Mha yr^{-1} (+0.2 % error).

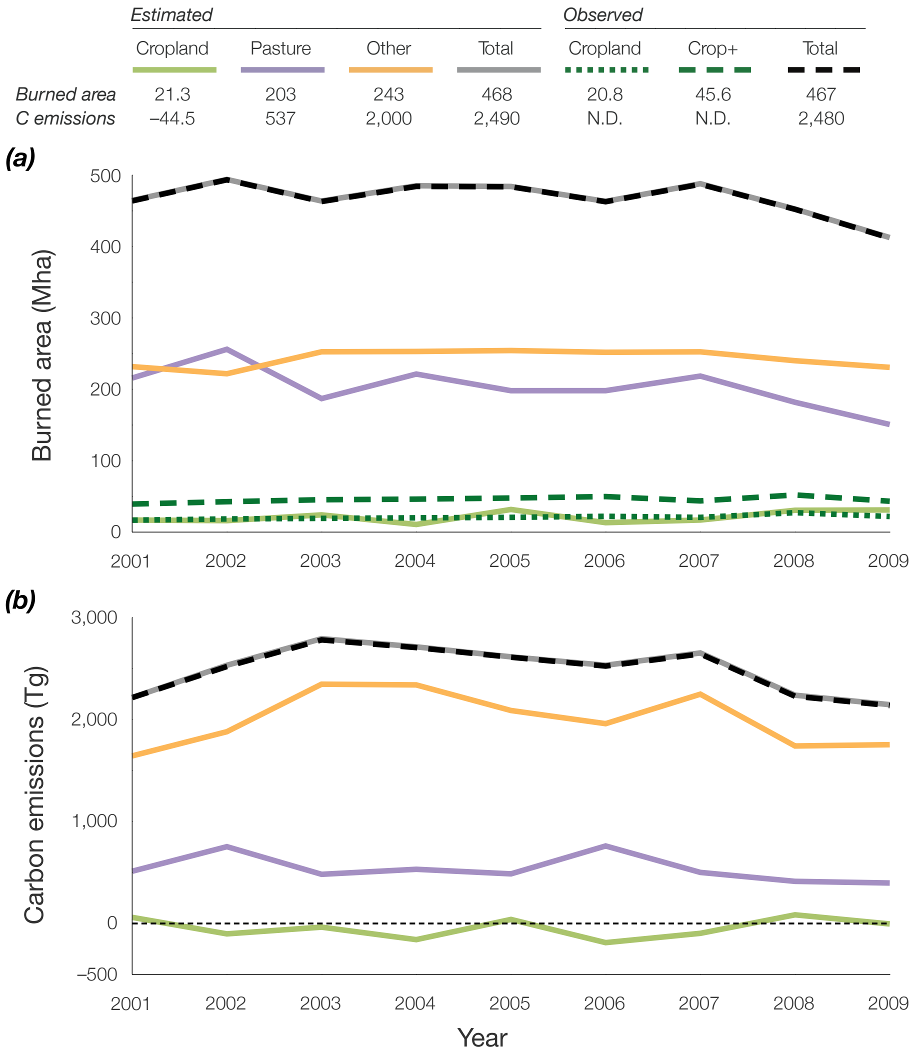

Figure 1.2: Observed and estimated annual time series of net observed and estimated global burned area (**a**; Mha) and C emissions (**b**; Tg = Mt). Numbers in table represent annual means. "N.D.": no data; "Crop+": cropland + cropland–natural mosaic. Corresponds to Fig. A.2.

The distribution of fire emissions across land-use types differs strongly from what might be expected based on their relative burned areas. Whereas annual burned area associated with non-agricultural land was only $\sim 20\,\%$ greater than that with pasture, non-agricultural land was responsible for over $260\,\%$ more fire C emissions (Fig. 1.2b). Emissions per area burned can be thought of as the product of fuel load and combustion completeness – i.e., the amount of dead and living biomass multiplied by the fraction combusted (Seiler and Crutzen, 1980). Fuel load should be higher on average for non-agricultural lands than for pasture because pastures do not have trees in densities comparable to more carbon-rich forest ecosystems. Moreover, although croplands had a net positive contribution to global burned area, they had a net negative effect on fire emissions (Fig. 1.2). This suggests that, even though less area would have burned with less cropland, the burning would be happening in more carbon-dense ecosystems. As with burned area, total global fire emissions were very slightly overestimated (by less than 0.4 %; Fig. 1.2b).

Figure 1.3 shows time series plots as in Fig. 1.2 but broken down by GFED region. Pasture can be seen to account for a sizable portion of burning in South America (NHSA and SHSA), Africa (NHAF and SHAF), central Asia (CEAS), and Australia (AUST). Overall, the algorithm reproduces the amount and interannual variability of total fire well at these large regional scales: on a scatter plot comparing the estimated and observed burned area of the 1512 GFED region-months (14 regions $\times$ 108 months), most points fall near the one-to-one line (linear regression y intercept $= -3.7 \times 10^{-3}$, slope $= 1.0008$, Pearson's $r = 0.9997$; Fig. A.3). The most apparent discrepancies compared to GFED3s occur in Europe (EURO) and the Middle East (MIDE), whose mean annual burned area totals are underestimated by $\sim$ 40 and $\sim$ 30 %, respectively. With respective mean annual observed burned areas of $\sim$ 11 200 and $\sim$ 15 800 km^2 (0.2 and 0.3 % of global fire activity), however, these are the least-burned GFED regions.

The net mean annual burned area associated with croplands, pasture, and other land is illustrated in the maps in Fig. 1.4. Pasture accounts for a large amount of burned area in the savannas of NHAF and SHAF, with NHSA, SHSA, CEAS, and AUST being highlighted to a lesser degree. Eastern Europe, northern Australia, various parts of sub-Saharan Africa, and especially India's Punjab state emerge as spots where cropland has a strong positive effect on burned area (Fig. 1.4a). Cropland has a net negative effect on burned area in other places – most notably Cambodia and southern Vietnam, Ethiopia and South Sudan, India, eastern Argentina, and southeastern Australia. These are mostly biomes where vegetation tends to be quite fire-prone, and thus where strong active and/or passive suppression due to cropland might be expected. Interestingly, pasture and non-agricultural lands are also seen to sometimes have net suppressive effects (Fig. 1.4b and c). In the case of pasture, this could be due to a passive effect – grazing pressure can reduce fuel loads, leading to slower-spreading and/or less-frequent fires (Cheney and Sullivan, 2009). Non-agricultural lands with net negative influence may result

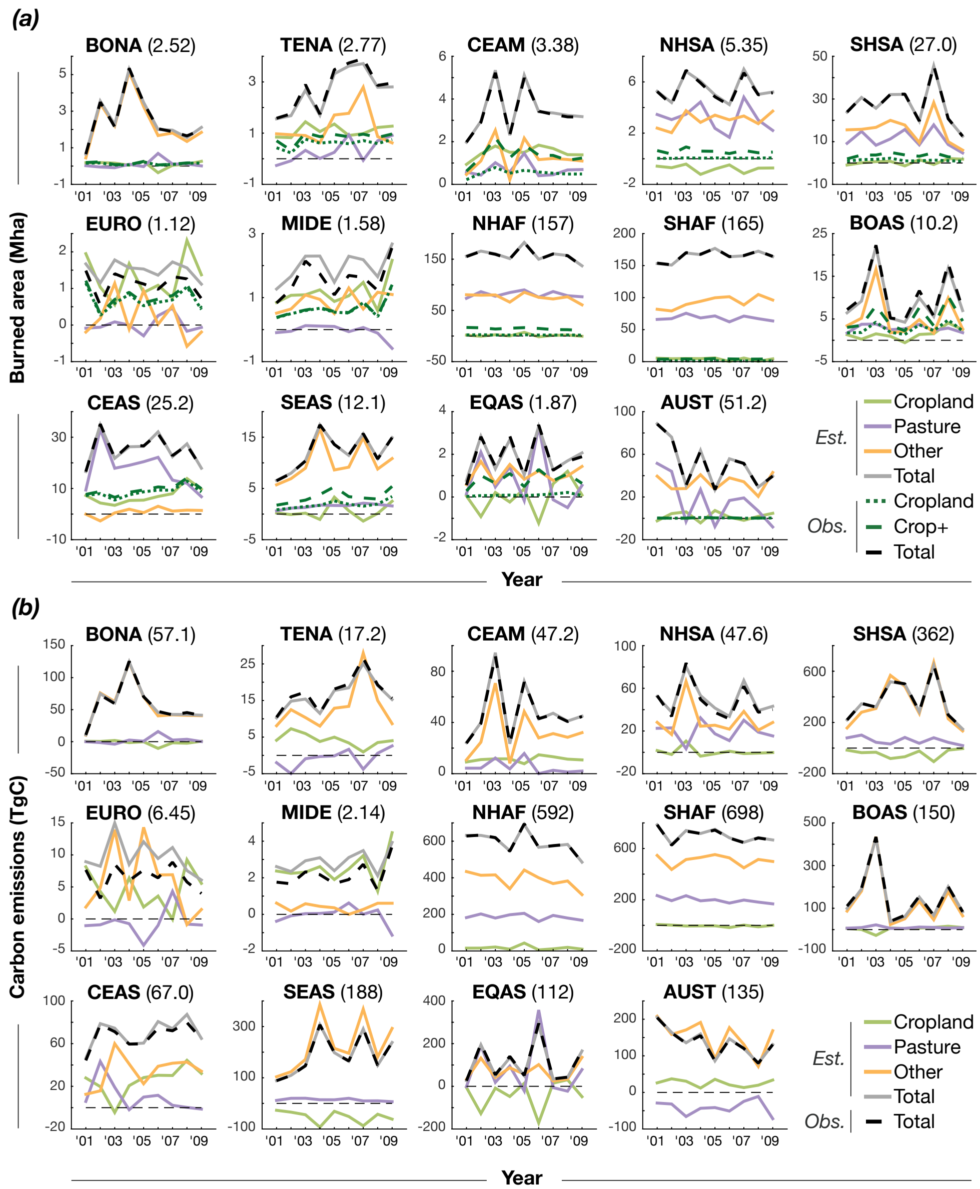

Figure 1.3: Annual time series of different fire types in each GFED region based on analysis of burned area (**a**; Mha) and C emissions (**b**; TgC). Numbers in parentheses next to region names represent mean annual observed fire there (either burned area or C emissions). "Crop+": cropland + cropland–natural mosaic. Corresponds to Fig. A.3.

from either active or passive suppression. People might use alternative management techniques to avoid fire use on cropland or pasture near valuable or protected forests, for example. Alternatively, if fire on pasture is at least to some extent unmanaged, less-flammable vegetation types such as forest or wetland could serve to break up pasture into disconnected patches and thus reduce how much it can burn. It is also important to remember that apparent negative influences might not represent any real process, being instead artifacts of this analysis (see figures in Appendix A).

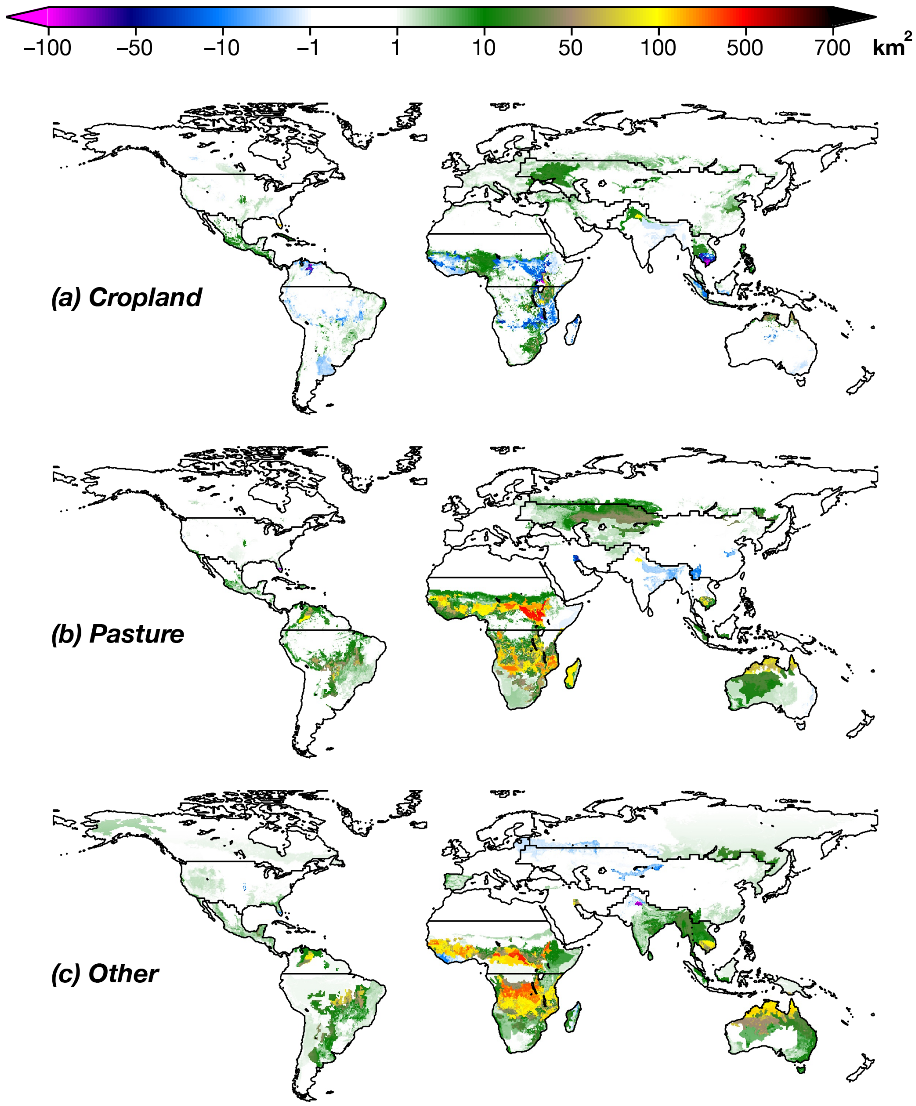

Figure 1.4: Maps of mean annual burned area (km^2) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. These are calculated from monthly maps generated by the equation $B_i = \widehat{F_k} A_{k,i}$ for each month and region. The results can be interpreted as how much more (or less) fire would be expected if the area of the given land cover were to double (and the others remain the same). Corresponds to Figure A.5. Compare with seasonal maps in Figures A.8–A.11.

Overall, the algorithm generates maps of total fire that broadly agree with the distribution of burning seen in the observations (Fig. 1.5). However, the spatial variation in burned area within regions is not fully captured; we discuss this further in Section 1.5.3.

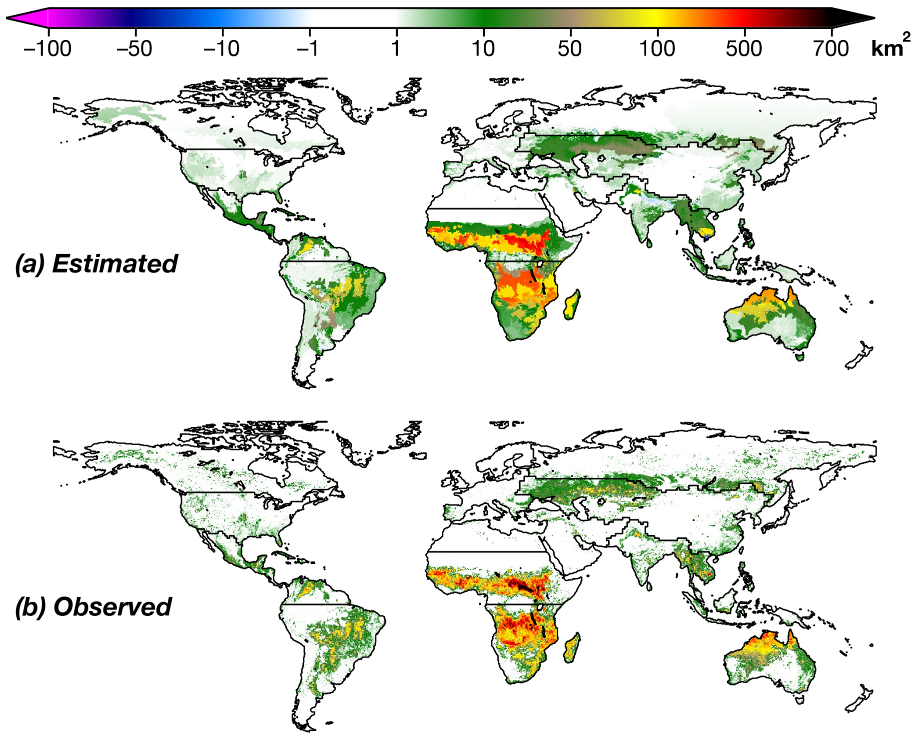

Figure 1.5: Maps of net mean annual total burned area (km^2): **(a)** estimated and **(b)** observed. Corresponds to Fig. A.6.

1.4.2 Fire timing

The previous results have shown the influence of different land-use types on fire at an annual level, but land use and management can also affect the seasonality of fire. Figure 1.6 shows, for each GFED region, the mean seasonality of estimated and observed burned area and carbon emissions as compared with observations. As expected based on the algorithm's performance with regard to annual total fire (Fig. 1.3), all regions except EURO and MIDE show good correspondence between observations and estimates of total fire.

Estimated cropland fire is sometimes higher or lower than GFED3s for cropland or cropland–natural mosaic. One reason for this is that the analysis may describe the net effect of cropland on fire, as discussed above. Another is that detection of cropland, especially of small fields, is difficult using moderate-resolution satellite imagery, such as the MODIS MCD12 data set used in GFED3s (Friedl et al., 2010). Klein Goldewijk et al. (2007), for example, had to deal with this in developing HYDE. In some regions – such as the contiguous 48 United States

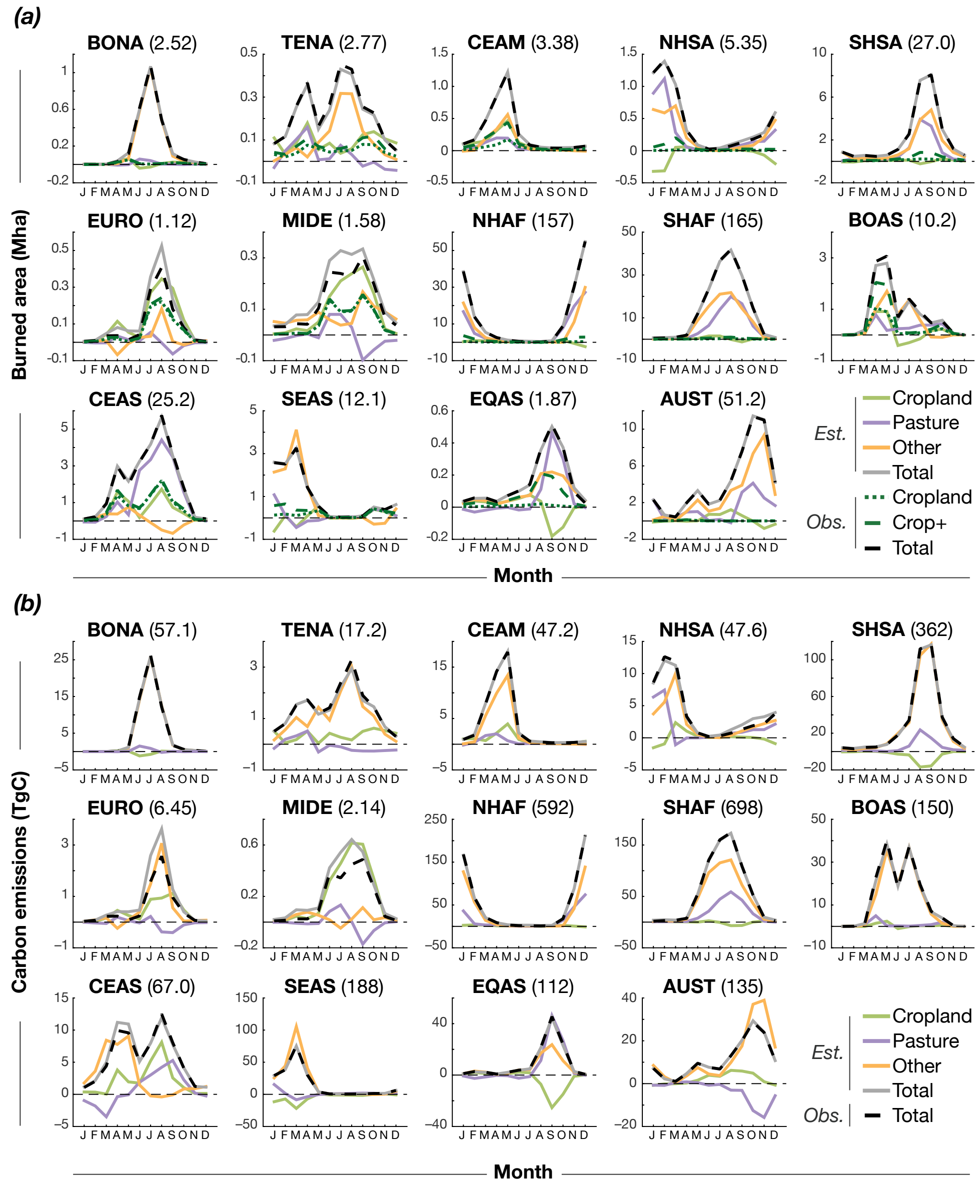

Figure 1.6: Seasonality of different fire types in each GFED region based on analysis of burned area (**a**; Mha) and C emissions (**b**; TgC). Numbers in parentheses next to region names represent mean annual observed fire there (either burned area or C emissions). Corresponds to Figure A.7.

(a.k.a. temperate North America, TENA), Europe (EURO), and central Asia (CEAS) – trends of estimated cropland burned area closely follow those from observations (Fig. 1.6). In other regions – such as Northern Hemisphere South America (NHSA) and equatorial Asia (EQAS) – cropland has an apparent negative influence on burned area for several months of the year. a comparison with observed cropland burning (of which there is little in such months) suggests that this is often a nearly pure signal of a suppressive effect. The effect appears especially strong in

EQAS during September and October, although the large amount of cropland–natural mosaic burning complicates interpretation there. Pasture sometimes has a similar effect, although rarely; this is most apparent in TENA, EURO, MIDE, and SEAS. In EURO and CEAS, even other lands sometimes have a net negative estimated burned area. As discussed above, negative influence of pasture and non-agricultural lands could reflect active and/or passive suppressive effects associated with these land-use/cover types.

Figures A.8–A.11 present another way to examine the seasonal changes in the influence of different land covers on burning. This presents an advantage over the regional time series discussed above where contrasting patterns exist within one GFED region. For example, Fig. A.9a shows that cropland is contributing to burned area in southwestern Australia from March to May, but is suppressing fire in the northern part of the continent. Figure 1.6 does not capture this pattern, instead making it appear as though cropland has no effect across the entire region of Australia and New Zealand (AUST).

The effect of different land uses on fire can be best explored and understood by examining patterns across a few regions. The savannas of western Africa have seen a good deal of remote sensing, anthropological, and ecological research regarding their fire regimes and thus provide a good example. The Sudanian savanna there experiences a distinct dry season from approximately October or November through April or May, during which it burns extensively (Laris, 2002; Kull and Laris, 2009). The fire regime is highly managed by people for agriculture and other purposes, with burning generally initiated early in the dry season and suppressed later. Early fires can have a number of benefits. For example, burning that occurs while the soil still has some residual moisture allows herbaceous regrowth, replenishing food availability for livestock ahead of the worst of the dry season (Mbow et al., 2000). Due to higher fuel moisture, these fires are also often easier to control than more intense burns under more flammable conditions later in the dry season. People often burn savanna early to fragment the burnable landscape, preventing late-season burns that can damage property and resources (Laris, 2002).

We isolated three regions (Fig. 1.7a) that mostly fall into the ecoregions "West Sudanian savanna" and "Guinean forest-savanna mosaic" according to Olson et al. (2001). Small amounts of other land cover types – including lowland and montane forests, flooded savanna, and Sahelian acacia savanna – are also included.

On average, this area sees a slight negative annual contribution of cropland to burned area – that is, cropland tends to reduce the amount of burning on pasture and other lands. Pasture contributes over a third of the observed annual burned area, with non-agricultural lands accounting for approximately twice that. Observed total burned area, which is matched almost perfectly by the estimate, peaks with pasture and non-agricultural burning in December (Fig. 1.7). As expected based on the literature on human fire management practices in this region (Mbow et al., 2000; Laris, 2002), most fire associated with pasture and non-agricultural land occurs in the early dry season – i.e., before January. Interestingly, though, the fire season for pasture seems to begin and end about a month earlier than that of non-agricultural land: from about October through January instead of November through February. Although early fire is often beneficial for all savanna in the region, the added impetus of burning early to create food for livestock appears to result in a distinct pattern. However, it is also possible that the October burning represents intentional burning of short-grass savanna, which is not actually used by livestock but may have been considered "pasture" in the land-use data (P. Laris, personal communication, 2015). An overall net suppressive effect of cropland is also evident. The strongest negative influence corresponds with both the December peak of non-cropland fire and the harvest (P. Laris, personal communication, 2015; Figs. 1.7, A.8–A.11). This emerges despite the fact that at least some cropland burning (including cropland–natural mosaic) was observed throughout the dry season (Fig. 1.7b). Even though there is some observed fire associated with cropland, then, there would be much more if cropland were replaced with pasture or non-agricultural land. This interpretation has assumed that negative values are meaningful, but similar patterns emerge using constrained $\widehat{F_k}$ values (Fig. A.12).

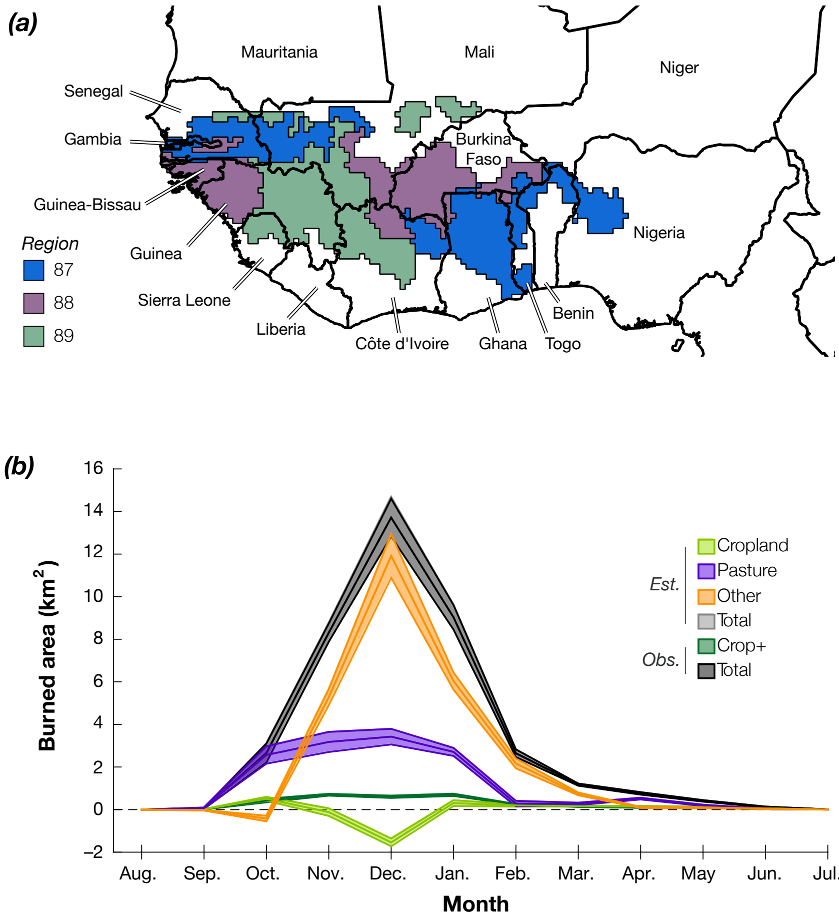

Figure 1.7: **(a)** Area included in the West African case study, color-coded by analysis region. **(b)** Mean seasonality of burned area in case study regions. Shading represents interannual variability (± 1 SEM). Note that the X axis begins in August. Corresponds to Figure A.12.

1.5 Discussion

1.5.1 First estimates of pasture-associated fire

Pasture fire accounts for about 43 % of global annual burned area and about 22 % of global C emissions from fire. Pasture burning is especially important in CEAS, NHSA, NHAF, SHAF, and SHSA, in each of which it accounts for over 40 % of annual burned area. These regions together comprise 81 % of mean annual burning. As with the global numbers, the fraction of annual fire emissions from pasture burning there is disproportionately small – only NHSA has pasture contributing more than 40 % of C emissions (Fig. 1.3b). These results are not qualitatively different in the analysis with $\widehat{F_k}$ values constrained to zero or above (Appendix A).

In most regions, the seasonality of pasture burning is roughly similar to that of non-agricultural land. a tendency for pasture to burn earlier than non-agricultural land is apparent in NHSA, EURO, MIDE, NHAF, SEAS, and to some extent AUST (Fig. 1.6). The seasonality of these two fire types is notably different in CEAS, where pasture fire peaks in August and non-agricultural fire peaks in May. During the peak of pasture burning in that region, non-agricultural land exerts a negative influence on total burning (Fig. 1.6). Some insight into the interplay of the different land-use types in this region, as well as the intricacies involved in interpreting the estimates from our method, can be gleaned from a more detailed look at pasture and other fire in CEAS. Most of the negative influence of non-agricultural land is concentrated in northern Kazakhstan and surrounding Russia. This is also the subregion where most pasture fire is concentrated during its July–August–September peak, which corresponds to the strongest negative influence of non-agricultural land. Taken together, these details suggest that there is at least some uncontrolled burning happening on pasture there at that time, since the presence of other land (presumably less-flammable vegetation types such as forest) appears to reduce pasture fire, likely by fragmenting the burnable landscape.

1.5.2 Input data quality

As with all data analysis, the performance of this algorithm is restricted by how well its input data represent the real world. Errors in the data sets of either land use or burned area will propagate through to the $\widehat{F_k}$ estimates and partitioned maps of fire by land-use type.

The first step in the development of the HYDE land-use data set was the production of a map of cropland and pasture representative of their distribution during the period 1990–2000. By reconciling remote-sensing maps of land cover with country-level area totals from the FAO, HYDE represented a significant advance over previous methods (Klein Goldewijk et al., 2007, 2010). However, the FAO numbers themselves may not be completely internally consistent, since they are compiled and reported by each country. A wide variety of ecosystem types and land-use patterns might all qualify as what the FAO terms "permanent pasture," and countries' standards of what to report likely differ (Klein Goldewijk et al., 2007). Differing methods of compilation introduce another source of uncertainty.

By incorporating active fire detections as an ancillary source of "burned area" information, the algorithm used in GFED3s was designed to avoid (as much as possible) the issue of fires much smaller than a single sensor pixel being excluded (Randerson et al., 2012). Even though GFED3s includes much more cropland fire than GFED3, it likely still misses much such burning. For example, McCarty et al. (2009) used fieldwork to inform a remote sensing estimate of cropland burning in the contiguous US and found that an average of more than 1.2 $\mathrm{Mha\,yr^{-1}}$ burned between 2003 and 2007; during the same period, GFED3s has only 0.67 $\mathrm{Mha\,yr^{-1}}$ (or 0.93 $\mathrm{Mha\,yr^{-1}}$ if also including cropland–natural mosaic). Moreover, the "small fires" improvement may not have improved the detection of burning underneath a relatively undamaged canopy, which poses a challenge even for active fire sensors and algorithms (Giglio, 2013). In regions of southern Africa with tree cover $\geq 21\,\%$, this was blamed for a 41 % underestimate of burned area in an assessment of the algorithm underlying most of GFED3 (Giglio et al., 2009); a similar assessment has not been performed for GFED3s.

1.5.3 Impacts of regional analysis

The specific set of regions chosen for this analysis can be important for the quality of the results. One aspect to consider is that analysis regions that are too extensive may encompass too many different fire patterns for any one set of $\widehat{F_k}$ values to describe well. This may have been the cause of the poor performance in EURO and MIDE with regard to total fire (Fig. 1.3): both include parts of one or more very large analysis regions (Fig. 1.1). Fire is much more frequently used to manage croplands in the eastern part of the large EURO analysis region than in the west (Lin et al., 2012). This could be due to different crops being grown, but this seems unlikely since wheat and maize comprise most of the cropland across the region (Leff et al., 2004). Instead, differences in cultural history, policies regulating residue burning, and economic

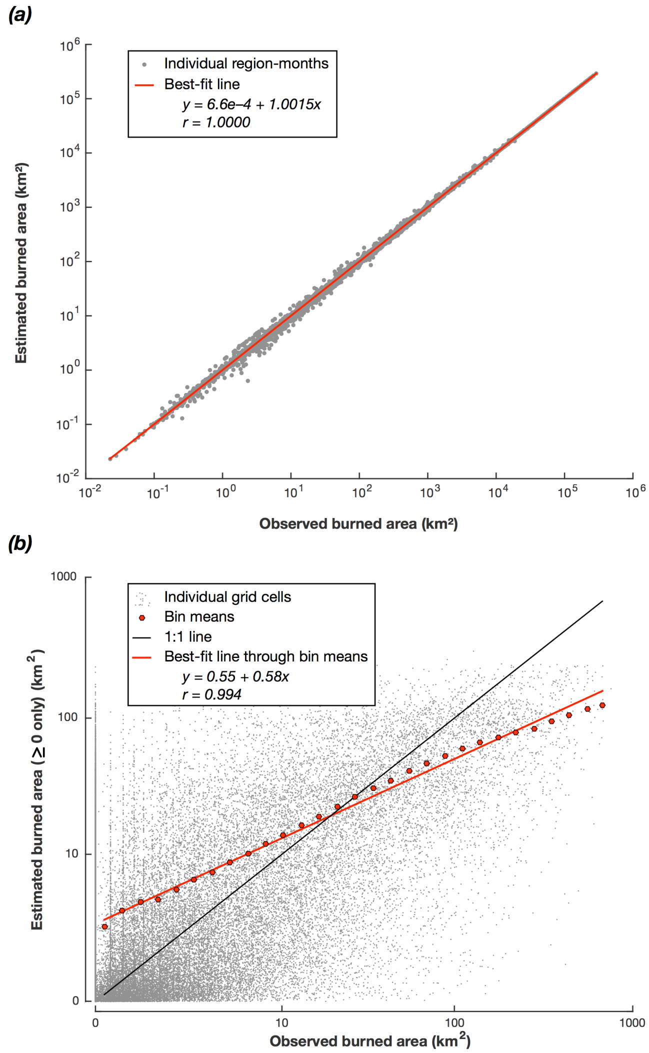

Figure 1.8: Scatter plots comparing estimated and observed total burned area. Gray points represent **(a)** each analysis region and month (region-month) or **(b)** individual gridcells ($\frac{1}{75}$ of cells chosen at random for plotting). Red lines represent the best-fit line from linear regression, with the regression in **(b)** fit to the red points, which represent mean observed and estimated values of gridcells in bins of observed burned area equally spaced along the X axis (with at least 100 gridcells required for a bin to be included). Values ≤ 0 not shown due to log-scale axes. Gridcells in region-months with no observed fire, where the analysis was not performed, were excluded from both plots and regressions. Corresponds to Figure A.13.

conditions probably play a large role. Breaking the large region into more fine-grained regions would likely better account for this heterogeneity in fire patterns and practices.

On the other hand, analysis regions that are too small – specifically, those that do not sample gridcells with a wide range of values for each land cover type – may serve to confound the results. In an extreme example, a region that had no cropland would be assigned $\widehat{F_c} = 0$. However, because no cropland was observed, the true effect cropland would have in the region might actually be different from zero. In a less extreme case, burning patterns might be controlled mostly by the influence of one dominant land cover type. This sort of effect could be at play in BOAS, for example, where (as discussed above) total regional burned area is estimated accurately despite its containing several large regions.

Another, more general consequence of the regional analysis is that spatial heterogeneity of burning within analysis regions is not well represented in the results. As expected based on the mathematics involved in the parameterization, the total estimated amount of fire at the regional level is usually quite accurate (Fig. 1.8a) – estimated total burned area was correct to within 5 % in 86 % of region-months with fire observed. A best-fit line through a plot of the total observed vs. estimated burned area of all region-months illustrates this. With a slope near one, intercept near zero, and high value of Pearson's r, most of the estimated means lie near the one-to-one line. On a finer-grained level, a best-fit line through the mean estimated burned area of bins of grid-cell-level observed burned area, equally spaced on a log scale, shows that the algorithm tends to overestimate burning where there is little observed fire and underestimate where observed burning is high (Fig. 1.8b), but the scatter of individual grid cells around these binned averages is large. Especially noticeable is the large number of gridcells with zero (or very little) observed fire that are overestimated by the algorithm. When calculated across all gridcells in all months, the coefficient of determination $R^2 = 0.356$, indicating that only just over a third of the variation in spatiotemporal patterns of fire can be explained by land-use distributions. More of the variability is due to factors governing fuel availability and moisture, such as net primary productivity, temperature, precipitation, and humidity (Bistinas et al., 2014; Lasslop et al., 2015b). In region-months where land cover distributions have very low explanatory power, the individual $\widehat{F_k}$ values should tend towards the total fraction of land burned.

The maps in Fig. 1.5 illustrate this problem in a more intuitive format. Although fire activity is usually well characterized at the level of the analysis region (as illustrated by Fig. 1.8a), Fig. 1.5 shows that it does not fully incorporate the heterogeneity evident in the observations as illustrated by Fig. 1.8b). Thus, interpretations of the maps in Fig. 1.4 should focus on general patterns without delving too deeply into gridcell by gridcell variation.

Finally, because the GFED region boundaries do not all correspond to those of the analysis regions, GFED regions without much fire are highly sensitive to inclusion of parts of analysis regions with too much or too little estimated fire. This also may have contributed to the poor performance in EURO and MIDE (Fig. 1.3). For example, Afghanistan (MIDE) is included in analysis region 26, "west-central Asian desert steppe" (AR26), which is not completely contained by MIDE. Afghanistan is an area of overestimate in AR26, and although it is balanced out by underestimates elsewhere in that region (especially along its northern boundary), MIDE only includes the overestimate. This effect, then, contributes to the net overestimate in MIDE.

1.6 Conclusions

The analysis presented here shows that agriculture does have far-reaching consequences on vegetation fire, often in ways not previously measured or considered at large scales. The widely acknowledged suppressive effect of cropland (Archibald et al., 2009; Andela and van der Werf, 2014) is quantified by broadening the scope of land-use associations with burning to include fire prevented on other land-use types. Pasture, previously not considered as a distinct land-use type in estimates of fire activity since it is not mapped globally at high resolution, is shown to account for nearly half of global annual burned area (Fig. 1.2a). Importantly, analysis at the regional and monthly level elucidates for the first time variations in management practices and other patterns across space and time. For example, although cropland has a net suppressive effect in parts of the world such as Southeast Asia, it enhances fire activity in regions such as southern Mexico (Fig. 1.4a). Even within a given region, such as the one examined in western Africa (Fig. 1.7), cropland can have either an enhancing or suppressive effect on fire, depending on the time of year (Figs. 1.7, A.8–A.11).

These new estimates of burning associated with cropland, pasture, and other land could be used for a variety of purposes. For example, a lack of data has contributed to cropland and pasture management burning being mostly ignored in global fire models (although see Li et al., 2013; Pfeiffer et al., 2013); the results from this work could inform the development of mechanisms to account for such practices. Future development of this algorithm could add terms to explicitly account for interactions between land uses, such as cropland suppressing fire on non-agricultural land. This would generate estimates of burning on cropland separate from its effect on other land-use types, further improving the utility of the results.

Chapter 2

The FINAL global fire model, version 1: Incorporating modern-day cropland and pasture burning practices

2.1 Abstract

The Fire Including Natural & Agricultural Lands model (FINAL), which for the first time explicitly simulates cropland and pasture management fires based on how people use burning to manage their land, is described. The non-agricultural fire module uses empirical relationships in a quasi-mechanistic framework to estimate burned area. A novel automated optimization routine is used to fit parameters in the non-agricultural module, which improves fidelity of the model to observations of non-agricultural fire. This represents the first time a global fire model has been designed to replicate the patterns of non-agricultural fire unpolluted by cropland and pasture burning. The agricultural fire component is forced with previously-derived estimates of cropland and pasture fire frequency (Ch. 1). Unsurprisingly, then, FINAL accurately simulates the amount, distribution, and seasonal timing of burned cropland and pasture over 2001–2009 (global totals: 0.434×10^6 and 2.02×10^6 $km^2\,yr^{-1}$ modeled, 0.454×10^6 and 2.04×10^6 $km^2\,yr^{-1}$ observed), although associated carbon emissions for cropland and pasture fire are overestimated (global totals: 0.297 $PgC\,yr^{-1}$ and 0.712 $PgC\,yr^{-1}$ modeled, 0.194 $PgC\,yr^{-1}$ and 0.538 $PgC\,yr^{-1}$ observed). The non-agricultural fire module is less accurate: It markedly underestimates global burned area (1.66×10^6 $km^2\,yr^{-1}$ modeled, 2.44×10^6 $km^2\,yr^{-1}$ observed) and carbon emissions (1.33 $PgC\,yr^{-1}$ modeled, 1.84 $PgC\,yr^{-1}$ observed), with some regions seeing too much fire and others too little. Taken as a whole, FINAL represents an important step in the development of global fire models, but room for improvement remains.

2.2 Introduction

Vegetation fire is an important force for the Earth system at local, regional, and global scales. It can shape ecosystems (Bond and Kelley, 2005; Staver et al., 2011a), affect human health (Johnston et al., 2012; Marlier et al., 2012; Hahn et al., 2014), exacerbate or mitigate anthropogenic climate change (Ward et al., 2012; Ciais et al., 2013), and cause direct economic damage (Doerr and Santín, 2013; Bryant and Westerling, 2014). Fire occurrence can even affect the likelihood of more burning, through positive and negative feedbacks resulting from fire's impact on weather, climate, and vegetation (Laurance and Williamson, 2001; Balch et al., 2008; Zhang et al., 2008). Anthropogenic climate change and increases in atmospheric carbon dioxide concentrations have already increased – or can be expected to increase – the frequency and severity of burning in some parts of the world, while other regions could see decreased burning (Gillett et al., 2004; Westerling et al., 2006; Flannigan et al., 2009; Krause et al., 2014)

However, fire does not exist solely at the interface of climate and vegetation. Humans play an important role in regulating the fire regimes of many regions around the world (Flannigan et al., 2009; Bowman et al., 2011). This can come about as a result of many processes, one of which is fire's use as tool to manage agricultural lands. Croplands can be burned to facilitate planting or harvest; for example, sugarcane is typically burned before being harvested, and farmers in many parts of the world burn their crop wastes in the field after harvest (Yevich and Logan, 2003). Pastures and rangelands often see regular burning to reinvigorate the soil and control non-palatable weeds (Uhl and Buschbacher, 1985; Laris, 2002).

The way people burn croplands and pasture in a given region can differ from how the ecosystems there would burn in the absence of humans, in terms of both frequency and seasonal timing (Le Page et al., 2010; Magi et al., 2012; Rabin et al., 2015). This is significant for modeling efforts because it suggests a decoupling of agricultural fire from the mechanisms governing non-agricultural fire. For example, whereas the fire regime of southern Mali might naturally be dominated by large burns late in the dry season, humans have imposed a regime of small, scattered early burning to avoid such hard-to-control fires (Laris, 2002, 2011).

A full accounting of the importance of vegetation fire to the Earth system at present as well as historically and into the future requires the use of dynamic global vegetation models (DGVMs). These simulate processes of vegetation establishment, growth, mortality, disturbance, and competition at large scales using varying levels of mechanism, which allows the regional- and global-level biogeochemical implications of ecosystem dynamics to be fully estimated. When DGVMs are coupled with models of the soil, atmosphere, and oceans, the resulting Earth system models (ESMs) even simulate how these major components of our planet interact with and feed back upon one another. To understand the complex nature of fire's role in the Earth system, then, realistic models of vegetation burning must be designed and incorporated into DGVMs.

However, previous development of global fire models has mostly ignored the effects that agricultural management burning can have on real-world fire patterns. Anthropogenic effects on fire most commonly are modeled as dependent solely on population density, not land use (e.g., Venevsky et al., 2002; Arora and Boer, 2005; Pechony and Shindell, 2009; Thonicke et al., 2010; Li et al., 2012; Melton and Arora, 2015; Hantson et al., 2016). Moreover, the effect of population density is only to increase or decrease the amount of fire relative to that which would occur naturally – not to affect the intra-annual timing of fire. There are a few exceptions. The LPJ-LMfire model (Pfeiffer et al., 2013) includes functions to simulate how pre-industrial societies could manage cropland and pasture using fire, but these depend on assumptions that do not make sense in today's technological environment. A fire model developed for the Community Land Model by Li et al. (2013) simulates cropland fire, with annual burned area based on socioeconomic data (population density and gross domestic product) and timing based on observations, but pasture is not simulated as a land cover/use type distinct from grassland. The HESFIRE model (Le Page et al., 2015) accounts for how the amount of human land use (cropland and urban areas) affects burning, but again pasture is not considered. Neither of these latter two models, moreover, take into account how human activity can affect the *timing* of fire.

To some extent, the neglect of pasture burning in particular can be attributed to a lack of data. Cropland and a number of other vegetation types can, like fire, be algorithmically mapped using medium-resolution satellite imagery. Overlaying maps of vegetation type and burned area allows the generation of observational datasets of fire activity on different land covers (e.g., Giglio et al., 2010). However, no such map of global pasture distribution exists – only maps at relatively coarse resolutions describing the fraction of each gridcell that is pasture (e.g., Ramankutty et al., 2008; Klein Goldewijk et al., 2010). Developers of global fire models have thus, when aiming to design and parameterize models of non-agricultural burning, been limited in their choice of observational data with which to constrain their models. The options have been to either focus on regions with low fractions of cropland and/or pasture (thus potentially biasing their parameterization towards parts of the world inhospitable to agriculture) or to use a dataset "polluted" with signals from cropland and/or pasture burning. However, we now have estimates of the amount of fire associated with cropland, pasture, and non-agricultural lands at regional scales (Ch. 1; Rabin et al., 2015). This presents an opportunity to create a fire model that not only explicitly simulates burning practices on cropland and pasture, but also to develop a model of non-agricultural burning based on a purer observational signal.

However, the choice of reference data is only the first step in model development. Model fitting, also referred to as optimization or parameterization, is also critical, and many different methods can be used. Empirical fire models have often been fitted against observations of weather, climate, vegetation state, and anthropogenic factors using regression-type methods (e.g., Archibald et al., 2009; Lehsten et al., 2010) or multidimensional search algorithms (Knorr et al., 2014). Unfortunately, because they do not account for the effect of fire on vegetation, the models resulting from these methods can produce unexpected results. Imagine two DGVM-coupled fire models that are identical except for a few parameter values. These might be expected to simulate different amounts of burned area: The difference in those parameters results in different rules governing fire's behavior. But because fire is simulated to burn vegetation, they also produce different environments in which fire occurs: All other things being equal, an ecosystem that burns more frequently will have less biomass. Because biomass makes up the fuel for vegetation fires, this fire-biomass feedback represents an indirect effect of changing parameter values on simulated burned area. Of these two effects – different rules and different environments – previous models fit using regression or multidimensional search algorithms (e.g., Archibald et al., 2009; Lehsten et al., 2010; Knorr et al., 2014) only accounted for the former.

Quasi-mechanistic models that include processes intended to represent real physical mechanisms controlling fire occurrence and spread – such as the one presented here – are often run interactively with vegetation during development, thus accounting for both effects of changing parameters. But although this process is performed in combination with data from the literature when possible, it is rather manual and based on trial and error. Ideally, model fitting would combine the best parts of these two approaches, algorithmically searching parameter space for the "best" set of values based on how the model actually performs. Recently, Le Page et al. (2015) used the Metropolis Markov Chain Monte Carlo method to do just this in fitting the HESFIRE model.

Here we describe the development and performance of a global fire model that uses the new disentangled estimates of burned area associated with cropland and pasture (Ch. 1) to explicitly simulate land management burning practices on those lands using derived climatologies. The model also includes a model of non-agricultural fire that is fit against the purer, non-agricultural burning data – i.e., observations excluding fire on cropland and pasture – using an algorithm that explores parameter space interactively with the fire and vegetation model.

2.3 Fire model

The Fire Including Natural & Agricultural Lands (FINAL) model comprises two different submodels, simulating separately fire on agricultural and non-agricultural land. Here we describe the model's structure, beginning with the land and vegetation model within which it has been developed, then detailing the separate setups used for simulating non-agricultural and agricultural fire, and finally explaining the simulation of fire's effects on vegetation.

2.3.1 LM3

LM3, run by the National Oceanographic & Atmospheric Administration Geophysical Fluid Dynamics Laboratory (NOAA-GFDL), is a state-of-the art global dynamic vegetation and land surface model that can be run either offline or interactively with atmosphere and oceans in GFDL's Earth System Model (Shevliakova et al., 2009; Dunne et al., 2013). It simulates five different live plant biomass pools: leaves, heartwood, sapwood, labile carbon, and fine roots. (The heartwood, sapwood, and labile carbon pools can also be thought of as combining to form the "stem" biomass pool.) One of five different plant "species," representing biome types with different physiological properties, is assigned to each point based on bioclimatic envelopes and amount of biomass.

One of LM3's most interesting features is its use of sub-gridcell units called tiles, which allow land in different land use types (and in different stages of recovery from land use) to have distinct simulated vegetation and soil. Gridcells can have one each of "natural," cropland, and pasture tiles, along with several "secondary" tiles representing land in different stages of recovery from wood harvesting or agricultural abandonment. (There are also non-vegetated tiles representing glaciers and lakes.) Tiles are not spatially arranged, instead existing effectively as a list within each gridcell. Wood harvest and land use transitions occur once per year. At the same time, secondary tiles are merged together if they have similar amounts of heartwood biomass; this prevents the computational burden from becoming unreasonable.

LM3's structure lends itself nicely to fire modeling: Tiles could allow LM3 to simulate the heterogeneity of vegetation that fire can create across a landscape, and cropland and pasture tiles could have fire occur in a completely different way than non-agricultural tiles. However, the original LM3 fire model was rather simplistic: Cropland and pasture did not burn at all, and elsewhere, fire would happen once per year based on fuel loading, drought, and historical fire frequency (Shevliakova et al., 2009). The next two sections will describe the structure of the new fire models developed for non-agricultural (natural and secondary; Sect. 2.3.2) and agricultural (cropland and pasture; Sect. 2.3.3) tiles.

2.3.2 Fire model: Non-agricultural land

The fire model for non-agricultural lands is based on that developed for the Community Land Model by Li et al. (2012, 2013). Total burned area (BA) in the natural and secondary fire model is calculated as the product of the number of fires (N_{fire}) and burned area per fire (BA_{pf}):

$$BA = N_{fire} \times BA_{pf}. \tag{2.1}$$

Number of fires

Lightning and humans both serve as sources of ignitions, some fraction of which actually become fires. Li et al. (2012) modeled their equation for the density of lightning ignitions after that elaborated by Prentice and Mackerras (1977). At each time step, the number of ignitions from lightning (I_n, ignitions km^{-2}) is a function of latitude (Λ, radians) and the density of lightning flashes (L, flashes km^{-2}):

$$I_n = L \times (5.16 + 2.16cos\left[3\Lambda\right])^{-1}. \tag{2.2}$$

The number of anthropogenic ignitions (I_a, ignitions km^{-2}) is a function of population density (people km^{-2}):

$$I_a = (\beta_{Ia} \times P_D) \times \left(6.8 \times {P_D}^{-0.6}\right) \tag{2.3}$$

With β_{Ia} representing the rate of ignitions per person at each time step and P_D representing population density (people km^{-2}), the first part of Equation 2.3 gives a starting value for density of anthropogenic ignitions per time step.[1] The second part of Equation 2.3 is intended

[1] Throughout the rest of this chapter, β will denote parameters determined during our optimization routine (Sect. 2.3.6). The final values of these parameters can be found in Table 2.3.

to represent the fact that each person can be expected to light fewer fires as population density increases (Venevsky et al., 2002).

To calculate the number of ignitions actually becoming fires (N_{fire}), the total number of ignitions ($A_T[I_n + I_a]$, where A_T is the area of the tile in km^2) is multiplied by five functions that vary from zero to one, representing the suppressive effects of relative humidity (f_{RH}), soil moisture (f_θ), aboveground biomass (f_{AGB}), temperature (f_T), and population density (f_{P_D}):

$$N_{fire} = A_T(I_n + I_a) \times f_{RH} \times f_\theta \times f_{AGB} \times f_T \times f_{P_D}. \tag{2.4}$$

Li et al. (2012) calculate the effect of relative humidity on number of fires as

$$f_{RH} = max\left(0, min\left[1, \frac{0.7 - RH}{0.7 - 0.3}\right]\right), \tag{2.5}$$

where RH (range 0–1) is the relative humidity in the tile. Relative humidity ceases limiting fire (i.e., $f_{RH} = 1$) below $RH = 0.3$, and it suppresses all fire above $RH = 0.7$. However, the artificial limitation of this formulation to the range $[0, 1]$ would cause problems during our parameterization, which requires a continuously differentiable function. Instead we used the Gompertz function:

$$f_{RH} = \exp\left(-\beta_{RH,1} \times \exp[-\beta_{RH,2} \times RH]\right). \tag{2.6}$$

This function also varies between zero and one, with the parameter $\beta_{RH,1}$ controlling the location of the curve along the X axis and and $\beta_{RH,2}$ determining the steepness of the function as it decreases from one to zero.

Li et al. (2012) formulate the effect of soil moisture on number of fires as

$$f_\theta = \exp\left(\pi \times \left[\frac{\theta}{\theta_e}\right]^2\right), \tag{2.7}$$

where θ is volumetric soil moisture and θ_e is a parameter determining the soil moisture level where approximately 95% of fires are suppressed. This is a continuously differentiable function, but for consistency we translated it (like f_{RH}) into a Gompertz function:

$$f_\theta = \exp\left(-\beta_{\theta,1} \times \exp[-\beta_{\theta,2} \times \theta]\right). \tag{2.8}$$

In addition to flammability as determined by fuel moisture, Li et al. (2012) calculate the effect of above-ground biomass on number of fires as

$$f_{AGB} = max\left(0, min\left[1, \frac{AGB - AGB_{lo}}{AGB_{up} - AGB_{lo}}\right]\right), \tag{2.9}$$

where AGB ($\mathrm{kgC\,m^{-2}}$) is the sum of aboveground[2] biomass in the heartwood, sapwood, labile carbon, live leaf, and leaf litter pools. The parameters ($\mathrm{kgC\,m^{-2}}$) determine the levels of aboveground biomass below which fire is impossible (AGB_{lo}) and above which biomass is no longer limiting (AGB_{up}). However, as with f_{RH}, the fact that this function is not continuously differentiable would create problems for parameterization, so we used a Gompertz function instead:

$$f_{AGB} = exp\left(-\beta_{AGB,1} \times exp[-\beta_{AGB,2} \times AGB]\right). \tag{2.10}$$

[2]LM3 assumes that 80% of the total biomass carbon in the heartwood and sapwood pools is in the aboveground stem, with the rest being in coarse roots.

The effect of temperature on number of fires is calculated as

$$f_T = max\left(0, min\left[1, \frac{T - T_{lo}}{T_{up} - T_{lo}}\right]\right), \tag{2.11}$$

where T ($^\circ$ C) is the temperature of the canopy. The T_* parameters ($^\circ$ C) serve the same purpose as the parameters in the original formulation of f_{AGB} (Eq. 2.9); that is, no fire can occur ($f_T = 0$) at or below T_{lo} and temperature does not limit fire ($f_T = 1$) at or above T_{up}. After Li et al. (2013), we set T_{lo} to –10 $^\circ$ C and T_{up} to 0 $^\circ$ C. Because we did not include this function in the optimization, we did not convert it to a Gompertz function as we did with f_{RH} and f_{AGB}.

The suppressive effect associated with increasing population density on all potential fires (as opposed to just anthropogenic ignitions, as accounted for in Eq. 2.3) is calculated as

$$f_{P_D} = 1 - (0.99 - 0.98 \times exp\left[-\beta_{PD} \times P_D\right]), \tag{2.12}$$

where P_D is human population density (people km^{-2}). $f_{P_D} \rightarrow 0.01$ as $P_D \rightarrow \infty$, and $f_{P_D} = 0.99$ where $P_D = 0$, after Li et al. (2012). β_{PD} determines the shape of the function between these limits.

Li et al. (2013) also included a suppressive effect of per-capita gross domestic product (GDP) on number of fires. This was based on the idea that relatively wealthy parts of the world might have more valuable property to protect and a better capacity for suppression than less developed regions. However, for several reasons, we chose not to include this function. First, although globally gridded maps of GDP exist for the past 25 years or so (van Vuuren et al., 2007), no existing data sets describe the distribution of economic status before 1990. Second, the functions elaborated by Li et al. (2013) are somewhat ad-hoc, not taking into account other variables that might be responsible for the observed relationships. Bistinas et al. (2014), for example, showed that an apparent relationship between GDP density (GDP per area) and burned area (Aldersley et al., 2011) can be better explained as an emergent property resulting from the effect of population density, which is a major influence on GDP density. That result does not deal with GDP per capita, of course, but it does indicate the care that must be taken to avoid confounding variables when modeling fire. We thus declined to include GDP effects on burning in our model.

Burned area per fire

Burned area per fire is calculated based on an approximation of individual fires having elliptical shapes, with the point of ignition being one focus and the fastest spread occurring along the major axis (Fig. 2.1). It is made up of three main components: Duration, shape, and rate of spread.

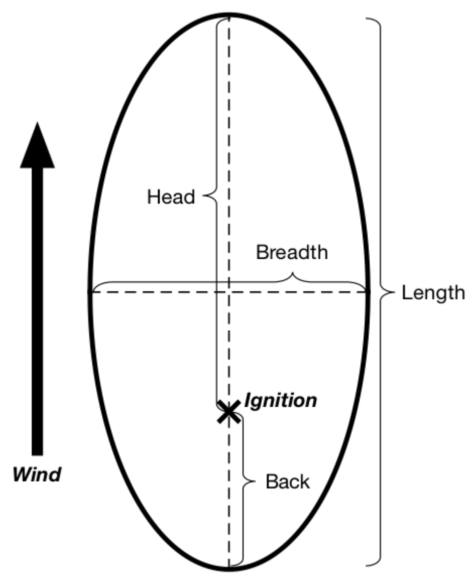

Figure 2.1: Approximation of fire as an ellipse.

Up to a certain point, fires become more elongated with increasing wind speed. That is, higher winds increase the length-to-breadth ratio LB (Fig. 2.1):

$$LB = 1 + 10 \times (1 - exp[-0.06W]), \tag{2.13}$$

where W is wind speed ($\mathrm{m\,s^{-1}}$) at 10 meters above ground level. High winds also increase rate of downwind spread relative to the rate of upwind spread, which can also be thought of as increasing the head-to-back ratio HB (Figure 2.1). HB is related to LB as

$$HB = \frac{LB + \sqrt{LB^2 - 1}}{LB - \sqrt{LB^2 - 1}}, \tag{2.14}$$

Forward rate of spread (ROS_f, $\mathrm{m\,s^{-1}}$) – i.e., spread rate downwind from an ignition – is a function of wind speed, fuel moisture, and vegetation type. Vegetation type ("species" *sensu* LM3) determines the maximum possible rate of spread in a tile. We initially defined maximum rate of spread for each species ($ROS_{max,sp}$) based on similar PFT-specific values used by Li et al. (2012 and Corrigendum): 0.4 $\mathrm{m\,s^{-1}}$ for C3 and C4 grass, 0.3 $\mathrm{m\,s^{-1}}$ for tropical and evergreen trees, and 0.22 $\mathrm{m\,s^{-1}}$ for temperate deciduous trees.[3] However, we included maximum rate of spread for tropical tree and C3 and C4 grass in the parameterization (β_{ROStt} and β_{ROSgr}, respectively; Sect. 2.3.6), so 0.4 $\mathrm{m\,s^{-1}}$ and 0.3 $\mathrm{m\,s^{-1}}$ represent their starting values. Their final values can be found in Table 2.3.

The rate of spread realized by any given fire increases with wind speed towards the limit of $ROS_{max,sp}$ according to the function $g(W)$:

$$gW = \frac{2LB}{1 + HB^{-1}} \times g_0, \tag{2.15}$$

where

$$g0 = \frac{1 + HB_{max}^{-1}}{2LB_{max}}, \tag{2.16}$$

[3]Note that although Li et al. (2012 and Corrigendum) actually used 0.22 $\mathrm{m\,s^{-1}}$ for all forest types other than needleleaf, we increased the maximum rate of spread in tropical tree tiles closer to that given by Li et al. (2012 and Corrigendum) for shrub PFTs (0.3 $\mathrm{m\,s^{-1}}$). This was done because the rate of spread in tropical savannas is much higher than that in tropical closed forests (especially moist forests), but LM3 has no "shrub" or "savanna" species, with the result that much of the world's tropical savannas are classified as "tropical tree."

Here, $LB_{max} = 11$ and $HB_{max} \approx 482$ are the limits of LB and HB as $W \to \infty$ (Equations 2.13 and 2.14).

Fires spread more slowly in wet conditions, so fuel moisture is considered in rate of spread. Li et al. (2012) multiplied rate of spread by f_{RH} (Equation 2.5) as well as $f_{RH}(\theta)$, the latter being identical to f_{RH} except with soil moisture (θ) replacing relative humidity (RH). However, we substituted $f_{RH}(\theta)$ with f_θ for simplicity and transparency. Thus, the complete equation for forward rate of spread is as follows:

$$ROS_f = ROS_{max,sp} \times g(W) \times f_{RH} \times f_\theta. \tag{2.17}$$

The final component of burned area per fire is the length of time between ignition and extinction. After Li et al. (2012), we set fire duration (d, seconds) to 24 hours (86,400 s).

$$BA_{pf} = \frac{\pi \times (ROS_f \times d)^2}{4 \times 10^6 \times LB} \times \left(1 + HB^{-1}\right)^2. \tag{2.18}$$

Li et al. (2013) also include functions that reduce burned area per fire based on population density and GDP per capita. We did not include either of these. The issues with using GDP per capita are described in Section 2.3.2 above. Population density might be considered a more trustworthy and meaningful statistic, but as with the GDP functions, the method used by Li et al. (2013) to describe the effect of population density on fire size was somewhat ad-hoc and did not take into account possible confounding factors. Moreover, our model optimization (Sect. 2.3.6) would have essentially seen the functions relating population density to number of fires and burned area per fire as one large, complicated function. For simplicity and parsimony, then, we did not include an effect of population density on burned area per fire.

Several limits are imposed on BA_{pf}. If the burned area calculated at a time step (i.e., $BA_{pf} \times N_{fire}$) is greater than the area of the tile that has not yet burned that day ($A_{t,un}$), BA_{pf} is adjusted for consistency:

$$BA_{pf} = \frac{A_{t,un}}{N_{fire}}. \tag{2.19}$$

Moreover, we add a limitation to fire size based on landscape fragmentation, based on the idea that fragmentation of the landscape into burnable and unburnable patches tends to prevent fires from reaching their maximum possible size (Archibald et al., 2009; Hantson et al., 2015). Maximum possible fire size as a function of tile size and fraction unburnable area in the gridcell is modeled after the function described by Pfeiffer et al. (2013):

$$BA_{pf,max} = A_t \times \left(1.003 + exp\left[16.607 - 41.503 \times \frac{A_{g,unburnable}}{A_g}\right]\right)^{-2.169}. \tag{2.20}$$

Here, A_g refers to the area of land (including nonvegetated “land” such as glaciers or lakes) in the gridcell, and $A_{g,unburnable}$ refers to the area of vegetated land in the gridcell other than cropland. $BA_{pf,max}$ is calculated at the end of each model day – after burning, tile splitting, and land-use transitions have occurred – and applied to the following day.

Burned area is calculated at every fast time step (30 model minutes) and accumulates throughout each day. At the end of each model day, burning occurs (Sect. 2.3.4).

2.3.3 Fire model: Cropland and pasture

Burned area on cropland and pasture tiles is estimated in a much more simplistic manner than that on natural and secondary tiles. At the beginning of each month, some fraction of each cropland and pasture tile burns according to a mean monthly climatology of burned fraction of cropland and pasture. These gridded climatology maps are based on results from the unpacking analysis (Ch. 1), which provided estimates of burned area associated with cropland, pasture, and other land on a month-by-month basis. The results presented in the main text of Chapter 1, with $\widehat{F_k}$ unconstrained, give the net effect of each land-use/cover type on burned area, including any suppressive effects cropland, for example, might have on burned area on non-agricultural land. However, here we use the results with $\widehat{F_k}$ constrained to non-negative values (App. A), which should provide a more reasonable estimate of how much burning actually occurs on each land cover type.[4]

2.3.4 Fire effects

Carbon in the leaves, stems, and aboveground litter of a burned tile is combusted (i.e., transferred to the smoke pool; Sect. 2.3.5) according to species-specific fractional combustion completeness (CC) values based on those used by Li et al. (2012). The remaining non-combusted biomass in leaves, stems, and fine roots is subjected to species- and pool-specific fractional mortality (M; i.e., transferred to above- or belowground litter), again based on values from Li et al. (2012). Combustion completeness and mortality values used here can be found in Table 2.1. Note that although the heartwood and sapwood pools are assumed to be 80% aboveground ("stems") and 20% belowground ("coarse roots"), CC_{stem} and M_{stem} are the same for both above- and belowground pools. This was necessary because LM3 assumes a constant 80%–20% split. However, fire-killed heartwood and sapwood is transferred to aboveground or belowground litter proportionally.

If less than 1 km^2 of a tile burns, the tile's biomass is reduced according to $CC \times BF$ and $M \times BF$, where BF is the burned fraction of the tile. This is the method that has been used by every other global fire model previously developed. However, it does not reflect the reality that an actual fire results in a mosaic where only part of the landscape has been burned. So when ≤ 1 km^2 burns in a given day, the model splits the tile into two new tiles – one burned and one unburned. This "fire tile splitting" occurs on all land cover types except cropland.

[4] For simplicity, the data from Chapter 1 may be referred to henceforth as the data from the "unpacking" analysis, or the "unpacked" data. The term "other land" may also be used to refer to non-agricultural land.

Table 2.1: Combustion completeness and mortality values for each "species" and tissue pool. Note that "stem" refers to both aboveground and belowground stem biomass, and that "root" refers only to fine roots.

	Combustion completeness				Mortality			
Species	Leaf	Stem	Root	Litter	Leaf	Stem	Root	Litter
C4 grass	0.85	1.00	0	0.85	0.85	0.00	0.20	0
C3 grass	0.85	1.00	0	0.85	0.85	0.00	0.20	0
Tropical tree	0.70	0.15	0	0.50	0.70	0.60	0.10	0
Temperate deciduous tree	0.70	0.10	0	0.45	0.70	0.55	0.07	0
Evergreen tree	0.75	0.20	0	0.55	0.75	0.65	0.13	0

2.3.5 Other changes

The implementation of daily fire and associated tile splitting necessitated many adjustments to parts of the LM3 codebase not dealing with fire directly. Previously, tiles would only be created and/or merged once per year, and secondary vegetation was the only land type allowed to have multiple tiles within a single gridcell. The code for land transitions needed to be reworked to allow daily splitting and merging. We also changed the code to allow all vegetation types, instead of just secondary land, to have multiple tiles (although we disabled this on cropland to reduce computational demand).The criteria for merging tiles were also altered to be based on aboveground biomass available for fire (AGB in Equation 2.9) instead of heartwood. Moreover, we changed the binning structure by which tiles are determined to have similar-enough biomasses to be merged. Previously, bin edges were located at 0.5, 1, 2, 3, 4, 5, 6, 8, 10, and 1000 $\mathrm{kgC\,m^{-2}}$. To better sample ranges of biomass where fuel is limiting, we replaced the first two bin edges with 0.1, 0.3, 0.5, 0.7, 0.9, and 1.1 $\mathrm{kgC\,m^{-2}}$. Finally, various aspects of carbon accounting throughout the model needed to be adjusted for daily tile splitting and merging.

More frequent fire also required other changes. Previously, grazing of pasture happened once per year, but in order to more reasonably simulate emissions from pasture fire we made grazing occur daily. For the main runs (`FINAL_V0` and `FINAL_V1`; Table 2.2), we also boosted the fraction of live leaf biomass removed by grazers from $\sim$0.07% $\mathrm{day^{-1}}$ to 4% $\mathrm{day^{-1}}$. This resulted in more realistic estimates of aboveground biomass in pasture, and of annual global consumption of biomass by grazers.

The "smoke pool," a virtual pool of carbon emitted by annual fire in the original LM3 implementation, was set to be emitted gradually over the course of the next year. The idea here was to avoid sudden unrealistic pulses of emissions, but now with daily fire it would be more unrealistic to pretend that it takes a year to emit the carbon from one day's burning. We thus adjusted the decay rate of the smoke pool for it to be drawn down over the course of a day rather than a year.

Finally, we used the CORPSE soil model (Sulman et al., 2014), which in addition to simulating the dynamics of soil organic matter also simulates leaf litter and coarse wood litter pools. This is important in some ecosystems where leaf litter is an important component of aboveground biomass and/or woody litter comprises a significant portion of fire emissions. However, the default setting for CORPSE is to simulate 15 different belowground soil cohorts, which poses a significant computational demand. To improve computational efficiency (especially important with the creation of so many tiles due to fire splitting on multiple land use types), we set CORPSE to simulate only one belowground soil cohort.

2.3.6 Parameter optimization

Simply copying parameters from the model described by Li et al. (2012, 2013) exactly was not possible for a number of reasons. First, here we have separated out both cropland and pasture from each other and from non-agricultural burning. Li et al. (2013), on the other hand, included special modules for cropland, deforestation, and peat fire – pasture burning being convolved with all other fire. Now that we have extracted from non-agricultural burning the influence of pasture, a significant source of fire activity that often differs from what might be expected under a totally "natural" fire regime, we expect to find different relationships between fire and its driving variables. Second, CLM is of course a different model than LM3, with its own idiosyncrasies and biases distinct from those of LM3. Although Li et al. (2012, 2013) strove to parameterize their equations based on independent data as much as possible, some functions were entangled with how their model itself worked. Third, as described in Section 2.3.5, we added some processes and removed others. Fourth, Li et al. (2012, 2013) tested their model against the GFED3 burned area dataset, whereas we used the GFED3s dataset, which has significantly more burned area than GFED3. Finally, Li et al. (2012, 2013) used different climatic forcing data than we did.

All these differences meant that we needed to reparameterize at least some parts of the non-agricultural fire model. Here we begin by briefly walking through the algorithm used to carry out the optimization, and then describe the parameters that we chose to optimize.

The Levenberg-Marquardt algorithm

We used the Levenberg-Marquardt method as the basis of our optimization routine. This algorithm uses the first derivatives of a performance metric with respect to each parameter to iteratively move through parameter space in search of a local minimum of the sum of squared errors. It starts with some initial guess, then evaluates the sum of squared errors S in non-agricultural burned area between the unpacked data and the estimates generated by the model:

$$S = \sum_{m=1}^{M} \sum_{i=1}^{N} \left(BA_{mod,i,m} - BA_{unp,i,m}\right)^2 . \tag{2.21}$$

Here, the summation is performed across all M months in the parameterization run period and all N sample gridcells selected for the optimization.

The algorithm then generates a new parameter set guess and the model is rerun. If the new guess decreases the sum of squared errors, it is "accepted," with a new guess then being generated based on it. If not, it is "rejected," and a new guess is generated based on the original guess. The details and derivation of how new guesses are calculated are described in Appendix B. To briefly summarize: Guesses are adjusted by interpolating between steps that would be generated by either the gradient descent method or the Gauss-Newton algorithm, leaning more towards the former when far from a minimum and the latter when near a minimum.

We initially selected 250 land cells at random from the LM3 grid, but rejected 9 for various reasons (all glacier, all lake, etc.). This left us with 241 gridcells which we would use for the optimization. Preliminary tests, however, revealed a few problems with the selection: A bias towards improving model fit in gridcells with strong model underestimation was evident (i.e, gridcells where the model simulated too much fire were undersampled), and the high northern latitudes – which make up a small fraction of global land area and an extremely small fraction of global fire activity – were judged to be oversampled. We got rid of 14 of those far northern gridcells (from Greenland and the Canadian tundra), then selected 23 new cells to bring us up to 250. The new cells were specifically selected from cells where a preliminary model run either underestimated or overestimated non-agricultural burned area relative to the unpacked data. Unfortunately, the model's performance in that preliminary run did not well match how the model actually performed in our optimization run. As such, we ended up oversampling areas of underestimation, leading to a bias towards making the model burn too much. We then culled the most extreme underestimated gridcells one by one until the sums of squared errors from underestimated and overestimated gridcells generated by the initial guess were approximately equal. This left us again with 241 gridcells, whose locations and initial sum of squared errors are shown in Figure 2.2a. A histogram of the mean annual error in burned area of the initial guess (Fig. 2.2b) shows that the positive and negative errors in this new dataset are approximately balanced.

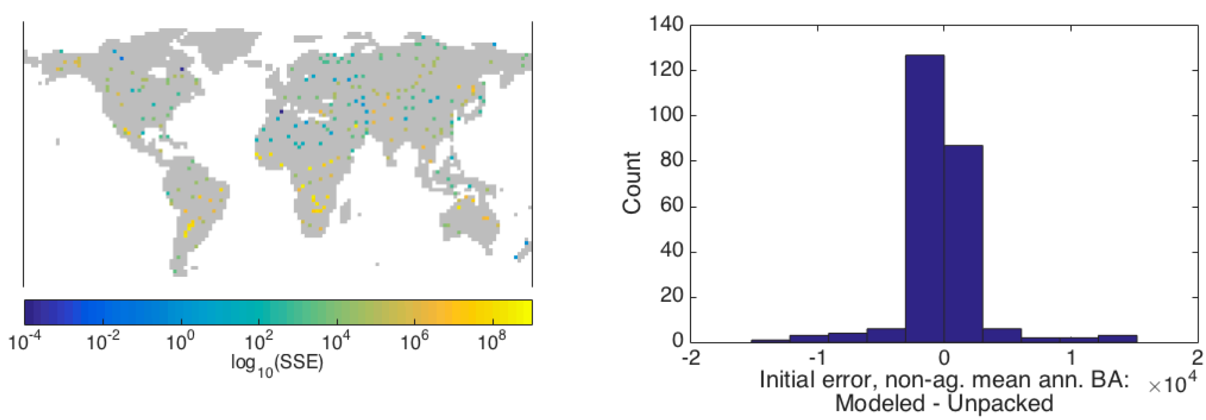

Figure 2.2: Summary of performance of initial guess in gridcells chosen for optimization with regard to non-agricultural burning. **(a)** Map of sum of squared errors. **(b)** Histogram of error in mean annual burned area.

The spinup run with which we generated initial conditions for the optimization is described later as `LM3_ORIG` (Sect. 2.4.1, Table 2.2). Note that we began the optimization runs in 1991 even though only the 2001–2009 data would be used for comparison to observations; the idea was to allow for the vegetation and fire regime in at least some of the gridcells (especially in regions where frequent fire is the norm) to equilibrate given the fire frequency of the new model.

Parameters chosen

From the equation for anthropogenic ignitions (I_a, Eq. 2.3), we optimized β_{Ia}, which can be thought of as controlling a sort of "baseline" value for how many ignitions each person can be expected to provide at each time step. Technically, we optimized $\beta_{Ia,m}$, which is describes the baseline number of ignitions per person per month instead of per timestep (of which there are 48 per day):

$$\beta_{Ia,m} = \beta_{Ia} \times 48 \times \frac{365}{12} \tag{2.22}$$

All other things being equal, higher values of $\beta_{Ia,m}$ result in more fires.

We also optimized β_{PD} from the function describing human suppression of all non-agricultural fires as a function of population density (f_{PD}, Eq. 2.12). All other things being equal, a higher value of this parameter would result in a faster approach of the fraction suppressed towards its upper limit.

Because the LM3 definition of a "species" to describe vegetation type is so broad, we thought it would be especially important to pay attention to several biome-specific maximum rate of spread parameters in FINAL. LM3's "tropical tree" type encompasses a wide range of real-world biomes, from tropical rainforests to semiarid shrublands. The rates of spread for fire in these systems are quite different, and so we included maximum rate of spread in tropical tree regions (β_{ROStt}) in the optimization. We also included the rate of spread in C3 and C4 grasslands (β_{ROSgr}), because preliminary testing showed strong overestimates in regions dominated by the C4 grass species especially.

Finally, we optimized parameters from f_{RH} ($\beta_{RH,1}$ and $\beta_{RH,2}$, Eq. 2.6), f_θ ($\beta_{\theta,1}$ and $\beta_{\theta,2}$, Eq. 2.8), and f_{AGB} ($\beta_{AGB,1}$ and $\beta_{AGB,2}$, Eq. 2.10). We generated initial guesses for these parameters by fitting Gompertz functions, with the upper asymptote set at 1, to the corresponding functions from Li et al. (2012). Fitting was performed using the MATLAB Curve Fitting Toolbox (MATLAB and Curve Fitting Toolbox Release 2014b, The MathWorks, Inc., Natick, Massachusetts, United States.)

2.4 Experimental setup and analysis

2.4.1 Experimental runs

Spinup of the land to pre-industrial conditions began with a "bare ground" scenario and ran for 300 years, during which climate forcings (Sect. 2.4.2) from 1948–1977 were repeatedly cycled through. During spinup, atmospheric CO_2 concentration was held constant at 286 ppm and land use was turned off. Next, we simulated years 1861–1947, using repeated 1948–1977 climate forcings but historical land use and atmospheric CO_2 concentration (Sect. 2.4.2). Finally, the model was run from 1948–1991 with historical climate forcings, land use, and atmospheric CO_2. This run – referred to as `LM3_ORIG` (Table 2.2) – provided initial conditions for other model runs, including the optimization. Note that the daily grazing intensity (Sect. 2.3.5) was set at its default value of ~0.07 for `LM3_ORIG`.

The new model (Sects. 2.3.2–2.3.5), with new parameters as described in Section 2.5.1 and Table 2.3, was run from 1948–2009 (`FINAL_V1`; Table 2.2). This run began with initial conditions as produced for the beginning of 1948 by the original LM3 run described above (`LM3_ORIG`). An experimental run with the complete new model structure but all settings as initially guessed in the parameterization (`FINAL_V0`) was also performed, comparison of which to `FINAL_V1` would allow us to explore where the optimization improved or worsened model performance. For both `FINAL_V0` and `FINAL_V1`, daily grazing intensity (Sect. 2.3.5) was set at 4%.

Finally, we tested the effect of "fire tiling" on equilibrium pre-industrial potential biomass by performing two additional pre-industrial spinups. As in `LM3_ORIG`, both used a constant atmospheric CO_2 concentration of 286 ppm with repeated 1948–1977 climate and no land use for 300 years. These two runs, however, used the non-agricultural fire model as described in Section 2.3.2 except with (a) all human effects turned off, and (b) all functions as originally formulated by Li et al. (2012, 2013). Both runs did use the changes made to other parts of LM3 described in Sect. 2.3.5. The only difference between the two runs was that one (`FINAL_PI_ASLI`) had fire tiling turned on for all cells that burned at least 1 km^2, and the other (`FINAL_PI_ASLI_NFT`) had fire tiling turned completely off (Table 2.2). We hypothesized that fire tiling would increase biomass density by creating "refugia" of unburned tiles where conditions become increasingly less flammable.

2.4.2 Input data

The LM3 land and vegetation model is run "offline" in this study, meaning that it is forced by a set of meteorological and radiation-related variables without any interaction between the land and atmosphere. The variables used here to force LM3 – daily precipitation, surface air pressure, specific humidity, wind vectors, and downward longwave and shortwave radiation – are taken from the dataset developed by Sheffield et al. (2006). All variables are interpolated to the spatial and temporal resolution of the LM3 fast time step, here set to 30 model minutes. Carbon dioxide (CO_2) concentrations are taken from Meinshausen et al. (2011). Historical data on land use transitions and wood harvesting come from the harmonized dataset created by Hurtt et al. (2011) for use in Earth system models. The mean distributions of cropland, pasture, and non-agricultural land in this study over 2001–2009 are presented in Figure 2.3.

In addition to these standard LM3 forcing data, the FINAL non-agricultural fire model also requires data on lightning. We used a gridded monthly climatology of lightning flash rate (flashes km^{-2}) based on data from the Lightning Imaging Sensor (LIS) and Optical Transient Detector (OTD) remote instruments. Specifically, we used the LIS/OTD Low-Resolution Monthly Time Series (LRMTS) described by Cecil et al. (2012). This dataset is provided at a $2.5° \times 2.5°$

Table 2.2: Experimental runs discussed in this chapter. "1–300" indicates that 300 years were simulated, but that these are not tied to any historical data, and thus they do not correspond to any actual historical years. "Graze rate" refers to the amount of non-wasted leaf biomass consumed each day by livestock on pasture tiles.

Name	Fire model	Years	Initial conditions	Climate	CO_2	Land use	Graze rate	Non-agri. fire: Humans	Agri. fire
`LM3_ORIG`	Original	"1–300"	Cold start	Repeated 1948–1977	286 ppm	Off	n/a	n/a	n/a
		1861–1947	—	Repeated 1948–1977	Historical	Historical	0.07%	n/a	n/a
		1948–1991	—	Historical	Historical	Historical	0.07%	n/a	n/a
`FINAL_PI_ASLI`	FINAL, all parameters and functions after Li et al. (2012, 2013)	"1–300"	Cold start	Repeated 1948–1977	286 ppm	Off	n/a	Off	n/a
`FINAL_PI_ASLI_NFT`	New w/ fire tiling disabled; all parameters and functions after Li et al. (2012, 2013)	"1–300"	Cold start	Repeated 1948–1977	286 ppm	Off	n/a	Off	n/a
`FINAL_V0`	New fire model structure; initial guess parameter set	1948–2009	As `LM3_ORIG`	Historical	Historical	Historical	4%	On	As unpacked
`FINAL_V1`	New fire model structure; optimized parameter set	1948–2009	As `LM3_ORIG`	Historical	Historical	Historical	4%	On	As unpacked

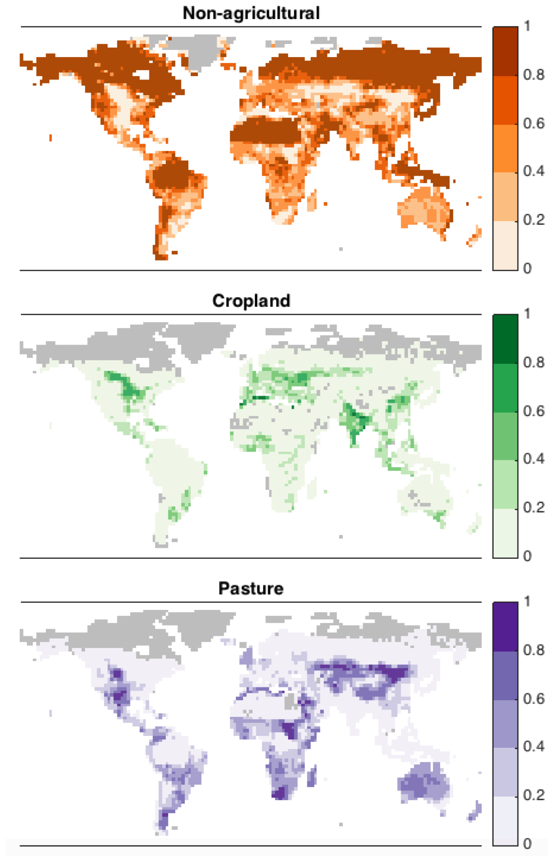

Figure 2.3: Mean fractional land cover of **(a)** non-agricultural land, **(b)**, cropland, and **(c)** pasture over 2001–2009 as simulated in model runs (after Hurtt et al., 2011). Gray cells did not contain any of the indicated land cover type.

resolution, which we interpolated to match the LM3 resolution of 2° latitude by 2.5° longitude. The version of LRMTS that we used, v2.3, included maps of flash rate for each month in the period 1996–2014. We found the average of each month (January, February, etc.) and used these to build our climatology.

Lastly, FINAL requires input data on population density. We used the historical population density estimates from HYDE 3.1 (Klein Goldewijk et al., 2010), coarsened from their original 5-minute resolution to the LM3 resolution (2° latitude by 2.5° longitude). We interpolated population density linearly between each time point in the HYDE dataset.

The agricultural fire model in FINAL requires climatologies of burned area associated with cropland and pasture, and the optimization routine required estimates of non-agricultural fire. For these, we used the global gridded maps of monthly burned area and emissions associated with cropland, pasture, and other (non-agricultural) land based on an algorithm that generated estimates for each of 134 regions around the world based on the GFED3s data (Randerson et al., 2012) as described in Chapter 1. Burned fraction for each gridcell in the unpacked data is adjusted here to produce the correct amount of burned area, accounting for the fact that the land cover distributions used in the unpacking (Ch. 1) differ slightly from those used in this study.

2.4.3 Analysis of results

The new model's performance in terms of recreating observed patterns of burned area and fire carbon emissions is evaluated here by comparison against GFED3s and the unpacked fire data. In addition to global totals of mean annual fire activity, we assess the spatial distribution of fire using maps of mean annual burned fraction and emissions.

The accuracy of seasonal fire trends is tested by comparing the difference between peak day of burned area simulated by the model with the peak as estimated by the unpacking analysis. This is quantified using mean phase difference, as described by Kelley et al. (2013). Each gridcell's annual pattern of fire can be described as a vector in the complex plane:

$$\boldsymbol{V_i} = (x_{m,i}, \theta_m)\,, \tag{2.23}$$

where $x_{m,i}$ is the mean burned area in month m for gridcell i, and θ_m is an arbitrary angle unique to month m and calculated for all gridcells as:

$$\theta_m = 2\pi \frac{(m-1)}{12}. \tag{2.24}$$

The mean vector $\boldsymbol{L_i}$ for each gridcell has end points that can be described in Cartesian coordinates as the origin and $(L_{x,i}, L_{y,i})$, where:

$$L_{x,i} = \sum_{m=1}^{12} x_{m,i} \cos\left(\theta_m\right) \tag{2.25}$$

and

$$L_{y,i} = \sum_{m=1}^{12} x_{m,i} \sin(\theta_m). \quad (2.26)$$

The phase P_i, defined where fire occurrence is not distributed evenly across all months, describes the mean timing of peak fire activity:

$$P_i = \arctan\left(\frac{L_{y,i}}{L_{x,i}}\right). \quad (2.27)$$

The day of the year associated with peak fire activity can be calculated as $\frac{P_i}{2\pi} \times 365$. Mean phase difference MPD, which is used here to describe the difference in timing of peak fire between model results and observations, is calculated as

$$MPD = \frac{1}{\pi} \arccos\left(\frac{\sum_{i=1}^{N} cos\left[P_{i,mod} - P_{i,obs}\right]}{N}\right), \quad (2.28)$$

where modeled and observed phases are designated with the subscripts *mod* and *obs*, respectively. MPD varies from zero to one, with $MPD = 0$ if all modeled peaks correspond exactly to observed peaks and $MPD = 1$ if all modeled peaks differ from observed peaks by the maximum possible amount (6 months).

2.5 Results

2.5.1 Optimized parameters

Figure 2.4 shows the progression of the parameter guesses, along with the sum of squared errors associated with each parameter set guess through the optimization. The sum of squared errors decreases rapidly for the first few iterations, but diminishing returns become apparent by about the fifth iteration (Fig. 2.4a). By the eleventh iteration, it did not seem that allowing iterations to continue would result in much improved sums of squared errors, and the optimization was manually halted. The original and final parameter values can be found in Table 2.3.

Table 2.3: Values of each optimized parameter, before (Initial) and after (Final) optimization.

	Initial	Final
$\beta_{RH,1}$	0.0062	0.011856898
$\beta_{RH,2}$	-9.1912	-0.172544308
$\beta_{\theta,1}$	0.0750	0.329099402
$\beta_{\theta,2}$	-6.3741	-6.967427375
$\beta_{AGB,1}$	7.3157	44.20896443
$\beta_{AGB,2}$	4.11	9.820100287
$\beta_{Ia,m}$	0.0035	0.002224368
β_{PD}	0.025	0.030732082
β_{ROSgr}	0.4	0.268421539
β_{ROStt}	0.3	1.018599996

The functions resulting from the new parameter set are visualized, in comparison with how they were in the Li et al. (2012, 2013) model as well as in the initial optimization guess, in Figure 2.5.

f_{AGB} saw its parameters increase markedly: both $\beta_{AGB,1}$, which translates the function along the X axis, and $\beta_{AGB,2}$, which controls the slope of the increase of f_{AGB} from low to high biomasses (Fig. 2.4b, c). The net effect relative to the original guesses was that the amount of fire allowed decreased at biomasses below about 0.3 $\mathrm{kg\,C\,m^{-2}}$ and increased between about 0.3 to 1.5 $\mathrm{kg\,C\,m^{-2}}$ (Fig. 2.5g).

The parameter controlling anthropogenic ignitions, $\beta_{Ia,m}$, decreased through the sixth guess, then increased to a level higher than initially guessed, before declining again to a low level by the end of the optimization (Fig. 2.4d). The density of anthropogenic ignitions I_a is thus decreased at all positive levels of population density (Fig. 2.5a). Moreover, the parameter β_{PD} – which controls anthropogenic suppression of burning f_{P_D} – increased (Fig. 2.4e), meaning that a larger fraction of ignitions (both lightning and anthropogenic) are suppressed wherever population density is greater than zero, though most noticeably between densities of 10–100 people $\mathrm{km^{-2}}$ (Fig. 2.5b). The net effect is to reduce unsuppressed anthropogenic ignitions (i.e., $I_a \times f_{P_D}$) relative to the initial guess, with the peak's location being mostly unchanged but its severity being modulated (Fig. 2.5e).

Four parameters relating to the effect of moisture on fire activity were optimized: $\beta_{RH,1}$ and $\beta_{RH,2}$, which control the effect of relative humidity f_{RH}, and $\beta_{\theta,1}$ and $\beta_{\theta,2}$, which control the effect of soil moisture f_{θ}. Altogether – i.e., taking into account moisture effects on both ignition success probability and rate of spread – fire is proportional to $(f_{RH} \times f_{\theta})^3$. Because f_{RH} and f_{θ} always appear together in the model equations, and because relative humidity and soil moisture might be expected to be strongly correlated, one might have expected the optimization to result in similar functions. However, the final shapes of f_{RH} and f_{θ} are quite different (Fig. 2.5c, d). $\beta_{\theta,1}$ increased and $\beta_{\theta,2}$ decreased (Fig. 2.4h, i), resulting in a stronger suppressive effect of soil moisture: Whereas the original function suppresses nearly all fire beginning at around $\theta = 0.65$, the new function reaches this point around $\theta = 0.35$ (Fig. 2.5d). Even in extremely dry soils where $\theta = 0$, $f_{\theta} = 0.7$ – meaning that around 30% of ignitions would be prevented from becoming spreading fires, and rate of spread would be reduced by 51%. f_{RH}, on the other hand, was effectively neutered: While $\beta_{RH,1}$ and $\beta_{RH,2}$ both increased (2.4), $\beta_{RH,2}$ increased so drastically that $f_{RH} \approx 1$ for all values of relative humidity (Fig. 2.5c). Figure 2.5f shows that the total effects of these shifts in these moisture functions are most extreme at low values of soil moisture, with low levels of relative humidity burning less and high levels of relative humidity burning more (all other things being equal). However, LM3 never produced the latter condition (Fig. 2.6d), and so low-humidity cells seem to have driven this trend.

Maximum rate of spread decreased more than 25% for grassland (Fig. 2.4k), a result which likely has to do with the model overestimating fire in these low-biomass systems. This parameter decreased sharply for most of the optimization, but as f_{AGB} appropriately began to take on more of the responsibility for regulating fire there, grassland maximum rate of spread began to increase back towards its initial guess. Maximum spread rate increased by over 300% for the "tropical tree" vegetation type (the initial guess for which was higher than that used by Li et al. (2012); Fig. 2.4j), which was due to a tendency towards underestimation of burned area in that biome.

Comparing the results of `FINAL_V0` with `FINAL_V1`, we can see that much of the improvement came in regions where the initial parameter set severely overestimated burned area (Fig. 2.7a–

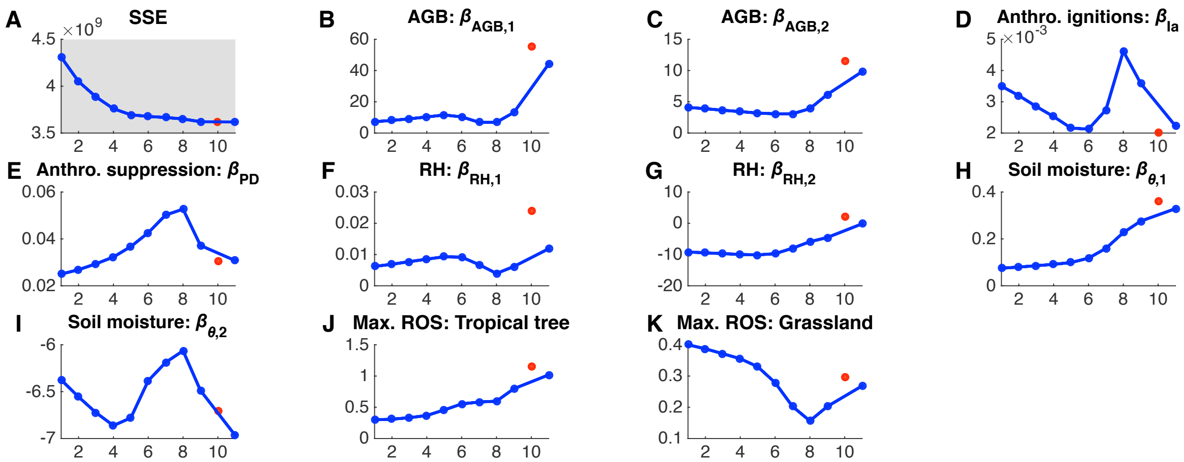

Figure 2.4: Trace plots showing the progression of sum of squared errors **(a)** and each of the ten parameters **(b–k)** over the length of the optimization. X-axes show iteration number, Y-axes show sum of squared errors or parameter guess value, and color of points indicate whether the associated parameter set guess was accepted (blue) or rejected (red).

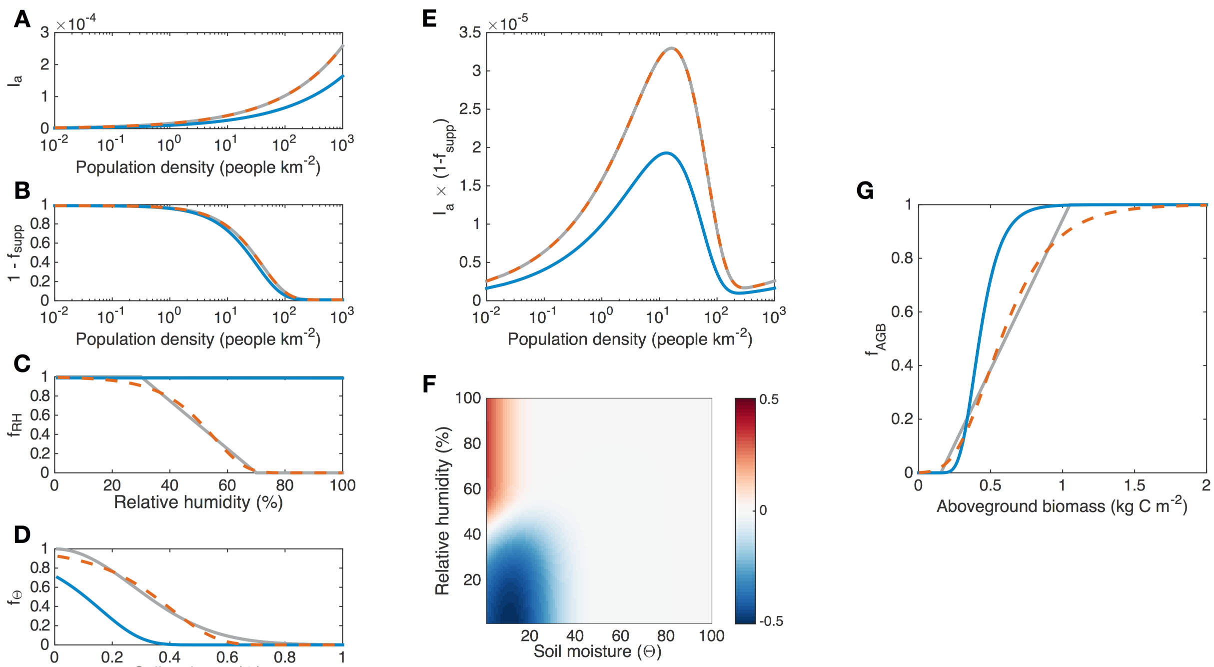

Figure 2.5: Changes in functions that got optimized, from original Li et al. (2012, 2013) functions (solid gray) to initial guesses with Gompertz-style functions where necessary (dashed red) to final parameter set (solid blue).

d). Performance worsened in other gridcells. A map of root mean squared error (Fig. 2.7e), which shows performance improvement as would be "seen" by the optimization algorithm for included gridcells, highlights a few cells in and around the tropical rainforests of Africa and

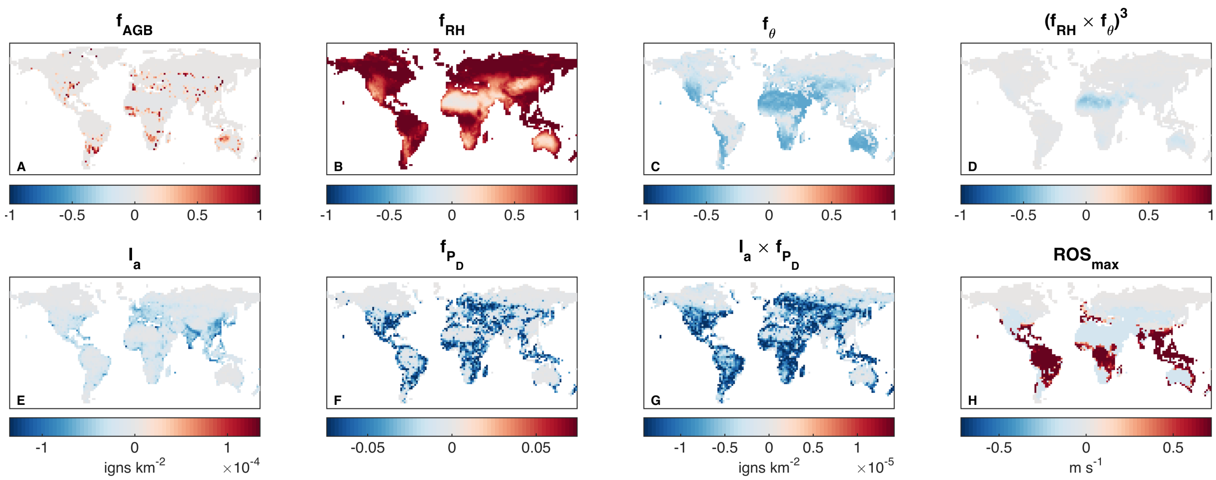

Figure 2.6: Difference in mean value of various fire model functions over 2001–2009 between FINAL_V0 and FINAL_V1. Red indicates regions where the function in FINAL_V1 allows more fire than in FINAL_V0; blue, less.

South America as areas where the performance metric increased markedly (indicating worsened performance) between the initial and final guesses.

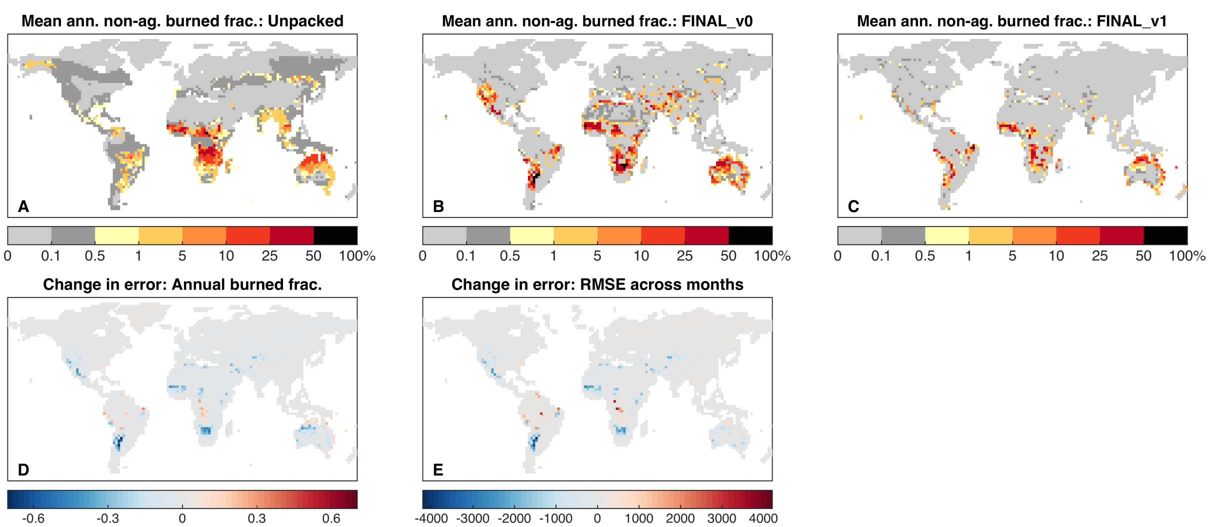

Figure 2.7: Improvement in non-agricultural fire model performance between the initial guess (run FINAL_V0) and the final parameter set (run FINAL_V1). **(a–c)** Mean annual burned fraction on non-agricultural lands from unpacking **(a)**, the initial guess **(b)**, and the final parameter set (**c**; identical to Fig. 2.8i.) **(d–e)** Difference between runs FINAL_V0 and FINAL_V1 in correspondence of modeled to unpacked non-agricultural burning as measured by mean annual burned fraction **(d)** and root mean squared error **(e)**. For **(d)** and **(e)**, blue indicates improvement by FINAL_V1 over FINAL_V0.

2.5.2 Model performance

Figure 2.8 compares maps of mean annual burned fraction (i.e., fraction of land area) from run `FINAL_V1` with those from GFED3s (Randerson et al., 2012) and the unpacking analysis. Figure 2.9a shows the difference in mean annual burned fraction between the model and the unpacked observations, against which the non-agricultural model was parameterized. Considering all land cover types together, the new fire model recreated the general pattern of annual fire activity well compared with both GFED3s (Randerson et al., 2012) and the unpacked data (Figs. 2.8a,b,f; 2.9a). The largest modeled overestimates relative to the unpacked data occurred in the grasslands and shrublands of western South America, the western Caatinga of northeast Brazil, and at various points throughout the African savannas (Fig. 2.9a). Most of the severe model underestimation relative to the unpacked data occurred in the African tropical savannas, as well as (to a lesser extent) the tropical savannas of northern Australia (Fig. 2.9a).

The modeled burned fractions of cropland and pasture match the unpacked numbers almost exactly (Figs. 2.9c,d), which is not surprising considering that the unpacked data were used to force the model on cropland and pasture tiles. There are some notable discrepancies, however. Specifically, there is too much cropland fire in one European gridcell and too little in several gridcells in northern Australia (Fig. 2.9c). Pasture fire did not experience such severe error in burned fraction anywhere (Fig. 2.9d).

The strong correspondence of modeled cropland and pasture fire with the unpacked observations (as expected since the latter were directly used to drive the former) suggests that the majority of the error seen in total burning must be associated with fire on non-agricultural lands. Indeed, although the non-agricultural fire model generally captured the worldwide distribution of fire – with tropical savannas, grasslands, and shrublands generally dominating burned area – the fit is by no means perfect (Fig. 2.9b). There are a number of regions where the model simulates little to no non-agricultural burning but the unpacked data show significant amounts of fire (Figs. 2.8b,f) . This phenomenon is especially noticeable in the eastern African savannas, the shrublands

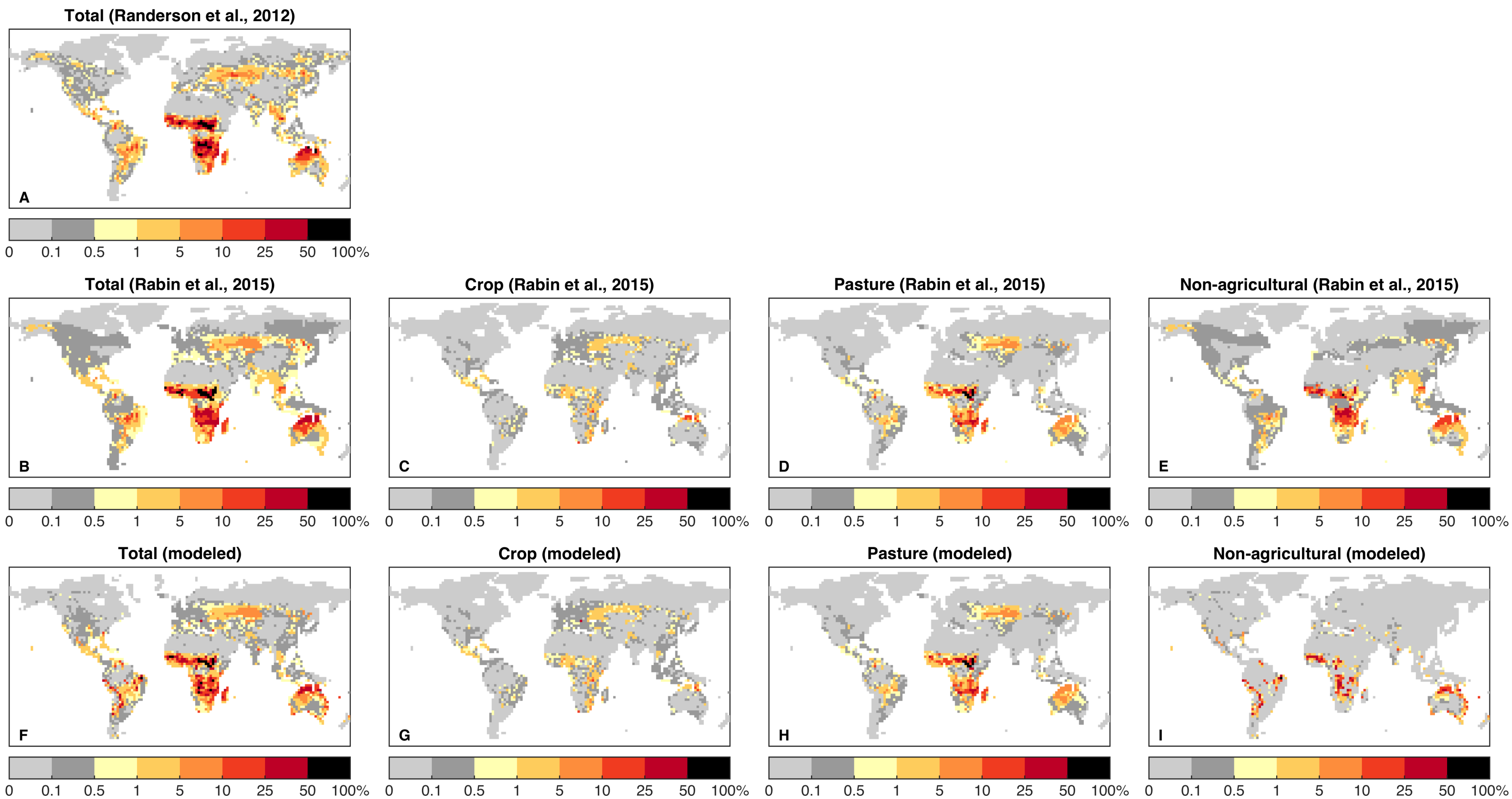

Figure 2.8: Mean annual burned fraction. **(a)**: From GFED3s (Randerson et al., 2012); **(b–e)**: observational estimates from unpacking analysis; **(f–i)**: Model-estimated. ([i] is identical to Fig. 2.7c.)

Table 2.4: Global mean annual burned area and associated carbon emissions, 2001–2009. (1) Midpoint of values for cropland burning with (0.208) and without (0.456) including cropland-natural mosaic.

	Burned area ($10^6\,km^2\,yr^{-1}$)				C emissions ($PgC\,yr^{-1}$)			
	T	C	P	O	T	C	P	O
GFED3s	4.68	0.332$^{(1)}$	—	—	2.48	n.d.	—	—
Unpacked	4.93	0.454	2.04	2.44	2.57	0.194	0.538	1.84
`FINAL_V0`	6.38	0.434	2.02	3.93	2.21	0.295	0.703	1.21
`FINAL_V1`	4.11	0.434	2.02	1.66	2.34	0.297	0.712	1.33

of western Australia, and throughout the tropical and temperate grasslands, savannas, and shrublands of South America.

Worldwide, the non-agricultural fire model underestimated burned area, with $1.66\times10^6\ km^2\,yr^{-1}$ simulated as having burned – an underestimate of 32% relative to the unpacked estimate. Unsurprisingly given the spatial results presented above, global averages for cropland and pasture were much better – $0.434 \times 10^6\ km^2\,yr^{-1}$ (4% underestimate) and $2.02 \times 10^6\ km^2\,yr^{-1}$ (1% underestimate), respectively. Mean annual global burned area across all land covers over 2001–2009 was modeled as $4.11 \times 10^6\ km^2\,yr^{-1}$, an underestimate of 12% relative to GFED3s and an underestimate of 17% relative to the unpacked total.[5] The time series of annual burned area over 2001–2009 for each land cover from the model (i.e., `FINAL_V1`) are compared with the GFED3s and unpacked estimates in Figure 2.10a.

Just as the model tended to underestimate total global burned area, it also underestimated carbon emissions from fire (Table 2.4). The 2.34 $PgC\,yr^{-1}$ simulated by the model represents an underestimate of 6% relative to GFED3s and of 9% relative to the unpacking data. This is again principally due to non-agricultural fire, for which the model simulated 1.33 $PgC\,yr^{-1}$ as opposed to the unpacked estimate of 1.84 $PgC\,yr^{-1}$ – an underestimate of 28%. Agricultural fire emissions were actually overestimated, with 0.297 $PgC\,yr^{-1}$ for cropland and 0.712 $PgC\,yr^{-1}$ for pasture – overestimates of 53% and 32% compared to the unpacked values of 0.194 $PgC\,yr^{-1}$ and 0.538 $PgC\,yr^{-1}$, respectively.

The spatial distribution of errors in total fire carbon emissions (Fig. 2.9e) generally reflects the distribution of errors in simulated burned area (Fig. 2.9a). As with burned area, there

[5]The unpacked estimate is greater than the value from GFED3s because of the constraint of $\widehat{F_k}$ values to ≥ 0 (Ch. 1).

are sizable regions where the model simulates little to no non-agricultural fire carbon emissions but the unpacked data show otherwise (Figs. 2.11e,i). Cropland fire emissions, as with burned area, are underestimated in northern Australia; there are also two regions in central Africa where cropland fire emissions are overestimated despite essentially correct annual burned fraction (Figs. 2.11c,g). The areas of slightly underestimated pasture burned fraction are not apparent in the map of pasture fire emissions error; large overestimates of emissions from pastures in the tropical savanna biome are instead the most apparent aberrations (Figs. 2.11d,h).

The non-agricultural fire model performed well in terms of simulating the within-year timing of burned area (Figs. 2.12e,i). This was reflected in the results for combined burning across all land cover types, which corresponded well with both GFED3s and unpacked burned area (Figs. 2.12a–b,f); the timing of peak model-estimated fire was 35 days later than observed for all fire combined as compared with total unpacked fire ($MPD = 0.19$), and 53 days later than observed for non-agricultural fire specifically ($MPD = 0.29$).

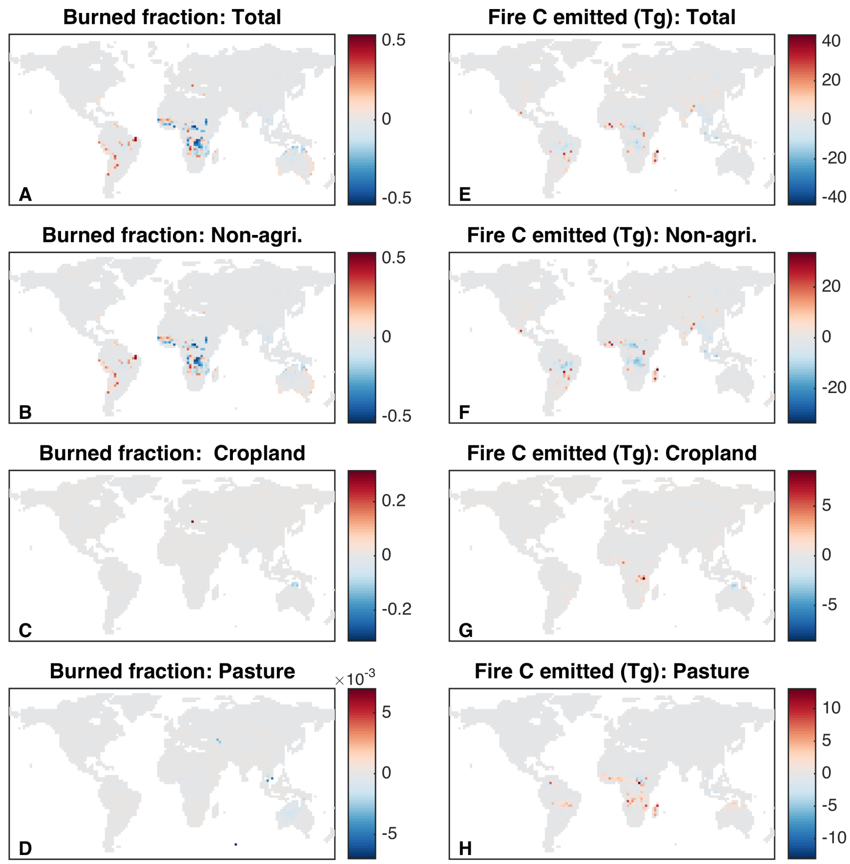

Figure 2.9: Absolute error in mean annual burned fraction **(a–d)** and fire carbon emissions **(e–h)** for each land cover type: Model-estimated minus observational estimates from unpacking analysis.

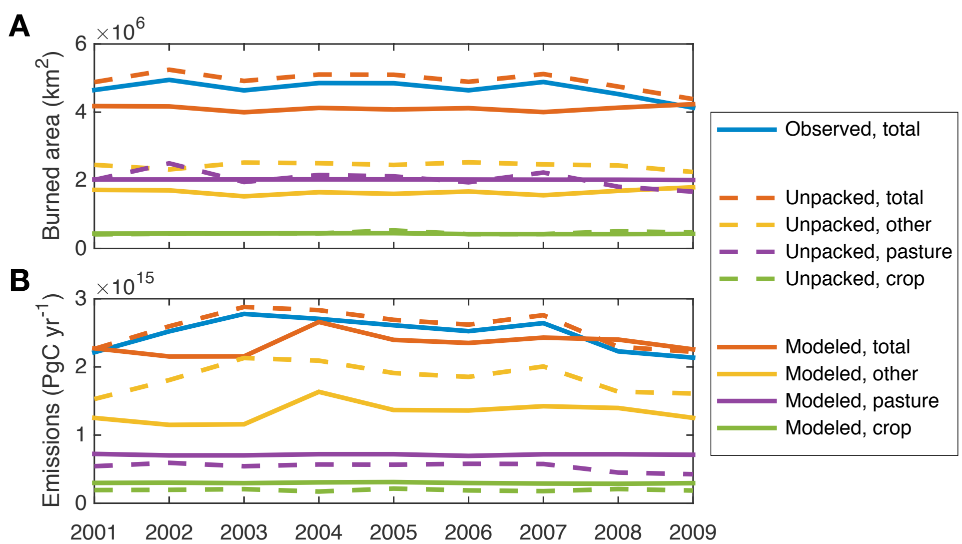

Figure 2.10: Annual time series of observed and model-estimated burned area (**a**, km^2) and fire carbon emissions (**b**, $\mathrm{PgC\,yr}^{-1}$) from 2001–2009. Dashed lines: Observational estimates of total and by-landcover fire emissions from Rabin et al. (2015). Solid blue lines: Observations of total emissions from GFED3s (Randerson et al., 2012). Other solid lines: Model-estimated total and by-landcover fire emissions.

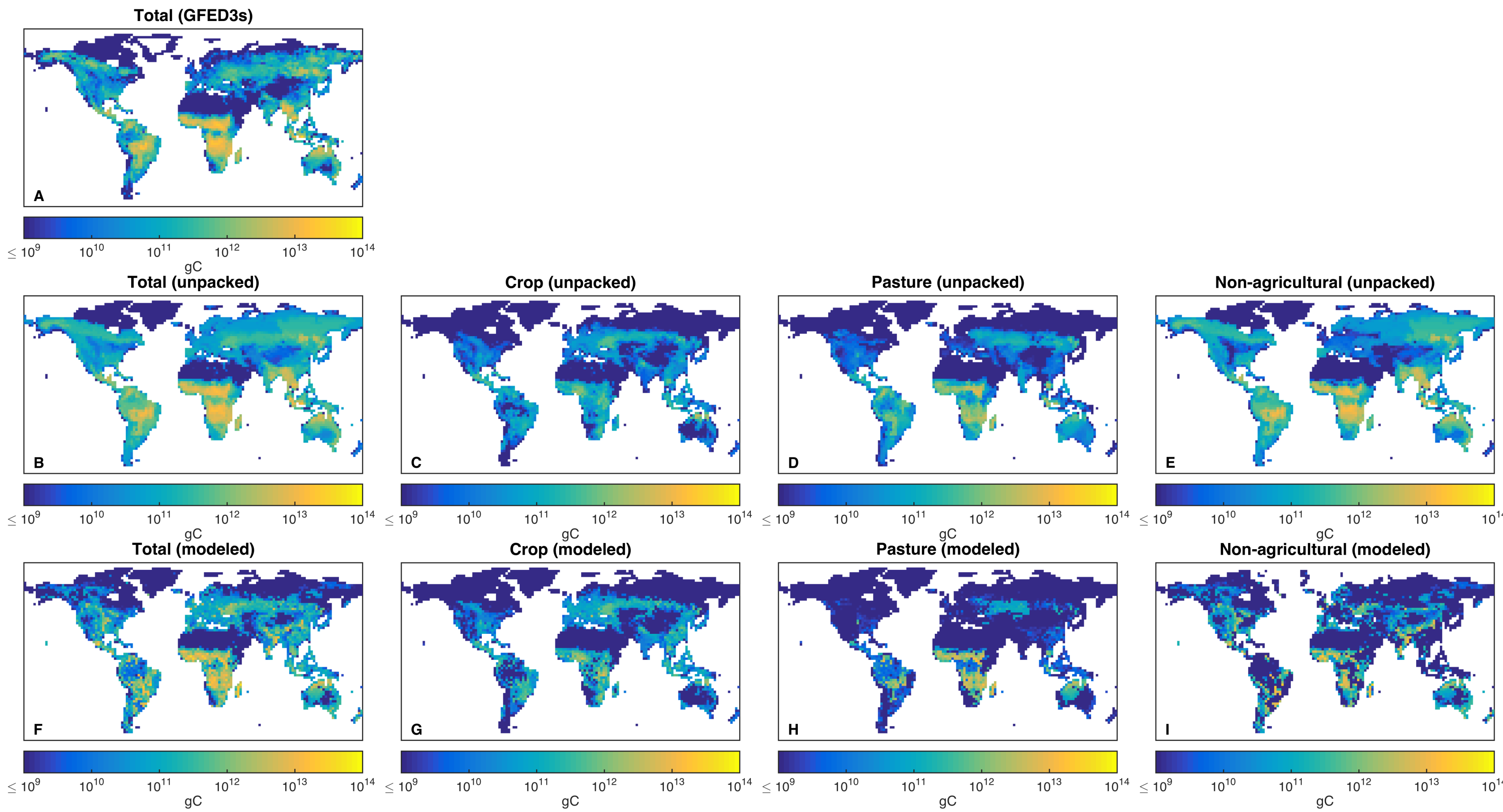

Figure 2.11: Mean annual fire carbon emissions. **(a)**: From GFED3s (Randerson et al., 2012); **(b–e)**: observational estimates from unpacking analysis; **(f–i)**: Model-estimated.)

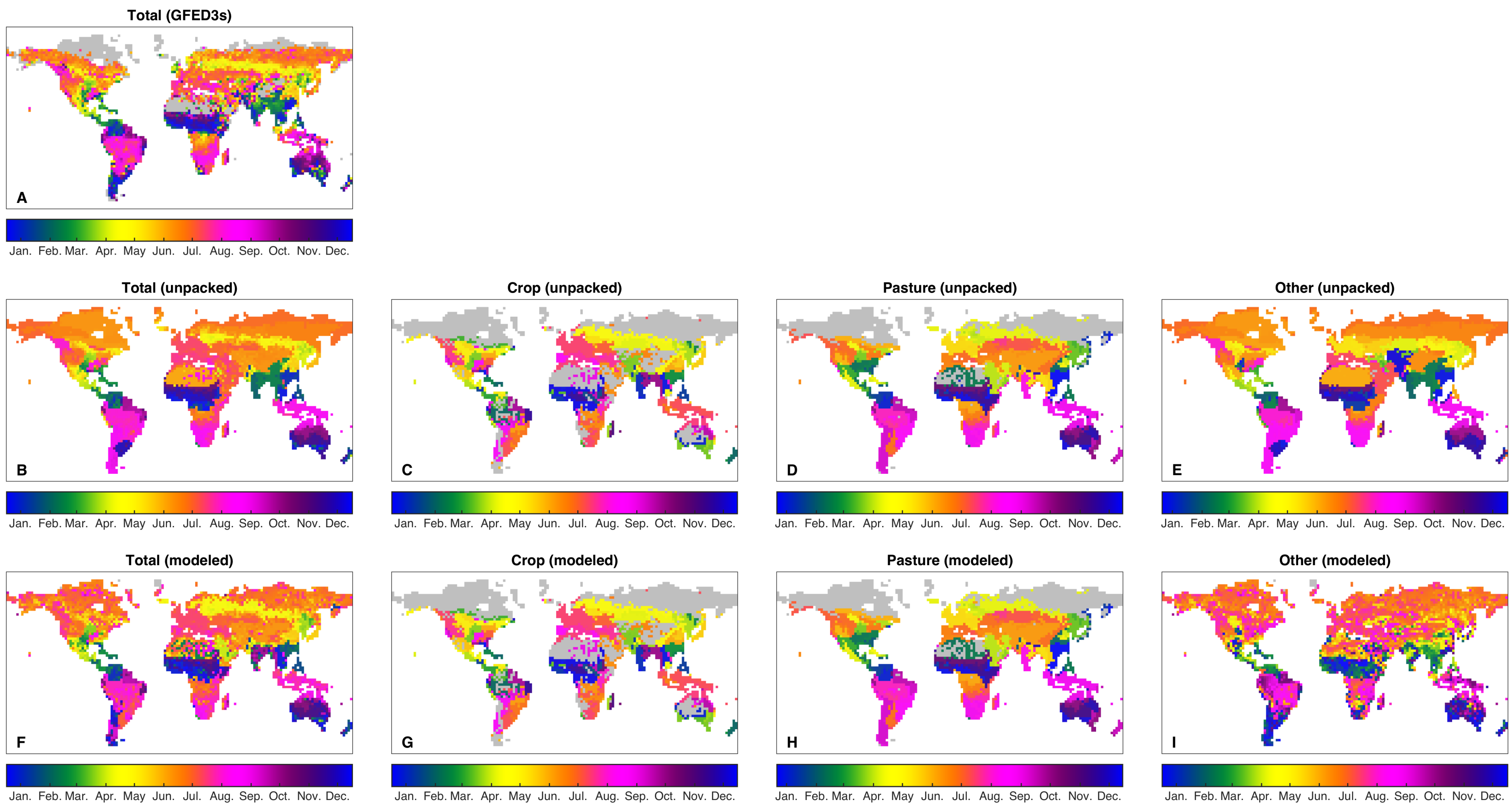

Figure 2.12: Mean timing of peak burned area **(a)**: From GFED3s (Randerson et al., 2012); (B–E): observational estimates from Rabin et al. (2015); (F–I): Model-estimated. Unburned cells in each map are colored gray. Tick marks and labels placed on the 15th of each month.

2.5.3 Effect of fire tiling on equilibrium pre-industrial potential biomass

The experiment described in Section 2.3.4 revealed that the creation of new tiles for burned areas greater than 1 km^2 had a noticeable effect on average global living biomass, which over years 271–300 of the experiment was simulated to be 1,020 GtC for the tiling-on run but only 918 GtC for the tiling-off run – an increase of more than 11% when using fire tiles. Figure 2.13 shows that this difference was not spread uniformly across the globe, instead being concentrated mostly in regions that today are tropical and temperate grasslands and savannas.

Figure 2.14 shows that our hypothesis was only somewhat correct: The biomass-enhancing effect of fire tiles did increase with tiling-off burned area, but only up to about a mean annual burned fraction of 10% of the gridcell. As mean annual burned fraction increased above that, the biomass-enhancing effect of fire tiles was diminished. In fact, above 10% mean annual burned fraction, it becomes much more common to see *decreased* live biomass density with fire tiling on. The possible significance of these results is discussed in more detail in Section 2.6.2.

2.6 Discussion and conclusions

2.6.1 Model performance in context

Modeled mean annual global burned area on non-agricultural lands (1.66×10^6 $km^2\,yr^{-1}$) was 32% lower than the unpacked estimate (2.44×10^6 $km^2\,yr^{-1}$; Table 2.4). The unpacking did tend to overestimate total global burned area by 0.25×10^6 $km^2\,yr^{-1}$, but even if we were to reduce the unpacked estimate of fire associated with non-agricultural land by that amount, the model would still fall short by 24%. While this indicates that there is room to improve FINAL, it is not an altogether unreasonable approximation at the global scale.

The tendency of `FINAL_V1` to underestimate total global 2001–2009 burned area is reflected in an underestimate of the associated carbon emissions – by 6% and 9%, respectively (Table 2.4). GFED3s and the unpacking data show respective average emissions densities of 0.53 and

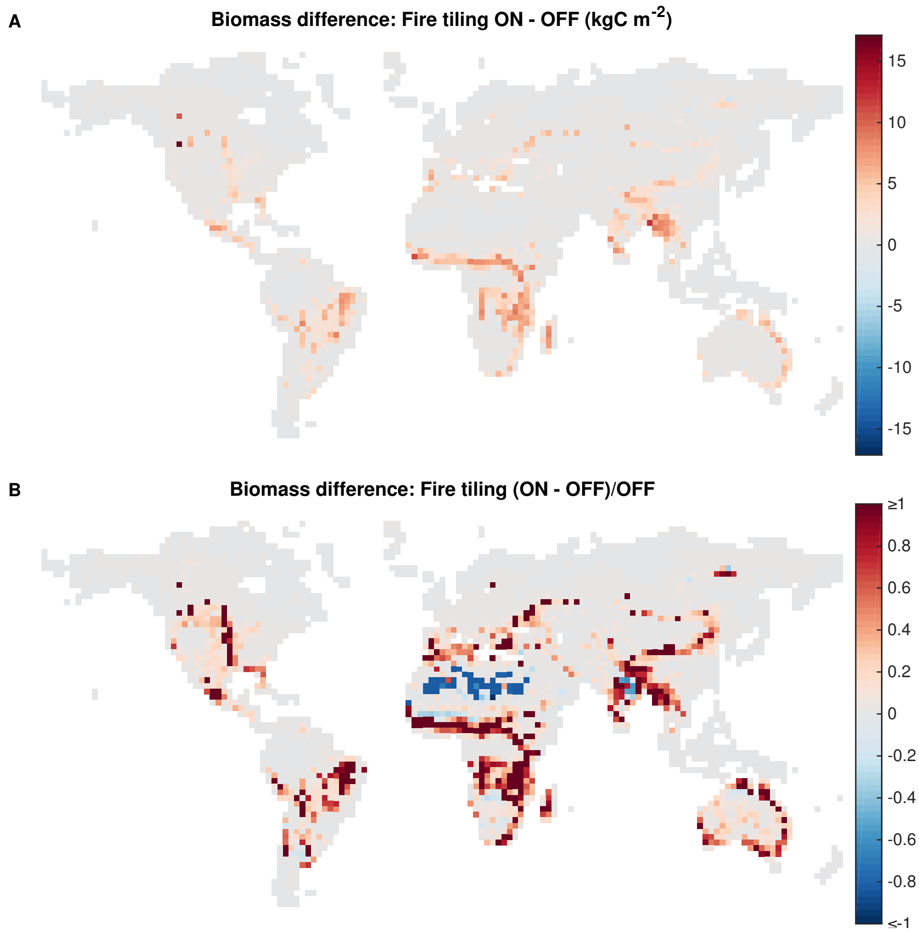

Figure 2.13: Difference in mean biomass density between years 271–300 experimental runs with fire tiles off and fire tiles on. Positive values indicate gridcells where the tiling-on run resulted in higher biomass. **(a)** Absolute difference ($\mathrm{kgC\,m^{-2}}$). **(b)** Fractional difference.

0.52 $\mathrm{kgC\,m^{-2}}$ of burning for all fire combined, whereas `FINAL_V1` gives 0.57 $\mathrm{kgC\,m^{-2}}$ (based on Table 2.4). The largest discrepancy in fire carbon emissions density between the modeled and unpacked estimates is on cropland, where `FINAL_V1` simulates 0.68 $\mathrm{kgC\,m^{-2}}$ but the unpacking analysis gives only 0.43 $\mathrm{kgC\,m^{-2}}$ (58% overestimate; Table 2.4). Emissions densities on pasture and non-agricultural land are also overestimated, respectively by 35% and 6.7%. This consistent pattern of overestimating the amount of emissions per burned area, especially on agricultural tiles, could be caused by a number of factors, which are discussed in Section 3.5.1.

In terms of spatial distribution, the model tends to over-cluster non-agricultural burned area relative to the unpacked estimate. That is, it tends (especially in savanna regions) to simulate a highly spatially heterogeneous distribution of non-agricultural burned area, with some areas

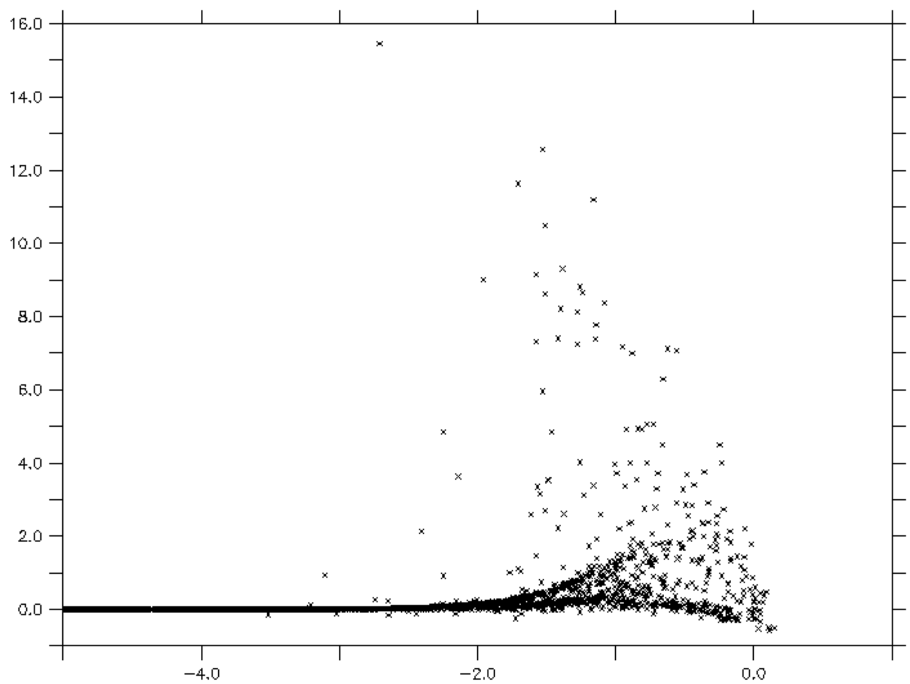

Figure 2.14: Scatter plot illustrating, over years 271–300 of the fire tiling experiment, the relationship between mean annual tiling-off burned fraction (X axis; $\log_{10} \mathrm{month}^{-1}$) and fractional difference in biomass density between tiling-on and tiling-off runs (Y axis, where positive values indicate greater biomass density in tiling-on run; $\frac{on-off}{off}$).

burning very little and others burning far too much (Fig. 2.8). It is important to consider, however, that although the unpacking method generates accurate estimates of total burned area at the level of each analysis region, the burning tends to be too evenly distributed within each region (Ch. 1). This results in an overly smooth map, as can be seen by comparing maps A and B in Figure 2.8. Non-agricultural burning in the real world might thus exhibit more spatial clustering than is apparent in Figure 2.8e. To get a sense of the spatial clustering of real-world non-agricultural fire, we have constructed a map of mean annual "GFED3s non-agricultural" burned fraction by subtracting unpacked cropland and pasture burned fraction from mean annual GFED3s total burned fraction.[6] We can then compare the coefficient of variation in 6×6 gridcell (12° latitude $\times$ 15° longitude) kernels across this map with similar maps for mean annual modeled and unpacked non-agricultural fire. As expected, the coefficient of variation is much higher in the GFED3s data than the unpacked data (Fig. 2.15), indicating stronger spatial clustering of non-agricultural fire in the real world. The fact that the model

[6]The exact numbers from this map are not very meaningful, since it is possible to have values less than zero in gridcells where unpacking estimated more cropland and pasture burning than all burning observed by GFED3s. The purpose of this exercise is only to examine spatial variation.

simulates more heterogeneity than the unpacked estimate, then, indicates that the model is capturing heterogeneity in fire drivers that are important to actual fire patterns. This is not to say, of course, that the heterogeneous patterns simulated by the model exactly match the observations – in some places they do not, as is apparent in Figure 2.8.

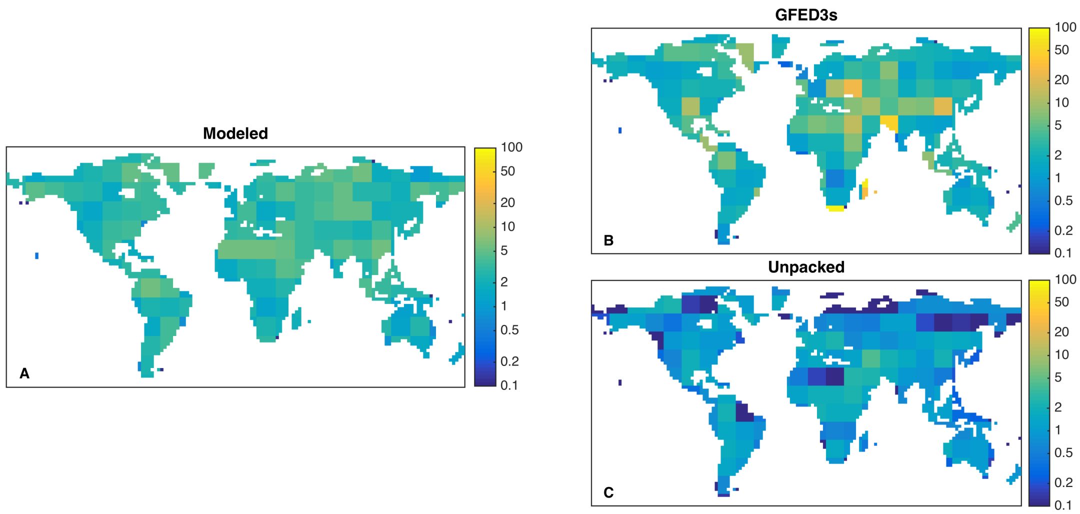

Figure 2.15: Coefficient of variation (standard deviation divided by mean) of non-agricultural burned fraction in 6×6 gridcell kernels (12° latitude $\times$ 15° longitude). **(a)** Modeled; **(b)** from artificially-constructed GFED3s non-agricultural fire data as described in text; **(c)** unpacked. Note log scale of color bars.

Although savanna regions may have shown the largest absolute difference in modeled vs. unpacked fire activity, smaller differences can be just as important in other areas. For example, the GFED3s and unpacked data show a mean annual burned fraction of 1–5% for the boreal forests of central Alaska and northwestern Canada (Figs. 2.8a–b,e), which would correspond to a mean fire return interval of 20–100 years. While this is a low rate of burning relative to, e.g., tropical savannas, it still represents an important process for the structure and function of that ecosystem. The non-agricultural fire model captures almost no boreal forest fire whatsoever (Fig. 2.8i), which should hamper the ability of LM3 to accurately simulate vegetation there. One possible contribution to this deficit is the importance of multi-day fires in the boreal region, whereas we followed Li et al. (2012) in assuming that all fires last 24 hours. This assumption is not well-supported by the literature. Korovin (1996) found that almost 60% of forest fires in Russia over 1947–1992 lasted longer than one day, and that fires lasting longer than 10 days accounted for nearly 70% of the burned forest area. Stocks et al. (2003) found a similar importance of very large (and thus presumably long-lasting) fires in Canada, with individual burns of more than 20,000 ha comprising over 65% of mean annual burned area over 1959–1997. Ideally, FINAL would replicate this pattern by explicitly modeling the duration of individual fires based on evolving weather conditions. Several global fire models have introduced such a component, but with mixed results. The LPJ-LMfire model developed by Pfeiffer et al. (2013), which allows fires to burn for about four hours per day until they experience significant precipitation, actually tends to *overestimate* boreal forest fire. The HESFIRE model (Le Page et al., 2015) also allows fires to burn indefinitely, calculating twice per day an extinction probability based on fuel load, attempted suppression intensity, landscape fragmentation, and weather conditions. However, like FINAL, HESFIRE simulates too little fire in the boreal region (Le Page et al., 2015).

2.6.2 Importance of fire tiling

Explicitly keeping track of fire history through the use of "fire tiles" resulted in a large increase in simulated biomass for potential vegetation (Sect. 2.5.3), especially in fire-prone tropical and temperate grasslands and savannas. This could be the result of a phenomenon where, without fire tiling, the biomass in a gridcell would become too low to allow much fire (i.e., f_{AGB} would be near zero; Eq. 2.10), at which point the gridcell would enter a fire-free period during which it would regain biomass until fire was again possible. With fire tiling, any part of the tile could have enough biomass to burn regardless of how recently some other part of the tile had burned. This is essentially the opposite effect of what we observed for intermediate levels of burning (and what we hypothesized). Because of the strength of the effect of fire tiling on biomass, more research to understand the underlying mechanisms driving the effect would be beneficial.

2.6.3 What do optimization results suggest?

Perhaps the most interesting result of the optimization was that it effectively excluded relative humidity from exerting any effect on fire activity, shifting all of the control of flammability to soil moisture (Fig. 2.5c, d). This suggests that, at the coarse spatiotemporal scale considered in global fire modeling, the moisture of the upper soil may be a much better proxy for fuel moisture than relative humidity. Physical mechanisms for this suggest themselves readily. Live fuels such as the herbaceous layer in grasslands and savannas have access to soil water that, even in the upper soil, likely fluctuates less over short time scales than relative humidity. Soil moisture might thus be a better predictor of the seasonal timing of fire. However, this may only have emerged here because our algorithm compared modeled to observed burned area on a monthly basis. If instead we had compared daily modeled versus observed burned area, relative humidity might have proven important.

The fact that soil moisture suppressive effect does not abate even for the driest soils – that is, $f_\theta(\theta = 0) \approx 0.7$ instead of 1 (Fig. 2.5d) – is another intriguing result. Because $f_\theta(\theta = 0) = exp\left(-\beta_{\theta,1}\right)$ (Eq. 2.8), it would have been reasonable to constrain $\beta_{\theta,1}$ during the optimization to prevent $f_\theta(\theta = 0)$ from being below 0.999 or some other value close to unity. Such a strategy would arguably even make physical sense – soil moisture can hardly limit fire if there is no moisture in the soil. This would presumably have the effect of increasing burned area at low soil moisture, but that might not be the case. It's possible that very few gridcells ever actually experienced such low soil moisture, and/or such cells were limited by other factors – chronically low soil moisture (or average conditions in regions that ever experienced such an extreme) would result in low aboveground biomass, for example. If true, this could mean that the result of $f_\theta(\theta = 0) \approx 0.7$ may have been essentially a spurious effect, since the algorithm would not have been very sensitive to f_θ at such low values of soil moisture. On the other hand, this might be a real effect, in which case there may be a more structural issue with the fire model. A simple scaling factor – some extra constant that reduces ignition density, for instance – could be a useful addition in that case, but would have the function of decreasing fire in all gridcells.

The result at the other end of the soil moisture spectrum – i.e., that soil moistures above about 0.35 prevent almost all fire from occurring, whereas the initial guess didn't restrict so severely until about $\theta = 0.65$ (Fig. 2.5d) – is also interesting. Le Page et al. (2015), in the manual phase of their model development, decided that soil moisture would prevent all burning above $\theta = 0.35$ as well. Although soil moisture in that model only affected rate of spread and not also ignition success rate as it does in our model, and although they also allowed relative humidity to affect rate of spread in a manner similar to Li et al. (2012), the fact that our optimization's result corresponded so closely with their parameter choice is intriguing.

Optimization resulted in fewer anthropogenic ignitions and stronger anthropogenic suppression for any given value of population density (Fig. 2.5a–b, e). This suggests that, by confounding non-agricultural fires with pasture fires, previous modeling efforts may have overestimated the contribution of humans to burning on non-agricultural land. That is, by extracting a "pure" non-agricultural fire signal, our study shows that pasture burning practices may have been responsible for much of what was once characterized as general anthropogenic fire, and that humans enhance fire on non-agricultural lands less than once believed. In terms of the general shape of net anthropogenic influence on non-agricultural fires – including the location and width of the peak – our results do not differ substantially from the function described by Pechony and Shindell (2009) or that used by Li et al. (2012; Fig. 2.5e). Knorr et al. (2014), on the other hand, used the Levenberg-Marquardt algorithm to fit a simple empirical fire model in a non-interactive fashion and found that the peak was actually located closer to a population density of 0.1 people km^{-2} than to the value of ~ 10 people km^{-2} that we found here.

Just like cropland and pasture fire management practices, the way people affect fire regimes on non-agricultural land varies around the world. Some fire models include a spatially-dependent human ignitions term (Thonicke et al., 2010) to account for this effect. Incorporating this geographic variation into FINAL could improve performance, but it would be important to do so based on independent analyses so as to avoid simply compensating for the model's errors.

2.6.4 Levenberg-Marquardt optimization: Lessons learned

One of the limitations of the Levenberg-Marquardt algorithm is that it can only "move downhill." At every iteration, it searches for new parameters in the direction of lower sum of squared errors from the current point in parameter space, even though the set of parameters with the lowest possible sum of squared errors may be in a totally different direction. As an analogy, imagine a person given the task of finding the point in a city with the lowest elevation above sea level.

Using a "downhill-only" algorithm, this person would literally walk downhill from their starting point and stop when they reach a point – the local minimum – where continued travel in any direction would be uphill. The person might more thoroughly search the city for its lowest point by occasionally turning uphill and/or randomly taking a bus once in a while to a totally different part of the city – analogous to the behavior of the Metropolis Markov Chain Monte Carlo or simulated annealing algorithms. Levenberg-Marquardt being a downhill-only algorithm is not a fatal flaw, of course, especially when the initial parameter set guess is well-informed based on the literature. It may well represent an improvement in methodology over the manual trial-and-error approach. But it is important to remember that Levenberg-Marquardt cannot be expected to produce the universally best possible parameter set.

Another, potentially more serious limitation of the Levenberg-Marquardt algorithm is its use of the sum of squared errors (SSE) as a metric to gauge model performance. While the setup used here does account for accuracy of burned area simulations in both space and time, SSE tends to result in a bias towards improving performance in gridcells where the model simulates burned areas much higher or much lower than observations. This tendency to reduce absolute error would be fine if the goal of optimization were to produce a model that accurately simulates burned area for its own sake, but *relative* error can be more reflective of how well the model simulates the state of the vegetation. For example, assume two hypothetical 1,000-km^2 gridcells: one dominated by tropical grassland where observations show 100% annual burning but the model simulates 25%, and one dominated by boreal forest where observations show 1% annual burning but the model simulates 0.25%. In both cases, the model is producing 75% less fire than what actually happens – a difference that could be extremely important to the simulated structure and function of both ecosystems. However, because the absolute error in the grassland gridcell (-750 km^2 yr^{-1}) is so much greater than that in the boreal forest gridcell (-7.5 km^2 yr^{-1}), the former will, all other things being equal, have a much greater influence on the direction and magnitude of the step towards the next parameter set guess.[7] Because the observations show that tropical savannas burn far more than any other biome, the absolute errors are highest there

[7] The fact that we used an equirectangular grid – with cells of constant size in units of latitude and longitude but not physical area – means that cells from high latitudes are much smaller than cells from the tropics, which would exacerbate the problem described here.

(Fig. 2.9). These regions thus likely drive most of the optimization, which could have led to the neglect of performance in, for example, the boreal region. An optimization algorithm that took relative error into account might thus improve performance in low-fire regions, while worsening it where fire is frequent.

Simply substituting an alternative measurement for SSE in a Levenberg-Marquardt context would be less than ideal for addressing this problem. In addition to being the performance metric – i.e., the statistic by which the algorithm determines whether a parameter set has resulted in improved model performance – SSE is an inherent part of the mathematics in the Levenberg-Marquardt algorithm generating the direction and size of the step from the most recently accepted guess to the next accepted guess (Appendix B). Using a different performance metric would still result in guesses designed to minimize the sum of squared errors, reducing the efficiency of the algorithm at best, and at worst resulting in searches orthogonal to the direction of improved performance. To most effectively avoid the problems inherent with SSE, a completely different algorithm – preferably one that can use any arbitrary performance metric – would be needed. The Markov Chain Monte Carlo method (MCMC) is one such option, which has the additional benefit as discussed above of being a global search algorithm. It has been widely used in the Earth sciences, including by Le Page et al. (2015) to fit a global fire model. Those authors used as their performance metric a combination of (a) accuracy of classification of gridcells into burned fraction bins[8] and (b) level of correspondence between model-simulated and observed interannual variability. However, being a global search, MCMC requires many iterations to converge on an optimal solution – Le Page et al. (2015) reported iteration counts of hundreds to over a thousand. The deeply model-interactive setup used here – where the complete model of soil, vegetation, and fire was forced with climatic data for nearly twenty model years – took around two hours per iteration with all gridcells being run in parallel, which made MCMC and similar many-iteration algorithms computationally infeasible.

The choice of gridcells and initial conditions is also extremely important to any automated model fitting algorithm. The strong effects we saw in preliminary optimization runs of including a few extra gridcells from badly-modeled regions make this quite clear. The process through which

[8]Bin edges used: 0%, 1%, 5%, 10%, 20%, 35%, and 50+% annual burned fraction.

we settled on our set of 241 gridcells was admittedly haphazard, and a more structured and informed approach would likely make the results more robust. Similarly, we did not experiment much with different initial parameter set guesses, but doing so is a good way to test model robustness (Knorr et al., 2014; Le Page et al., 2015).

2.6.5 Moving towards a prognostic model of anthropogenic fire practices

FINALv1 represents an important step toward including cropland and pasture fire management in Earth system models. However, it is limited in its utility by its reliance on set climatologies of burning. The assumption is that, for example, pastoralists in one part of west Africa burn the exact same fraction of their pastures every January, some other fraction every February, and so on. This obviously does not allow for any interannual variability in burning associated with variation in flammability and associated difficulty in controlling burns. This deficiency is evident in Figure 2.10, where modeled pasture burned area (solid purple line) exhibits much less variation across the study period than shown by the unpacked data (dashed purple line).

Perhaps more importantly, however, the use of a climatology based on just nine years of observations makes it difficult to justify the use of the model very far into the past or future. Economic development can result in changes in technology, types of crops, and legislative priorities (banning crop fires, for example), all of which can affect the amount and timing of agricultural fire. Climate change has and will continue to affect the timing, length, and quality of growing seasons; the associated impacts on planting and harvest date will affect the timing of crop residue burning, and people will shift the timing of burns to match the shifting phenology of pasture vegetation. It is thus important to understand what information people consider in their decisions of whether, when, and how much to burn. Literature reviews and new research could shed light on indigenous methods for climate forecasting based on changes in the weather and vegetation (e.g., Kagunyu et al., 2016), as well as how these cues might be tied to the timing of prescribed fire for various purposes (e.g., Laris, 2002). Advanced analytical methods could also be applied to climate and fire history observations to look for lagged, region-specific relationships of agricultural burning with weather at weekly to monthly time scales.

A more complete model of cropland and pasture burning practices would need to consider all of these factors in order to be useful for investigating the historical and future impact of fire on the Earth system.

Chapter 3

Carbon consequences of pasture fire management

3.1 Abstract

The Fire Including Natural & Agricultural Lands model (FINAL; Ch. 2) simulates, for the first time, the way that people manage both cropland and pasture using fire. Despite comprising a substantial portion of Earth's land, pasture has not previously been considered as a kind of land cover with possibly distinct fire regimes from grassland or savanna not being grazed by livestock. Here, we use FINAL to explore the effect of pasture fire management on carbon cycling and reservoirs at regional and global scales. We find that incorporating pastoral burning practices boosts simulated burned area by 66% (from 1.22×10^6 km^2 yr^{-1} to 2.02×10^6 km^2 yr^{-1}) and fire carbon emissions by over 550% (from 109 TgC yr^{-1} to 712 TgC yr^{-1}), bringing simulated pasture fire activity much more in line with observations (2.02×10^6 km^2 yr^{-1}, 538 TgC yr^{-1}). Despite the large relative change in simulated pasture fire emissions, the net exchange of carbon between terrestrial ecosystems and the atmosphere at a global level is not appreciably affected. Important differences in carbon fluxes and vegetation structure are seen, however, across tropical and temperate grasslands, woodlands, and savannas. These differences could be critical for accurately representing the distribution of these biomes in relation to each other and to forests.

3.2 Introduction

Humans are causing serious changes to the Earth's climate by increasing the concentrations of greenhouse gases in the atmosphere, especially carbon dioxide (CO_2). Between 2000 and 2009, our combustion of fossil fuels and use of concrete emitted 7.8 ± 0.6 PgC yr^{-1} in the form of CO_2 emissions (90% confidence interval; Ciais et al., 2013). Over the same time period, it has been estimated that we effectively emitted 1.1 ± 0.8 PgC yr^{-1} as a net effect of land use and management such as deforestation, logging, and agriculture (Ciais et al., 2013). Terrestrial ecosystems reduced the overall impact of these emissions by taking up 2.6 ± 1.2 PgC yr^{-1}, but the extent to which this effect will continue is in question. Climate and land use change could increase rates of disturbance, decomposition, and respiration while also decreasing photosynthesis – potentially shifting the terrestrial biosphere from a sink of CO_2 to a source.

The difficulty of this problem becomes immediately apparent when considering how challenging it is to even simulate present-day conditions. For example, in the uncertainty ranges given above: A 90%-confidence estimate of fossil fuel and concrete CO_2 emissions represents a range of less than ±8%; the estimate of our net land use effect, on the other hand, varies by over ±70%. To explore the future of the terrestrial carbon sink, we must first improve our understanding of what drives carbon cycling in terrestrial ecosystems. This represents a complex challenge because of the intricate ways various parts of the terrestrial biosphere interact, as well as the interconnections between ecosystems and other parts of the Earth system.

Vegetation fire sits at the intersection of a number of important Earth system processes. The combustion of vegetation releases greenhouse gases, mostly in the form of CO_2, as well as aerosols that can variously contribute to warming, cooling, and even changes in cloud formation and precipitation. Fire can also be an important regulator of vegetation state, such as by mediating semi-permanent shifts from forests to lower-biomass savannas. Climate change can be expected to increase burning in some parts of the world and decrease it in others, with potential consequences for the status of the terrestrial biosphere as a carbon sink.

But in addition to our effect on fire through climate, we also directly control the fire regime in many parts of the world to suit our management needs. Farmers often burn before planting or harvest, and pastoralists use fire to promote the growth of nutritious plants for their livestock. These management fires often differ from how fires would occur naturally, because agriculturists sometimes require burning at different times of the year – or at a different frequency – than their land would experience naturally. Because cropland and pasture together cover over a third of the Earth's surface, and because that number expected to increase over the coming century, the net effect of these burning practices on carbon cycling could be significant.

The Fire Including Natural & Agricultural Lands model version 1 (FINALv1), which combines a model of burning on non-agricultural land with estimates of burned area associated with cropland and pasture, was developed in part to explore this question. The design and performance of the model were described in Chapter 2. Here, we present the results of experimentation using FINAL to test the impact of cropland and pasture burning on terrestrial carbon cycling.

3.3 Methods

3.3.1 Modeling

The land, vegetation, and fire model used in this chapter was identical to that described in the previous chapter (Sect. 2.3). We also used the same input data (Sect. 2.4.2) and generated initial conditions the same way (Sect. 2.4.1).

In this chapter, we present the results of an additional experiment designed to test the role of pasture management fire in altering fire regimes, as well as to assess the sensitivity thereto of the fire model. Specifically, we performed a "pasture as non-agricultural" run (`PASTASOTHR`; Table 3.1) that was identical to `FINAL_V1` (Table 2.2; referred to in this chapter as `PASTASPAST`) except for the fact that the non-agricultural fire model was used on pasture instead of the pasture burning climatology derived from the unpacked data (Ch. 1).

Table 3.1: Experimental runs discussed in this chapter. PASTASPAST is referred to in Chapter 2 as FINAL_V1 (Table 2.2).

	PASTASPAST	PASTASOTHR
Fire model	New, optimized	New, optimized
Years	1948–2009	1948–2009
Init. conditions	As LM3_ORIG	As LM3_ORIG
Climate	Historical	Historical
CO_2	Historical	Historical
Land use	Historical	Historical
Non-agri. fire: Humans	On	On
Cropland fire	As unpacked	As unpacked
Pasture fire	As unpacked	As non-ag. land

3.3.2 Analysis of results

Because land carbon dynamics are so complex, we will present results from a number of different variables describing both the state of vegetation and the rate of various fluxes between the land and the atmosphere. Gross primary productivity (gpp) refers to the rate at which plants absorb carbon from the atmosphere in photosynthesis. Plants also emit carbon as they burn sugars in respiration (resp), and so net primary productivity (npp) is used to denote the net amount of carbon brought in to the biosphere by plants. Microbes decomposing dead biomass emit carbon from the litter and soil, a process referred to as heterotrophic respiration (rh). The net flux of carbon between the biosphere and atmosphere considering these processes is referred to as net ecosystem production (nep), but that does not paint a complete picture. Carbon emissions from combustion (burnCemitRate) and emissions associated with harvested or grazed biomass (crop/past/wood_harv_rate) also serve to move carbon from terrestrial ecosystems to the atmosphere. The net flux considering net ecosystem productivity, fire disturbance, and land use is referred to as net biosphere production (nbp). These and other fluxes, along with a number of living and dead biomass pools describing the amount of carbon present in ecosystems, are summarized in Table 3.2.

It is important to note that net biosphere production (nbp) is the single statistic that can best describe the net effect of pasture management fire on the land carbon balance. Counterintuitively, although fire obviously transfers much carbon from living and dead vegetation to the

Table 3.2: Modeled carbon pools and fluxes discussed in this chapter.

Abbreviation	Description	Type
agb	Aboveground living biomass (bl+blv+0.8*[bwood+bsw])	Pool
bgb	Belowground living biomass (br+0.2*[bwood+bsw])	Pool
bl	Living biomass in leaves	Pool
blv	Living biomass in labile C store	Pool
br	Living biomass in fine roots	Pool
bsw	Living biomass in sapwood	Pool
btot	Total living biomass	Pool
burnCemitRate	Fire combustion rate	Flux
bwood	Living biomass in heartwood	Pool
co2litter, leaf	Decomposition of leaf litter	Flux
co2litter, wood	Decomposition of woody litter	Flux
co2soil	Decomposition of soil	Flux
crop_harv_rate	Emissions associated with harvested crop biomass	Flux
gpp	Gross primary productivity	Flux
litter, leaf	Dead biomass in leaf litter	Pool
litter, wood	Dead biomass in woody litter	Pool
nbp	Net biosphere production (nep-[burnCemitRate+totl_harv_rate])	Flux
nep	Net ecosystem production (npp+rh)	
npp	Net primary productivity	Flux
past_harv_rate	Emissions associated with grazed pasture biomass	Flux
resp	Primary respiration	Flux
rh	Heterotrophic respiration	Flux
totl_harv_rate	Emissions associated with all harvesting and grazing (crop_harv_rate + past_harv_rate + wood_harv_rate)	Flux
wood_harv_rate	Emissions associated with harvested wood biomass	Flux

atmosphere via combustion, much of what burns is dead material that would have been emitted eventually via decomposition. On balance, fire can have a net neutral or even positive effect on ecosystem carbon retention (Yi et al., 2013; Yue et al., 2015, 2016). It is thus important to examine measurements beyond a simple tallying of carbon emissions from fire.

Table 3.3: Global mean annual burned area and associated carbon emissions. (1) Midpoint of values for cropland burning with (0.208) and without (0.456) including cropland-natural mosaic.

Metric	Land cover	GFED3s	Unpacked	Mod. past. as non-ag.	Mod. past. as unp.
Burned area (10^6 km^2)	Total	4.68	4.93	2.41	4.11
	Crop	0.332$^{(1)}$	0.454	0.434	0.434
	Pasture	—	2.04	1.22	2.02
	Non-ag.	—	2.44	1.86	1.66
C emissions (PgC)	Total	2.48	2.57	0.935	2.34
	Crop	n.d.	0.194	0.298	0.297
	Pasture	—	0.538	0.109	0.712
	Non-ag.	—	1.84	0.528	1.33

3.4 Results

3.4.1 Burned area and C emissions

The default fire model (`PASTASPAST`), as expected, simulated the burned area associated with pasture quite accurately relative to the unpacked estimates with which it was forced (Fig. 2.9, Table 3.3). When allowing pasture to burn according to the non-agricultural fire model (`PASTASOTHR`), simulated pasture burned area dropped by 40%, representing a much stronger underestimate relative to the unpacked data than the default run (Table 3.3). On non-agricultural land, however, `PASTASOTHR` simulated 12% more burned area, even though the non-agricultural fire model was the exact same for both runs (Table 3.3). One possible explanation of this could have to do with biomass levels after land abandonment. Pasture that burns less frequently, as in `PASTASOTHR`, would have more fuel available and thus be more likely to burn after being abandoned and transitioned to non-agricultural land where low fuel loads might otherwise limit fire.

The pasture fire carbon emissions total modeled by `PASTASPAST` of 0.712 PgC yr^{-1} represents a 32% overestimate compared to the unpacked estimate of 0.538 PgC yr^{-1} (Table 3.3). The great majority of the overestimate is associated with pasture burning in tropical savanna regions, although some pasture in Sudan shows an underestimate (Fig. 2.9h). `PASTASOTHR` reduced simulated fire carbon emissions even more drastically than it did burned area: Total global emissions were 0.109 PgC yr^{-1}, 60% (1.41 PgC yr^{-1}) lower than in `PASTASPAST` (Table 3.3).

Pasture and non-agricultural fire emissions were 0.603 $\mathrm{PgC\,yr^{-1}}$ (85%) and 0.802 $\mathrm{PgC\,yr^{-1}}$ (60%) lower, respectively, while cropland fire emissions increased slightly (0.001 $\mathrm{PgC\,yr^{-1}}$, 0.3%; Table 3.3).

Basing pasture fire on the unpacked data instead of letting pasture tiles burn like non-agricultural tiles noticeably improved pasture fire seasonality (Fig. 3.1), from an average global error of 68 days later than expected to an average error of 2.0 days too late.

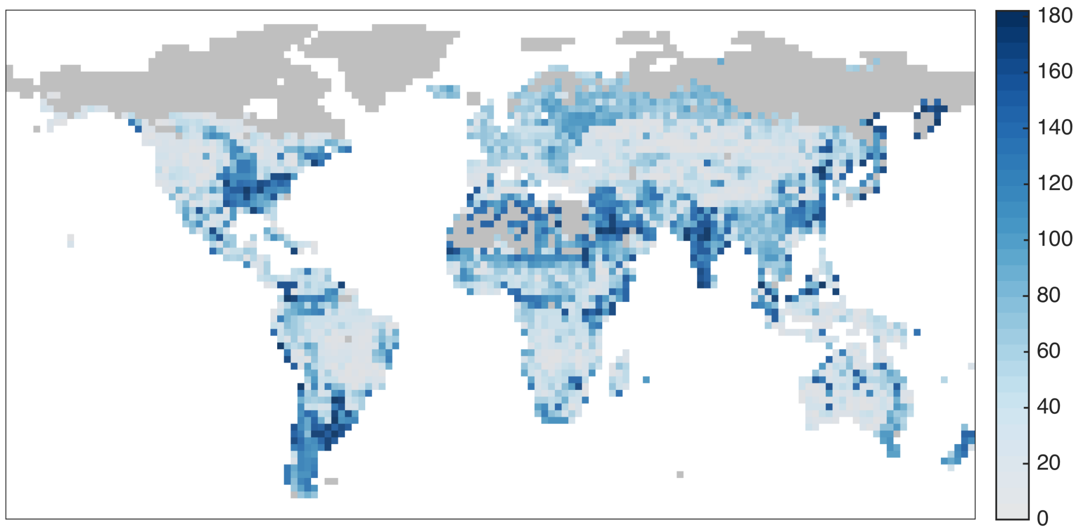

Figure 3.1: Improvement (days) in absolute error of simulated mean peak pasture burning when switching from fire as non-agricultural to unpacking-based fire. Dark gray: Pasture did not burn at all in one or both, or no pasture was present.

3.4.2 Annual C balance

At a global level, forcing pasture with the unpacked burned area caused no difference in net primary productivity (`npp`) but did marginally reduce heterotrophic respiration relative to the `PASTASOTHR` run, although this resulted in no significant change in simulated net ecosystem productivity (Fig. 3.2). `PASTASPAST` did generate significantly more fire carbon emissions than `PASTASOTHR` (Table 3.3) but did not appreciably alter the flux from total harvested biomass (totl_harv; Fig. 3.2). Overall, mean annual global net biosphere production (nbp) over 2001–2009 was reduced by 0.23 $\mathrm{PgC\,yr^{-1}}$ (6.2%) when using the unpacked burning as compared with fire as simulated by the non-agricultural model (Table D.1), but with a relatively large amount of interannual variability (Fig. 3.2).

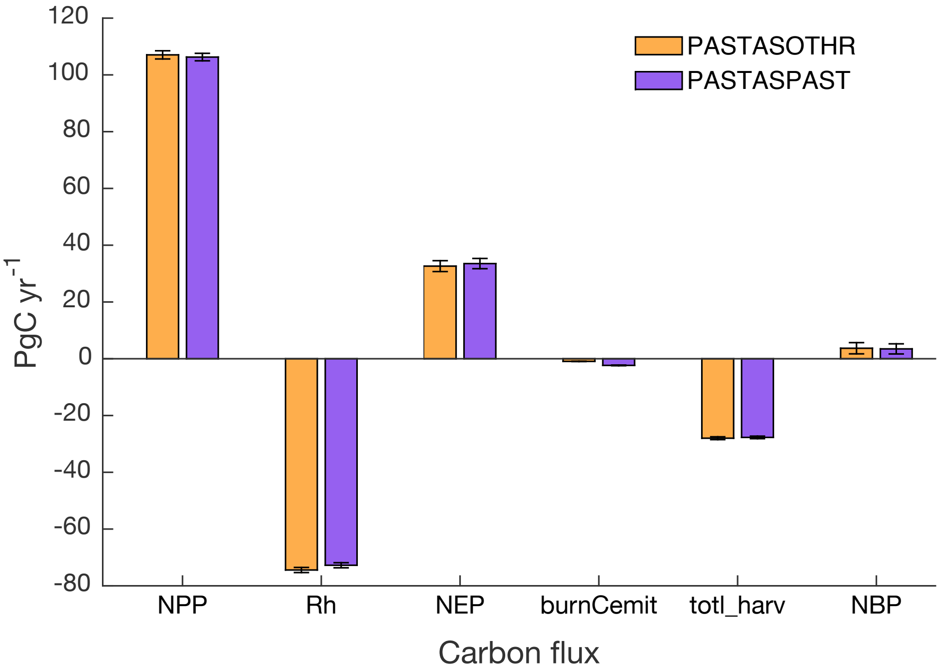

Figure 3.2: Mean annual global carbon fluxes simulated over 2001–2009. Error bars indicate 95% confidence interval based on Student's t distribution.

These effects of pasture management fire are not evenly distributed across the globe. When the differences between `PASTASOTHR` and `PASTASPAST` are mapped in terms of carbon fluxes (Fig. D.3) and pools (Fig. D.5), it is apparent that the spatial distribution of the most important contributing gridcells regions reflects where pasture fire emissions are large (2.11). These values can be misleading, however, when considering the impact on vegetation – an ecosystem of low productivity can be greatly impacted by even a small difference in fire frequency. The largest impacts of pasture burning practices on GPP and NPP relative to their values in `PASTASOTHR`, for example, occur in the Asian temperate zone (Figs. 3.3a–b).

Similarly, although the net effect of pasture burning practices on net primary productivity (`npp`) was negligible at the global scale, it does appear to have been regionally important. The east Sudanian savanna, where pasture burns more than anywhere else in the world, saw strongly increased net primary productivity, while other parts of Africa and Asia saw a reduction (Figs. 3.3, D.1–D.2).

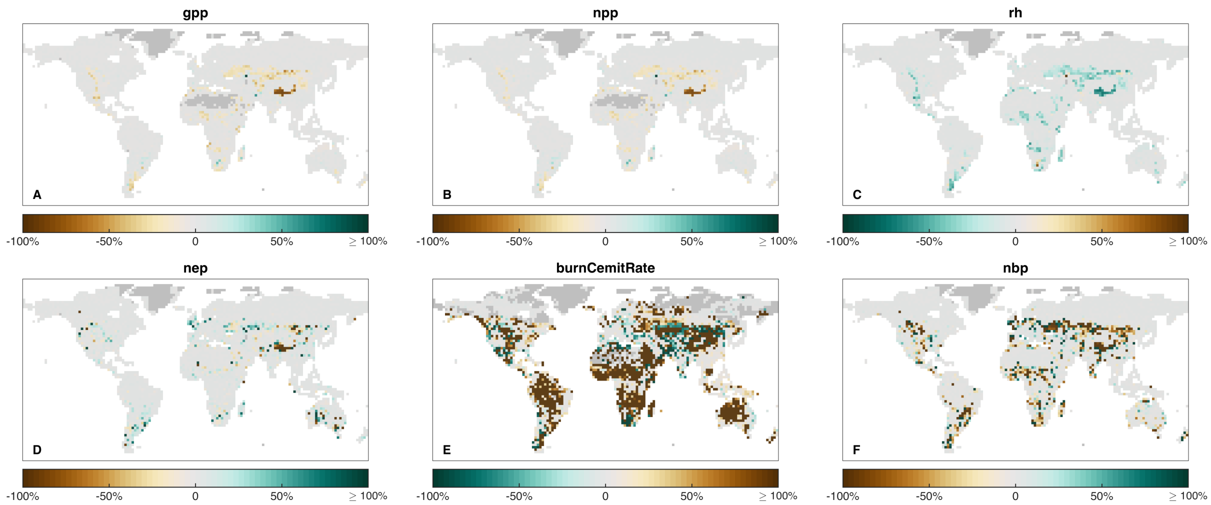

Figure 3.3: Percent difference in selected mean annual carbon fluxes between PASTASPAST and PASTASOTHR, 2001–2009. Green indicates increased ecosystem C retention in PASTASPAST; brown, reduced. Dark gray indicates land cells with no vegetation, or a value of zero for the given variable in PASTASOTHR. Note that this is for all vegetation, not just pasture. See Figures D.1–D.2 for separate maps of these and other fluxes for each run, and Figure D.3 for absolute difference.

Living and dead biomass pools were also affected by the simulation of pasture management fire. Madagascar and the steppes of the Tibetan plateau have much lower biomass simulated when forced with the unpacked pasture fire than when modeled to burn as non-agricultural lands (Figs. D.4–D.5). Important *relative* changes in biomass pools are widespread; pastures throughout Eurasia, the African tropical savannas, and temperate grasslands all see much less biomass in the PASTASPAST run in relative terms.

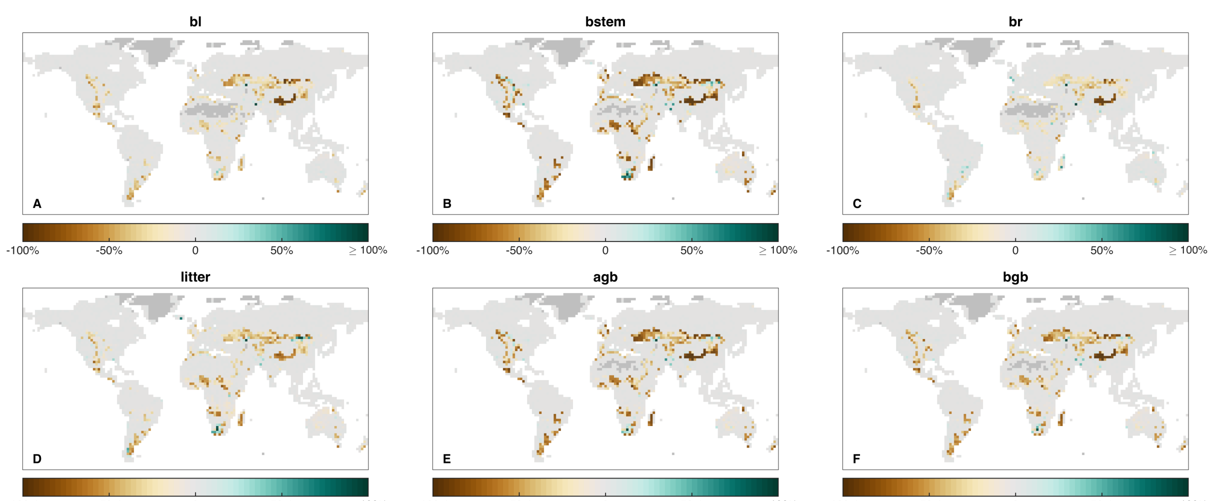

Figure 3.4: Percent difference in selected mean carbon pools between PASTASPAST and PASTASOTHR, 2001–2009. Green indicates increased ecosystem C pools in PASTASPAST. Dark gray indicates land cells with no vegetation, or a value of zero for the given variable in PASTASOTHR. Note that this is for all vegetation, not just pasture. See Figure D.4 for separate maps for each run, and Figure D.5 for absolute difference.

3.5 Discussion and conclusions

3.5.1 Exploring pasture fire carbon emissions

Carbon emissions from pasture fires, as with all fires, are the product of three quantities: burned area, aboveground biomass, and combustion completeness. Because the model simulates burned pasture area so accurately (Table 3.3), either or both of the latter two could have contributed to the overestimation of pasture fire emissions.

Aside from fire, a major determinant of aboveground biomass on pasture is the intensity with which it is grazed. With the rate of grazing set to 4% of leaf biomass each day, the `PASTASPAST` run simulated the consumption by livestock of 1.54 $\mathrm{PgC\,yr^{-1}}$ globally over 2001–2009. This compares favorably with previously-published estimates of carbon flows to livestock. Wirsenius (2000) estimated that domesticated grazers consumed 1.33 PgC in 1990, not counting draft animals. Krausmann et al. (2008), working on the year 2000, estimated that livestock (including draft animals) consumed 1.9 PgC. Haberl et al. (2007) estimated that the average grazing pressure on pasture for the year 2000 was 41 $\mathrm{gC\,m^{-2}}$, which again compares favorably with the simulated value from `PASTASPAST` of 45 $\mathrm{gC\,m^{-2}\,yr^{-1}}$ over 2001–2009.

On average over 2001–2009, `PASTASPAST` simulated 3.4 $\mathrm{kgC\,m^{-2}}$ of aboveground biomass on pastures, including both live vegetation and dead material. This was broken down into live leaves (0.22 $\mathrm{kgC\,m^{-2}}$), live stems (0.94 $\mathrm{kgC\,m^{-2}}$), leaf litter (0.45 $\mathrm{kgC\,m^{-2}}$), and dead woody material (1.8 $\mathrm{kgC\,m^{-2}}$); these pools are mapped for the world's major pasture regions in Figure 3.5. In their work in the Waikato region of New Zealand – a moist, temperate ecosystem dominated by C3 grasses – Hanna et al. (1999) defined active pastures as containing no more than 0.2 $\mathrm{kgC\,m^{-2}}$ of live leaves or 0.15 $\mathrm{kgC\,m^{-2}}$ of dead material. `PASTASPAST` simulated less than 0.1 $\mathrm{kgC\,m^{-2}}$ of live leaf tissue in New Zealand, and indeed the world's temperate pastures seem to satisfy the ≤ 0.2 $\mathrm{kgC\,m^{-2}}$ criterion (Fig. 3.5a). The tropics generally see much higher modeled pasture leaf biomass; in all cases, leaf biomass does not much exceed 0.25 $\mathrm{kgC\,m^{-2}}$ (Fig. 3.5a). Uhl and Kauffman (1990) describe a pasture in eastern Amazonia with 0.6 $\mathrm{kgC\,m^{-2}}$

of nonwoody material; this is close to the simulated value of combined live and dead leaf C (Fig. 3.5a,c) in the regions listed above. Kauffman and Cummings (1998), looking at three other pastures in Amazonia, found a range of 0.8–1.5 kgC m^{-2} of fine fuels, which included both live and dead leaf material as well as fine woody debris. Again, this corresponds well with our results (Fig. 3.5a,c), although we do not simulate fine woody debris. Kauffman and Cummings (1998) also found 1.3–5.2 kgC m^{-2} of large downed trunks remaining from the initial clearance of forest for pasture; the simulation produces levels of woody litter in that range for pastures in the Atlantic Forest region of Brazil and in southern China (Fig. 3.5d).

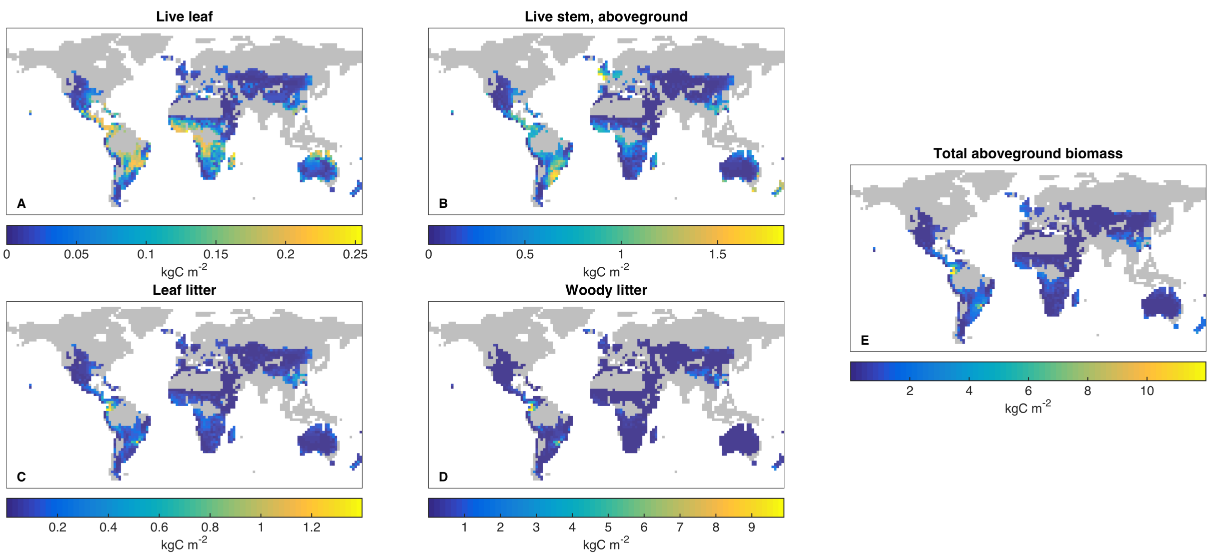

Figure 3.5: Mean aboveground carbon pools on pasture over 2001–2009. Gridcells composed of <20% pasture are shown in gray.

LM3 does seem to have overestimated pasture biomass in tropical savanna regions, however. Savadogo et al. (2007) found a mean of 0.045 kgC m^{-2} in the tree and bush savanna of Burkina Faso, where LM3 using `PASTASPAST` simulates live biomass pools (leaf + stem) of up to about 0.5 kgC m^{-2} (Fig. 3.5a, b). Savadogo et al. (2007) also found a mean of 0.07 kgC m^{-2} of dead material there, whereas our model simulated values of around 0.2–0.3 kgC m^{-2} (Fig. 3.5c, d). Ottmar et al. (2001) found that land in the Cerrado with a significant herbaceous layer (*campo limpo*, *campo sujo*, and *cerrado ralo*) generally tended to have less than 1 kgC m^{-2} of aboveground live and dead biomass; our model simulated about 1–1.5 kgC m^{-2} (Fig. 3.5e). It is not clear whether the sites examined by Ottmar et al. (2001) were actively grazed; if not, pastures there would be expected to have even less biomass, in which case LM3's overestimate would be more pronounced.

A widespread overestimation of biomass in tropical savannas would at least partially explain the tendency toward overestimated pasture fire carbon emissions there (Fig. 2.9d, h). Because most of the world's pasture fire occurs in this biome (Fig. 2.8), it would also explain the 32% overestimate of mean annual global pasture fire carbon emissions (Table 2.4). Excess simulated plant matter in tropical savannas could result from any or all of several factors. It is possible, for example, that grazing intensity there is unrealistically low. Although the global amount of grazed vegetation seems to have been simulated well (as discussed above), much variation likely exists among regions in how intensely land is grazed. This is not captured by the assumption in our model of a 4% daily grazing rate. Combustion completeness values being too low would also lead to too-high estimates of aboveground biomass, but the possible effect of this on estimated emissions is unclear. Increasing combustion completeness would increase fire emissions in the short term, but as any individual pasture tile grew older and approached equilibrium biomass, fire emissions might be no different. That is, decreased biomass with increased combustion completeness might not change emissions density.

Lastly, the fact that FINAL does not explicitly simulate fire associated with land clearance likely contributes to its overestimation of cropland and pasture fire emissions density. In the version of LM3 used here, however, burning is not assumed to occur during land use transitions. Instead, biomass can be either harvested or wasted. Harvested wood biomass goes to one of three long-lived virtual emissions pools, while wasted biomass is transferred to litter. But in reality, wood remaining after harvest (also known as slash) is often burned, especially in the high-biomass moist tropical forest biome. The emissions involved are significant: Tropical deforestation burns were estimated by van der Werf et al. (2010) to contribute up to 15% of global annual fire CO_2-C emissions on average. Instead of breaking this out into a separate flux, LM3 and FINAL are conflating land clearance fire emissions with the emissions from subsequent burning of the cleared land for agricultural management. This is unfortunately not a mere accounting quirk; the use of one or two burns to get rid of most of the remaining slash wood means that fire emissions spike soon after land clearance, whereas LM3 and FINAL simulate a gradual decrease over time. However, the frontier regions of moist tropical forests do not exhibit as much error in cropland and pasture fire carbon emissions as is seen in tropical savannas (Fig. 2.9e,f), and so the relative importance of this land clearance behavior to carbon fluxes at a global scale is limited.

3.5.2 Impact of pasture management fire on terrestrial ecosystems

Pasture burning practices may not significantly affect terrestrial carbon cycling at the global scale (Fig. 3.2), but they can be quite important for certain regions (Figs. 3.3–3.4). Parts of Eurasia saw the mean aboveground biomass of *all* vegetation – not just pasture – reduced to nearly zero with the implementation of realistic pasture burning (Fig. 3.4e). Biomass reductions nearly as extreme are apparent throughout both temperate and tropical grassland and savanna regions. Notably, much of the difference is the result of lower stem biomass (Fig. 3.4b). This suggests that the correct simulation of pasture fire frequency could be an important factor in the successful modeling of the savanna-forest boundary in large parts of the world. Indeed, fire is one factor that has been highlighted by some as critical in this dynamic, as demonstrated by large-scale empirical analyses (Staver et al., 2011b), theoretical modeling (Staver et al., 2011a; Staver and Levin, 2012), and experiments using dynamic global vegetation models (Bond et al., 2005; Lasslop et al., 2015a; Lehsten et al., 2016). Although other empirical analyses at continental to global scales have suggested that climate and/or soil quality might be more important to (or at least more directly predictive of) woody biomass than fire (Zeng et al., 2014; Lehmann et al., 2014), fire does appear to be a critical mechanism without which many grasslands and savannas – especially in less arid regions – become much woodier (O'Connor et al., 2014).

FINAL also shows that pasture burning practices significantly shift the seasonal timing of fire in many parts of the world (Fig. 3.1). The within-year timing of burns can be important for plant community composition, especially with regard to invasive vs. native species (Scmisseur and Miller, 1985; Mbow et al., 2000; Keeley, 2006). It can also have important ramifications for the temporal pattern of aerosol concentrations in the atmosphere (Marlier et al., 2014). Although not considered here, these short-lived chemical species can affect atmospheric chemistry and radiative balance (Penner et al., 1991; Ramanathan and Carmichael, 2008; Ward et al., 2012; Bond et al., 2013), change physical properties of land and the atmosphere, and harm human health (Rittmaster et al., 2006; Johnston et al., 2012; Marlier et al., 2012; Reisen et al., 2015). Understanding the amount and exact timing of aerosol emissions is important for meteorological forecasting and preparation for public health crises, so the improvement in timing shown here could have far-reaching consequences.

Conclusion

My research throughout my PhD candidacy has been focused on understanding how people manage agricultural land through fire, and how this management can affect terrestrial carbon cycling. Although cropland burning can be an important source of greenhouse gases and other pollutants – especially for certain parts of the world – it is a system whose ecology is completely controlled by humans. For that reason I have been especially interested in the biogeochemical and ecological consequences of management fire for pasture and rangeland. Truly exploring the regional variation in net global magnitude of these consequences required the use of a global fire model in conjunction with a dynamic global vegetation model.

Unfortunately, at that point there were no observations of pasture fire at the large spatial scales required to calibrate and test a global fire modeled. My co-authors and I thus set out to generate, for the first time, estimates of the regionally varying patterns of fire frequency and timing associated with pasture. We found pasture burning to make up a large portion of the fire that happens every year around the world, and that this sometimes occurs at a different time of year than burning on non-agricultural land. In the process, we developed a technique that allows us to quantify the fire-suppressing effect that people can have through landscape fragmentation. This research is included as Chapter 1, which was published last year in *Biogeosciences* (Rabin et al., 2015).

These data allowed me to then begin developing the first global fire model that would explicitly simulate agricultural burning and its effects on vegetation and biogeochemistry. Along the way, I designed a new model for non-agricultural burning that for the first time explicitly simulates the heterogeneity in vegetation structure across a landscape caused by fires occurring in different places at different times. As described in Chapter 2, this new model – which I have dubbed the Fire Including Natural & Agricultural Lands model, or FINAL – successfully captures the general global patterns of fire frequency and timing.

By parameterizing the non-agricultural fire model in FINAL against the novel estimates we had generated (Ch. 1) of burning associated with non-agricultural lands, I removed the contaminating effect that pasture burning had had on our understanding of the processes controlling fire in other ecosystems. In Chapter 3, I tested the effects of pasture management burning by experimentally applying this "purified" non-agricultural fire model to pasture instead of forcing pasture to burn according to their actual observed pattern. The results show that human control of pasture burning does indeed affect carbon pools and fluxes of carbon between the land and atmosphere, especially for certain parts of the world.

However, FINAL is by no means final – there is more work to be done.

In this first version, FINAL simulates cropland and pasture burning using a climatological forcing derived from the observational data my co-authors and I generated previously (Ch. 1). This is less than ideal for a number of reasons. First, humans are not such precise managers that they burn the exact same fraction of their pasture every single year. As currently implemented, the agricultural burning in FINAL does not allow any interannual variability, despite evidence of that in the observations (Ch. 1). More importantly, however, the observational data only cover the period 2001–2009. As climatic conditions, socioeconomic patterns, and land use types and intensities change into the future, we can expect that the timing and frequency with which land managers burn their land (or allow it to burn) will change as well. Historical patterns of agricultural management fire were also undoubtedly different. A substantial but critical challenge for future work in this area is to understand what actually drives the agricultural burning that we see, so that a prognostic version of FINAL – one that can account for shifting seasons, improvements in technology, changing crop choices, etc. – can be used to simulate past and future patterns of fire.

Finally, human-controlled burning of pasture and rangeland has more ecological consequences that simple measures of carbon pools and fluxes used here can describe. Future work should couple FINAL with a dynamic vegetation model that incorporates the demographic and competitive processes that are so important for vegetation structure and species distribution. This would allow a more informed and nuanced discussion of the ecological and biogeographical consequences of agricultural fire management for ecosystems around the world.

Bibliography

A. Aldersley, S. J. Murray, and S. E. Cornell. Global and regional analysis of climate and human drivers of wildfire. *Science of the Total Environment*, 409(18):3472–3481, Aug. 2011. doi: 10.1016/j.scitotenv.2011.05.032. URL `http://dx.doi.org/10.1016/j.scitotenv.2011.05.032`.

N. Andela and G. R. van der Werf. Recent trends in African fires driven by cropland expansion and El Niño to La Niña transition. *Nature Climate Change*, 4(9):791–795, Aug. 2014. doi: 10.1038/nclimate2313. URL `http://www.nature.com/doifinder/10.1038/nclimate2313`.

S. Archibald, D. P. Roy, B. W. van Wilgen, and R. Scholes. What limits fire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology*, 15(3):613–630, 2009. doi: 10.1111/j.1365-2486.2008.01754.x. URL `http://doi.wiley.com/10.1111/j.1365-2486.2008.01754.x`.

S. Archibald, A. C. Staver, and S. A. Levin. Evolution of human-driven fire regimes in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 109(3): 847–852, 2012. URL `http://www.pnas.org/content/109/3/847.short`.

S. Archibald, C. E. R. Lehmann, J. L. Gomez-Dans, and R. A. Bradstock. Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences*, 110 (16):6442–6447, Apr. 2013. doi: 10.1073/pnas.1211466110. URL `http://www.pnas.org/cgi/doi/10.1073/pnas.1211466110`.

V. K. Arora and G. Boer. Fire as an interactive component of dynamic vegetation models. *Journal of Geophysical Research*, 110, 2005.

J. K. Balch, D. C. Nepstad, P. M. Brando, L. M. Curran, O. Portela, O. de Carvalho, and P. Lefebvre. Negative fire feedback in a transitional forest of southeastern Amazonia. *Global Change Biology*, 14(10):2276–2287, Oct. 2008. doi: 10.1111/j.1365-2486.2008.01655.x.

I. Bistinas, D. Oom, A. C. L. Sá, S. P. Harrison, I. C. Prentice, and J. M. C. Pereira. Relationships between human population density and burned area at continental and global scales. *PLoS One*, 8(12):e81188, Dec. 2013. doi: 10.1371/journal.pone.0081188.t003. URL `http://dx.plos.org/10.1371/journal.pone.0081188.t003`.

I. Bistinas, S. P. Harrison, I. C. Prentice, and J. M. C. Pereira. Causal relationships versus emergent patterns in the global controls of fire frequency. *Biogeosciences*, 11(18):5087–5101, 2014. doi: 10.5194/bg-11-5087-2014. URL http://www.biogeosciences.net/11/5087/2014/.

T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. T. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research-Atmospheres*, 118(11):5380–5552, June 2013. doi: 10.1002/jgrd.50171. URL http://doi.wiley.com/10.1002/jgrd.50171.

W. J. Bond and J. E. Kelley. Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution*, 20(7):387–394, July 2005. doi: 10.1016/j.tree.2005.04.025. URL http://linkinghub.elsevier.com/retrieve/pii/S0169534705001321.

W. J. Bond, F. I. Woodward, and G. F. Midgley. The global distribution of ecosystems in a world without fire. *New Phytologist*, 165(2):525–538, 2005. doi: 10.1111/j.1469-8137.2004.01252.x. URL http://doi.wiley.com/10.1111/j.1469-8137.2004.01252.x.

D. M. J. S. Bowman, J. K. Balch, P. Artaxo, W. J. Bond, M. A. Cochrane, C. M. D'antonio, R. DeFries, F. H. Johnston, J. E. Keeley, and M. A. Krawchuk. The human dimension of fire regimes on Earth. *Journal of Biogeography*, 38(12):2223–2236, 2011. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2699.2011.02595.x/full.

B. P. Bryant and A. L. Westerling. Scenarios for future wildfire risk in California: links between changing demography, land use, climate, and wildfire. *Environmetrics*, 25(6):454–471, July 2014. doi: 10.1002/env.2280. URL http://doi.wiley.com/10.1002/env.2280.

D. J. Cecil, D. E. Buechler, and R. J. Blakeslee. Gridded lightning climatology from TRMM-LIS and OTD: Dataset description. *Atmospheric Research*, 135-136(C):404–414, 2012. doi: 10.1016/j.atmosres.2012.06.028. URL http://dx.doi.org/10.1016/j.atmosres.2012.06.028.

P. Cheney and A. Sullivan. *Grassfires: Fuel, Weather and Fire Behaviour*. CSIRO Publishing, 2nd edition, 2009. doi: 10.1111/j.1442-9993.2009.02037.x. URL http://doi.wiley.com/10.1111/j.1442-9993.2009.02037.x.

P. Ciais, C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Quéré, R. Myneni, S. Piao, and P. Thornton. Carbon and Other Biogeochemical Cycles. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA, 2013.

M. A. Cochrane and M. D. Schulze. Forest fires in the Brazilian Amazon. *Conservation Biology*, 12(5):948–950, 1998. URL http://www.jstor.org/stable/10.2307/2387566.

P. M. Cox, P. P. Harris, C. Huntingford, R. A. Betts, M. Collins, C. D. Jones, T. E. Jupp, J. A. Marengo, and C. A. Nobre. Increasing risk of Amazonian drought due to decreasing aerosol pollution. *Nature*, 453(7192):212–215, May 2008. doi: 10.1038/nature06960. URL http://www.nature.com/doifinder/10.1038/nature06960.

P. J. Crutzen and M. O. Andreae. Biomass Burning in the Tropics: Impact on Atmospheric Chemistry and Biogeochemical Cycles. *Science*, 250(4988):1669–1678, Dec. 1990. doi: 10.2307/2878539?ref=no-x-route:2460ed5c39e79cae59cdd5513c69f3bd.

S. H. Doerr and C. Santín. Wildfire: A burning issue for insurers? Technical report, Sept. 2013.

J. P. Dunne, J. G. John, E. Shevliakova, R. Stouffer, J. P. Krasting, S. Malyshev, P. C. D. Milly, L. Sentman, A. J. Adcroft, W. Cooke, K. Dunne, S. M. Griffies, R. W. Hallberg, M. J. Harrison, H. Levy, A. T. Wittenberg, P. J. Phillips, and N. Zadeh. GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics. *Journal of Climate*, 26:2247–2267, 2013. doi: 10.1175/JCLI.

FAO. Concepts & Definitions, Feb. 2015. URL http://faostat.fao.org/site/375/default.aspx.

M. D. Flannigan, M. Krawchuk, and W. J. de Groot. Implications of changing climate for global wildland fire. *International Journal of Wildland Fire*, 18:483–507, 2009. URL http://www.publish.csiro.au/?paper=WF08187.

M. Friedl, D. McIver, J. Hodges, X. Zhang, D. Muchoney, A. Strahler, C. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83:287–302, 2002. URL http://linkinghub.elsevier.com/retrieve/pii/S0034425702000883.

M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1):168–182, Jan. 2010. doi: 10.1016/j.rse.2009.08.016. URL http://dx.doi.org/10.1016/j.rse.2009.08.016.

L. Giglio. *MODIS Collection 5 Active Fire Product User's Guide Version 2.5*, Mar. 2013. URL https://earthdata.nasa.gov/sites/default/files/field/document/MODIS_Fire_Users_Guide_2.5.pdf.

L. Giglio, G. R. van der Werf, J. T. Randerson, G. Collatz, and P. Kasibhatla. Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics*, 6:957–974, 2006. URL http://hal-insu.archives-ouvertes.fr/hal-00295869/.

L. Giglio, T. Loboda, D. P. Roy, B. Quayle, and C. O. Justice. An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, 113(2):408–420, Feb. 2009. doi: 10.1016/j.rse.2008.10.006. URL http://dx.doi.org/10.1016/j.rse.2008.10.006.

L. Giglio, J. T. Randerson, G. R. van der Werf, P. Kasibhatla, G. Collatz, D. C. Morton, and R. DeFries. Assessing variability and long-term trends in burned area by merging multiple satellite fire products. *Biogeosciences*, 7(3):1171–1186, 2010.

N. P. Gillett, A. J. Weaver, F. W. Zwiers, and M. D. Flannigan. Detecting the effect of climate change on Canadian forest fires. *Geophysical Research Letters*, 31(18):L18211, 2004. doi: 10.1029/2004GL020876. URL http://doi.wiley.com/10.1029/2004GL020876.

H. Haberl, K. H. Erb, and F. Krausmann. Human appropriation of net primary production (HANPP), 2007. URL http://isecoeco.org/pdf/2007_march_hanpp.pdf.

M. B. Hahn, R. E. Gangnon, C. Barcellos, G. P. Asner, and J. A. Patz. Influence of Deforestation, Logging, and Fire on Malaria in the Brazilian Amazon. *PLoS One*, 9(1):e85725, Jan. 2014. doi: 10.1371/journal.pone.0085725.t003. URL http://dx.plos.org/10.1371/journal.pone.0085725.t003.

M. M. Hanna, D. A. Steyn-Ross, and M. Steyn-Ross. Estimating Biomass for New Zealand Pasture Using Optical Remote Sensing Techniques. *Geocarto International*, 14(3):89–94, Sept. 1999. doi: 10.1080/10106049908542121. URL http://www.tandfonline.com/doi/abs/10.1080/10106049908542121.

S. Hantson, G. Lasslop, S. Kloster, and E. Chuvieco. Anthropogenic effects on global mean fire size. *International Journal of Wildland Fire*, 24(5):589–596, 2015. doi: 10.1071/WF14208. URL http://www.publish.csiro.au/?paper=WF14208.

S. Hantson, A. Arneth, S. P. Harrison, D. I. Kelley, I. C. Prentice, S. S. Rabin, S. Archibald, F. Mouillot, S. R. Arnold, P. Artaxo, D. Bachelet, P. Ciais, M. Forrest, P. Friedlingstein, T. Hickler, J. O. Kaplan, S. Kloster, W. Knorr, G. Lasslop, F. Li, S. Mangeon, J. R. Melton, A. Meyn, S. Sitch, A. Spessa, G. R. van der Werf, A. Voulgarakis, and C. Yue. The status and challenge of global fire modelling. *Biogeosciences Discussions*, 0:1–30, Jan. 2016. doi: 10.5194/bg-2016-17. URL http://www.biogeosciences-discuss.net/bg-2016-17/.

G. C. Hurtt, L. P. Chini, S. Frolking, R. A. Betts, J. Feddema, G. Fischer, J. P. Fisk, K. Hibbard, R. A. Houghton, A. Janetos, C. D. Jones, G. Kindermann, T. Kinoshita, K. Klein Goldewijk, K. Riahi, E. Shevliakova, S. Smith, E. Stehfest, A. Thomson, P. Thornton, D. P. van Vuuren, and Y. P. Wang. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, 109(1-2):117–161, Aug. 2011. doi: 10.1007/s10584-011-0153-2. URL http://www.springerlink.com/index/10.1007/s10584-011-0153-2.

F. H. Johnston, S. B. Henderson, Y. Chen, J. T. Randerson, M. E. Marlier, R. S. DeFries, P. L. Kinney, D. M. J. S. Bowman, and M. Brauer. Estimated global mortality attributable to smoke from landscape fires. *Environmental Health Perspectives*, 120(5):695–701, 2012. doi: 10.1289/ehp.l. URL http://escholarship.org/uc/item/9s67c46r.pdf.

A. Kagunyu, S. Wandibba, and J. G. Wanjohi. The use of indigenous climate forecasting methods by the pastoralists of Northern Kenya. *Pastoralism*, pages 1–6, Mar. 2016. doi: 10.1186/s13570-016-0054-0. URL http://dx.doi.org/10.1186/s13570-016-0054-0.

J. B. Kauffman and D. Cummings. Fire in the Brazilian Amazon 2. Biomass, nutrient pools and losses in cattle pastures. *Oecologia*, 113:415–427, 1998. URL http://www.springerlink.com/index/HD90X79RVG9Y2P9C.pdf.

J. E. Keeley. Fire Management Impacts on Invasive Plants in the Western United States. *Conservation Biology*, 20(2):375–384, Apr. 2006. doi: 10.1111/j.1523-1739.2006.00339.x. URL http://doi.wiley.com/10.1111/j.1523-1739.2006.00339.x.

D. I. Kelley, I. C. Prentice, S. P. Harrison, H. Wang, M. Simard, J. B. Fisher, and K. O. Willis. A comprehensive benchmarking system for evaluating global vegetation models. *Biogeosciences*, 10(5):3313–3340, 2013. doi: 10.5194/bg-10-3313-2013. URL http://www.biogeosciences.net/10/3313/2013/.

K. Klein Goldewijk, G. van Drecht, and A. F. Bouwman. Mapping contemporary global cropland and grassland distributions on a 5×5 minute resolution. *Journal of Land Use Science*, 2(3): 167–190, Oct. 2007. doi: 10.1080/17474230701622940. URL http://www.tandfonline.com/doi/abs/10.1080/17474230701622940.

K. Klein Goldewijk, A. Beusen, G. Van Drecht, and M. De Vos. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, 20(1):73–86, Sept. 2010. doi: 10.1111/j.1466-8238.2010.00587.x. URL http://doi.wiley.com/10.1111/j.1466-8238.2010.00587.x.

W. Knorr, T. Kaminski, A. Arneth, and U. Weber. Impact of human population density on fire frequency at the global scale. *Biogeosciences*, 11(4):1085–1102, 2014. doi: 10.5194/bg-11-1085-2014. URL http://www.biogeosciences.net/11/1085/2014/.

S. Korontzi, J. McCarty, T. Loboda, S. Kumar, and C. Justice. Global distribution of agricultural fires in croplands from 3 years of Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Global Biogeochemical Cycles*, 20(GB2021), 2006. doi: 10.1029/2005GB002529. URL http://www.agu.org/pubs/crossref/2006/2005GB002529.shtml.

G. N. Korovin. Analysis of the distribution of forest fires in Russia. In J. G. Goldammer and V. V. Furyaev, editors, *Fire in Ecosystems of Boreal Eurasia*, pages 112–128. Dordrecht, The Netherlands, 1996. ISBN 0-7923-4137-6.

A. Krause, S. Kloster, S. Wilkenskjeld, and H. Paeth. The sensitivity of global wildfires to simulated past, present, and future lightning frequency. *Journal of Geophysical Research: Biogeosciences*, 119:312–322, Apr. 2014. doi: 10.1002/(ISSN)2169-8961.

F. Krausmann, K.-H. Erb, S. Gingrich, C. Lauk, and H. Haberl. Global patterns of socioeconomic biomass flows in the year 2000: A comprehensive assessment of supply, consumption and constraints. *Ecological Economics*, 65(3):471–487, Apr. 2008. doi: 10.1016/j.ecolecon.2007.07.012. URL http://linkinghub.elsevier.com/retrieve/pii/S0921800907004053.

C. A. Kull and P. Laris. Fire ecology and fire politics in Mali and Madagascar. In M. A. Cochrane, editor, *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics*, pages 171–226. 2009.

P. Laris. Burning the seasonal mosaic: Preventative burning strategies in the wooded savanna of southern Mali. *Human Ecology*, 30(2):155–186, June 2002. URL `http://www.springerlink.com/index/PQ98FDF8JJB892EM.pdf`.

P. Laris. Humanizing Savanna Biogeography: Linking Human Practices with Ecological Patterns in a Frequently Burned Savanna of Southern Mali. *Annals of the Association of American Geographers*, 101(5):1067–1088, Sept. 2011. doi: 10.1080/00045608.2011.560063. URL `http://www.tandfonline.com/doi/abs/10.1080/00045608.2011.560063`.

G. Lasslop, V. Brovkin, S. Kloster, and C. H. Reick. Tree cover bistability in the MPI Earth system model due to fire-vegetation feedback. In *EGU General Assembly ...*, 2015a. URL `http://adsabs.harvard.edu/abs/2015EGUGA..17.9831L`.

G. Lasslop, S. Hantson, and S. Kloster. Influence of wind speed on the global variability of burned fraction: a global fire model's perspective. *International Journal of Wildland Fire*, 22:959–969, 2015b. doi: 10.1071/WF15052. URL `http://www.publish.csiro.au/?paper=WF15052`.

W. Laurance and G. Williamson. Positive feedbacks among forest fragmentation, drought, and climate change in the Amazon. *Conservation Biology*, 15(6):1529–1535, 2001. URL `http://onlinelibrary.wiley.com/doi/10.1046/j.1523-1739.2001.01093.x/full`.

Y. Le Page, D. Oom, J. M. N. Silva, P. Jönsson, and J. M. C. Pereira. Seasonality of vegetation fires as modified by human action: Observing the deviation from eco-climatic fire regimes. *Global Ecology and Biogeography*, 19:575–588, 2010. doi: 10.1111/j.1466-8238.2010.00525.x. URL `http://doi.wiley.com/10.1111/j.1466-8238.2010.00525.x`.

Y. Le Page, D. C. Morton, B. Bond-Lamberty, J. M. C. Pereira, and G. Hurtt. HESFIRE: a global fire model to explore the role of anthropogenic and weather drivers. *Biogeosciences*, 12 (3):887–903, 2015. doi: 10.5194/bg-12-887-2015. URL `http://www.biogeosciences.net/12/887/2015/`.

B. Leff, N. Ramankutty, and J. A. Foley. Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, 18(1):–, 2004. doi: 10.1029/2003GB002108. URL `ftp://ftp.ciat.cgiar.org/Agroecosystems/incoming/Elizabeth/Leff_etal_2004.pdf`.

C. E. R. Lehmann, T. M. Anderson, M. Sankaran, S. I. Higgins, S. Archibald, W. A. Hoffmann, N. P. Hanan, R. J. Williams, R. J. Fensham, J. Felfili, L. B. Hutley, J. Ratnam, J. San Jose, R. Montes, D. Franklin, J. Russell-Smith, C. M. Ryan, G. Durigan, P. Hiernaux, R. Haidar, D. M. J. S. Bowman, and W. J. Bond. Savanna Vegetation-Fire-Climate Relationships Differ Among Continents. *Science*, 343(6170):548–552, Jan. 2014. doi: 10.1126/science.1247355. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1247355`.

V. Lehsten, P. Harmand, I. Palumbo, and A. Arneth. Modelling burned area in Africa. *Biogeosciences*, 7(10):3199–3214, 2010. doi: 10.5194/bg-7-3199-2010. URL `http://www.`

biogeosciences.net/7/3199/2010/.

V. Lehsten, A. Arneth, A. Spessa, K. Thonicke, and A. Moustakas. The effect of fire on tree–grass coexistence in savannas: a simulation study. *International Journal of Wildland Fire*, 25 (2):137–10, 2016. doi: 10.1071/WF14205. URL http://www.publish.csiro.au/?paper=WF14205.

J. Lenihan, C. Daly, D. Bachelet, and R. Neilson. Simulating broad-scale fire severity in a dynamic global vegetation model. *Northwest Science*, 72(4):91–101, 1998. URL http://www.sysecol2.ethz.ch/Refs/EntClim/L/Le146.pdf.

K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2:164–168, 1944.

F. Li, X. D. Zeng, and S. Levis. A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model. *Biogeosciences*, 9(7):2761–2780, 2012. doi: 10.5194/bg-9-2761-2012. URL http://www.biogeosciences.net/9/2761/2012/.

F. Li, S. Levis, and D. S. Ward. Quantifying the role of fire in the Earth system—Part 1: Improved global fire modeling in the Community Earth System Model (CESM1). *Biogeosciences*, 10(4):2293–2314, 2013. doi: 10.5194/bg-10-2293-2013. URL http://www.biogeosciences.net/10/2293/2013/.

H.-W. Lin, Y. Jin, L. Giglio, J. A. Foley, and J. T. Randerson. Evaluating greenhouse gas emissions inventories for agricultural burning using satellite observations of active fires. *Ecological Applications*, 22(4):1345–1364, 2012. URL http://www.esajournals.org/doi/full/10.1890/10-2362.1.

B. I. Magi, S. Rabin, E. Shevliakova, and S. W. Pacala. Separating agricultural and non-agricultural fire seasonality at regional scales. *Biogeosciences*, 9(8):3003–3012, Aug. 2012. doi: 10.5194/bg-9-3003-2012. URL http://www.biogeosciences.net/9/3003/2012/.

M. E. Marlier, R. S. DeFries, A. Voulgarakis, P. L. Kinney, J. T. Randerson, D. T. Shindell, Y. Chen, and G. Faluvegi. El Niño and health risks from landscape fire emissions in southeast Asia. *Nature Climate Change*, 3(2):131–136, 2012. doi: 10.1038/nclimate1658. URL http://www.nature.com/nclimate/journal/vaop/ncurrent/full/nclimate1658.html?WT.mc_id=FBK_NatureClimate.

M. E. Marlier, A. Voulgarakis, D. T. Shindell, G. Faluvegi, C. L. Henry, and J. T. Randerson. Atmospheric Environment. *Atmospheric Environment*, 89(C):158–168, June 2014. doi: 10.1016/j.atmosenv.2014.02.039. URL http://dx.doi.org/10.1016/j.atmosenv.2014.02.039.

J. Marlon, P. J. Bartlein, C. Carcaillet, D. G. Gavin, S. P. Harrison, P. E. Higuera, F. Joos, M. J. Power, and I. C. Prentice. Climate and human influences on global biomass burning over the past two millennia. *Nature Geoscience*, 1:697–702, 2008. URL http://www.nature.com/ngeo/journal/v1/n10/abs/ngeo313.html.

D. W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963. doi: 10.1137/0111030. URL http://epubs.siam.org/doi/abs/10.1137/0111030.

C. Mbow, T. T. Nielsen, and K. Rasmussen. Savanna Fires in East-Central Senegal: Distribution Patterns, Resource Management and Perceptions. *Human Ecology*, 28(4):561–583, Dec. 2000. doi: 10.2307/4603372?ref=no-x-route:7766736a94bb613bc42c1a58dc02439d.

M. Meinshausen, S. J. Smith, K. Calvin, J. S. Daniel, M. L. T. Kainuma, J.-F. Lamarque, K. Matsumoto, S. A. Montzka, S. C. B. Raper, K. Riahi, A. Thomson, G. J. M. Velders, and D. P. van Vuuren. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, 109(1-2):213–241, Aug. 2011. doi: 10.1007/s10584-011-0156-z. URL http://www.springerlink.com/index/10.1007/s10584-011-0156-z.

J. R. Melton and V. K. Arora. Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0. *Geoscientific Model Development*, 9:323–361, 2015. doi: 10.5194/gmd-9-323-2016.

G. Myhre, D. T. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, and H. Zhang. Anthropogenic and natural radiative forcing. pages 659–740. Cambridge, United Kingdom and New York, NY, USA, Dec. 2013. URL http://www.climatechange2013.org/images/report/WG1AR5_Chapter08_FINAL.pdf.

H. B. Nielsen. Damping parameter in Marquardt's method. Technical Report IMM-REP-1999-05, Lyngby, Denmark, 1999.

H. B. Nielsen. marquardt.m. page 2, 2001. URL http://www2.imm.dtu.dk/projects/hbn_software/marquardt.m.

T. G. O'Connor, J. R. Puttick, and M. T. HOFFMAN. Bush encroachment in southern Africa: changes and causes. *African Journal of Range & Forage Science*, 31(2):67–88, July 2014. doi: 10.2989/10220119.2014.939996. URL http://www.tandfonline.com/doi/abs/10.2989/10220119.2014.939996.

D. M. Olson, E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. Powell, E. C. Underwood, J. A. D'amico, I. Itoua, H. E. Strand, and J. C. Morrison. Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, 51(11):933–938, 2001. doi: 10.1641/0006-3568(2001)051%5B0933:TEOTWA%5D2.0.CO;2.

R. D. Ottmar, R. E. Vihnanek, H. S. Miranda, M. N. Sato, and S. M. A. Andrade. Stereo Photo Series for Quantifying Cerrado Fuels in Central Brazil—Volume I. Technical Report Gen. Tech. Rep. PNW-GTR-519, 2001.

O. Pechony and D. T. Shindell. Fire parameterization on a global scale. *Journal of Geophysical Research*, 114(D16115), Aug. 2009. doi: 10.1029/2009JD011927.

J. E. Penner, S. J. Ghan, and J. J. Walton. The role of biomass burning in the budget and cycle of carbonaceous soot aerosols and their climate impact. In J. S. Levine, editor, *Chapman Conference on Global Biomass Burning: Atmospheric, Climatic, and Biospheric Implications*, pages 387–393, Williamsburg, VA, USA, 1991.

M. Pfeiffer, A. Spessa, and J. O. Kaplan. A model for global biomass burning in preindustrial time: LPJ-LMfire (v1.0). *Geoscientific Model Development*, 6(3):643–685, 2013. doi: 10.5194/gmd-6-643-2013. URL `http://www.geosci-model-dev.net/6/643/2013/`.

S. Prentice and D. Mackerras. The ratio of cloud to cloud-ground lightning flashes in thunderstorms. *Journal of Applied Meteorology*, 16:1–6, 1977.

S. J. Pyne, P. L. Andrews, and R. D. Laven. Fire and culture. In *Introduction to Wildland Fire*, pages 213–307. John Wiley & Sons, 1996a.

S. J. Pyne, P. L. Andrews, and R. D. Laven. Fire ecology. In *Introduction to Wildland Fire*, pages 171–212. John Wiley & Sons, New York, 1996b.

S. S. Rabin, B. I. Magi, E. Shevliakova, and S. W. Pacala. Quantifying regional, time-varying effects of cropland and pasture on vegetation fire. *Biogeosciences*, 12(22):6591–6604, 2015. doi: 10.5194/bg-12-6591-2015. URL `http://www.biogeosciences.net/12/6591/2015/`.

V. Ramanathan and G. Carmichael. Global and regional climate changes due to black carbon. *Nature Geoscience*, 1(4):221–227, 2008. URL `http://www.nature.com/ngeo/journal/vaop/ncurrent/full/ngeo156.html`.

N. Ramankutty, A. T. Evan, C. Monfreda, and J. A. Foley. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, 22(1), Jan. 2008. doi: 10.1029/2007GB002952. URL `http://www.agu.org/pubs/crossref/2008/2007GB002952.shtml`.

J. T. Randerson, Y. Chen, G. R. van der Werf, B. M. Rogers, and D. C. Morton. Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research*, 117 (G4):G04012, Dec. 2012. doi: 10.1029/2012JG002128. URL `http://www.agu.org/pubs/crossref/2012/2012JG002128.shtml`.

F. Reisen, S. M. Duran, M. Flannigan, C. Elliott, and K. Rideout. Wildfire smoke and public health risk. *International Journal of Wildland Fire*, 24(8):1029–16, 2015. doi: 10.1071/WF15034. URL `http://www.publish.csiro.au/?paper=WF15034`.

R. Rittmaster, W. L. Adamowicz, B. Amiro, and R. T. Pelletier. Economic analysis of health effects from forest fires. *Canadian Journal of Forest Research*, 36(4):868–877, Apr. 2006. doi: 10.1139/x05-293. URL `http://www.nrcresearchpress.com/doi/abs/10.1139/x05-293`.

P. Savadogo, D. Zida, L. Sawadogo, D. Tiveau, M. Tigabu, and P. C. Odén. Fuel and fire characteristics in savanna–woodland of West Africa in relation to grazing and dominant grass type. *International Journal of Wildland Fire*, 16(5):531–539, 2007. doi: 10.1071/WF07011. URL `http://www.publish.csiro.au/?paper=WF07011`.

E. Scmisseur and R. Miller. *Prescribed Fire for Eastern Oregon Rangelands: Management Considerations*. Agricultural Experiment Station, Oregon State University, Corvallis, Corvallis, OR, Sept. 1985.

A. C. Scott, D. M. J. S. Bowman, W. J. Bond, S. J. Pyne, and M. E. Alexander. Pyrogeography—temporal and spatial patterns of fire. In *Fire on Earth: An Introduction*, pages 113–129. West Sussex, UK, Jan. 2014. ISBN 978-1-119-95357-9.

W. Seiler and P. J. Crutzen. Estimates of gross and net fluxes of carbon between the biosphere and the atmosphere from biomass burning. *Climatic Change*, 2(3):207–247, 1980. URL `http://link.springer.com/article/10.1007/BF00137988`.

J. Sheffield, G. Goteti, and E. F. Wood. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(13):3088–3111, July 2006. doi: 10.1175/JCLI3790.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JCLI3790.1`.

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

A. C. Staver and S. A. Levin. Integrating theoretical climate and fire effects on savanna and forest systems. *The American Naturalist*, 180(2):211–224, Aug. 2012. doi: 10.1086/666648. URL `http://www.jstor.org/stable/info/10.1086/666648`.

A. C. Staver, S. Archibald, and S. Levin. Tree cover in sub-Saharan Africa: rainfall and fire constrain forest and savanna as alternative stable states. 92(5):1063–1072, 2011a. URL `http://www.esajournals.org/doi/abs/10.1890/10-1684.1`.

A. C. Staver, S. Archibald, and S. A. Levin. The global extent and determinants of savanna and forest as alternative biome states. *Science*, 334(6053):230–232, Oct. 2011b. doi: 10.1126/science.1210465. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1210465`.

B. J. Stocks, J. A. Mason, J. B. Todd, E. M. Bosch, B. M. Wotton, B. D. Amiro, M. D. Flannigan, K. G. Hirsch, K. A. Logan, D. L. Martell, and W. R. Skinner. Large forest fires in Canada, 1959–1997. *Journal of Geophysical Research*, 108(D1):8149, 2003. doi: 10.1029/2001JD000484. URL `http://doi.wiley.com/10.1029/2001JD000484`.

B. N. Sulman, R. P. Phillips, A. C. Oishi, E. Shevliakova, and S. W. Pacala. Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2. *Nature Climate Change*, 4(12):1099–1102, Nov. 2014. doi: 10.1038/nclimate2436. URL `http://www.nature.com/doifinder/10.1038/nclimate2436`.

K. E. Taylor, R. J. Stouffer, and G. A. Meehl. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498, Apr. 2012. doi: 10.1175/BAMS-D-11-00094.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00094.1`.

K. Thonicke, A. Spessa, I. C. Prentice, S. P. Harrison, L. Dong, and C. Carmona-Moreno. The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: Results from a process-based model. *Biogeosciences*, 7(6):1991–2011, 2010.

M. K. Transtrum and J. P. Sethna. Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization. Jan. 2012. URL http://arxiv.org/abs/1201.5885.

C. Uhl and R. Buschbacher. A disturbing synergism between cattle ranch burning practices and selective tree harvesting in the eastern Amazon. *Biotropica*, 17(4):265–268, 1985. URL http://www.jstor.org/stable/2388588.

C. Uhl and J. Kauffman. Deforestation, fire susceptibility, and potential tree responses to fire in the eastern Amazon. 71(2):437–449, 1990. URL http://www.esajournals.org/doi/abs/10.2307/1940299.

G. R. van der Werf, J. T. Randerson, L. Giglio, G. J. Collatz, M. Mu, P. S. Kasibhatla, D. C. Morton, R. S. Defries, Y. Jin, and T. T. van Leeuwen. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics*, 10(23):11707–11735, Dec. 2010. doi: 10.5194/acp-10-11707-2010. URL http://www.atmos-chem-phys.net/10/11707/2010/.

D. P. van Vuuren, P. L. Lucas, and H. Hilderink. Downscaling drivers of global environmental change: Enabling use of global SRES scenarios at the national and grid levels. *Global Environmental Change*, 17(1):114–130, Feb. 2007. doi: 10.1016/j.gloenvcha.2006.04.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0959378006000501.

S. Venevsky, K. Thonicke, S. Sitch, and W. Cramer. Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study. *Global Change Biology*, 8(10):984–998, 2002.

D. S. Ward, S. Kloster, N. M. Mahowald, B. M. Rogers, J. T. Randerson, and P. G. Hess. The changing radiative forcing of fires: global model estimates for past, present and future. *Atmospheric Chemistry and Physics*, 12(22):10857–10886, 2012. doi: 10.5194/acp-12-10857-2012. URL http://www.atmos-chem-phys.net/12/10857/2012/.

A. L. Westerling, H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam. Warming and earlier spring increase western U.S. forest wildfire activity. *Science*, 313(5789):940–943, Aug. 2006. doi: 10.1126/science.1128834. URL http://www.sciencemag.org/content/313/5789/940.short.

S. Wirsenius. *Human Use of Land and Organic Materials: Modeling the Turnover of Biomass in the Global Food System*. PhD thesis, Department of Physical Resource Theory, Chalmers University of Technology and Göteborg University, Göteburg, Sweden, 2000. URL http://scholar.google.com/scholar?q=related:q10vTqZKo1sJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

R. Yevich and J. A. Logan. An assessment of biofuel use and burning of agricultural waste in the developing world. *Global Biogeochemical Cycles*, 17(4,1095), 2003. doi: 10.1029/2002GB001952. URL http://www.agu.org/pubs/crossref/2003/2002GB001952.shtml.

Y. Yi, J. S. Kimball, L. A. Jones, R. H. Reichle, R. Nemani, and H. A. Margolis. Recent climate and fire disturbance impacts on boreal and arctic ecosystem productivity estimated using a satellite-based terrestrial carbon flux model. *Journal of Geophysical Research: Biogeosciences*, 118(2):606–622, May 2013. doi: 10.1002/jgrg.20053. URL `http://doi.wiley.com/10.1002/jgrg.20053`.

C. Yue, P. Ciais, P. Cadule, K. Thonicke, and T. T. van Leeuwen. Modelling the role of fires in the terrestrial carbon balance by incorporating SPITFIRE into the global vegetation model ORCHIDEE – Part 2: Carbon emissions and the role of fires in the global carbon balance. *Geoscientific Model Development*, 8(5):1321–1338, 2015. doi: 10.5194/gmd-8-1321-2015. URL `http://www.geosci-model-dev.net/8/1321/2015/`.

C. Yue, P. Ciais, D. Zhu, T. Wang, S. S. Peng, and S. L. Piao. How have past fire disturbances contributed to the current carbon balance of boreal ecosystems? *Biogeosciences*, 13(3):675–690, 2016. doi: 10.5194/bg-13-675-2016. URL `http://www.biogeosciences.net/13/675/2016/`.

Z. Zeng, A. Chen, S. Piao, S. S. Rabin, and Z. Shen. Environmental determinants of tropical forest and savanna distribution: A quantitative model evaluation and its implication. *Journal of Geophysical Research: Biogeosciences*, 119:1432–1445, 2014. doi: 10.1002/(ISSN)2169-8961. URL `http://onlinelibrary.wiley.com/doi/10.1002/2014JG002627/full`.

Y. Zhang, R. Fu, H. Yu, R. E. Dickinson, R. N. Juarez, M. Chin, and H. Wang. A regional climate model study of how biomass burning aerosol impacts land-atmosphere interactions over the Amazon. *Journal of Geophysical Research*, 113(D14):D14S15, June 2008. doi: 10.1029/2007JD009449. URL `http://doi.wiley.com/10.1029/2007JD009449`.

Appendix A

Chapter 1 supplementary tables and figures

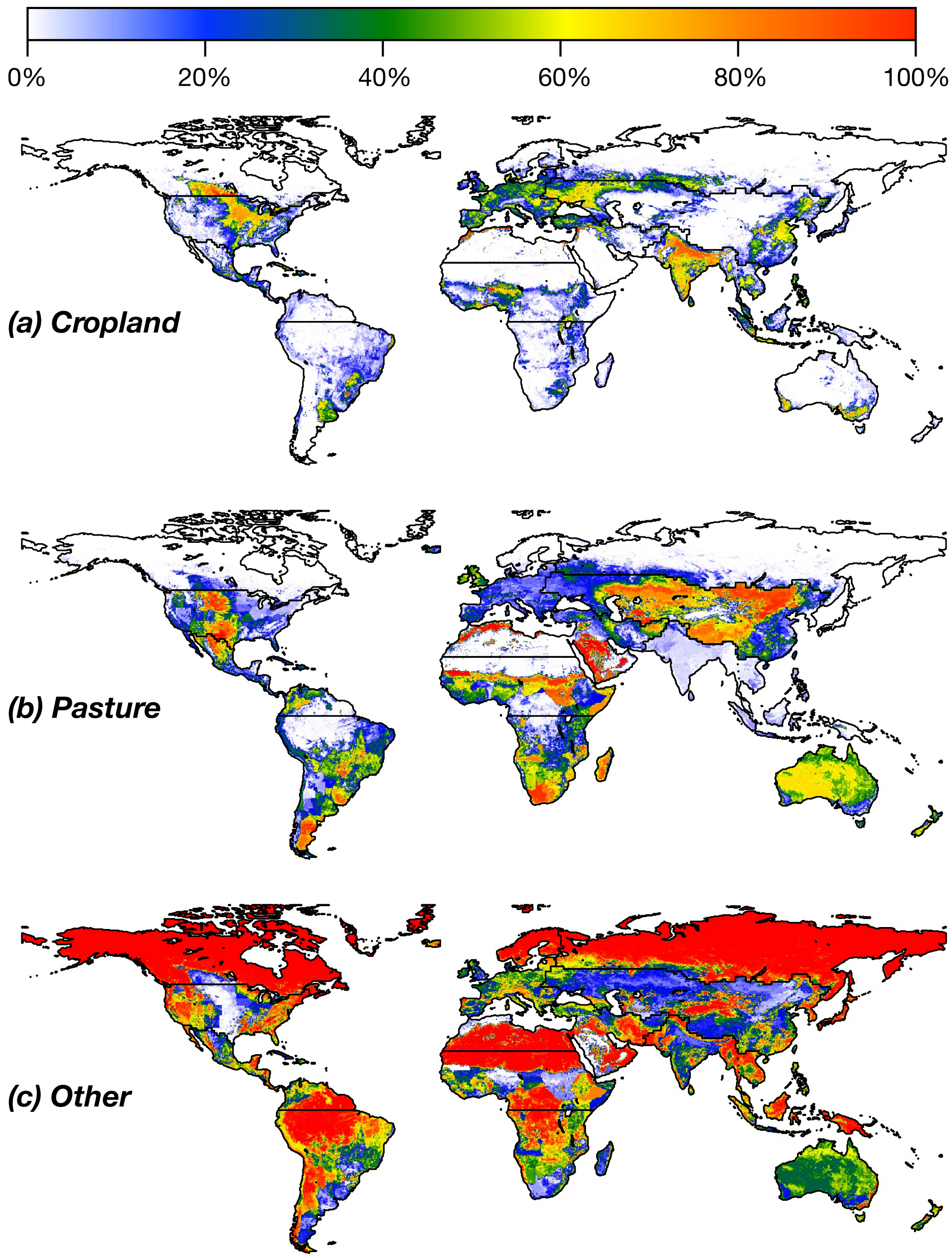

Figure A.1: Maps of mean land cover distributions from the HYDE dataset for the years 2001–2009: **(a)** cropland, **(b)** pasture, and **(c)** other land.

	Estimated				*Observed*		
	Cropland	Pasture	Other	Total	Cropland	Crop+	Total
Burned area	43.5	194	239	476	20.8	45.6	467
C emissions	194	540	1,850	2,580	N.D.	N.D.	2,480

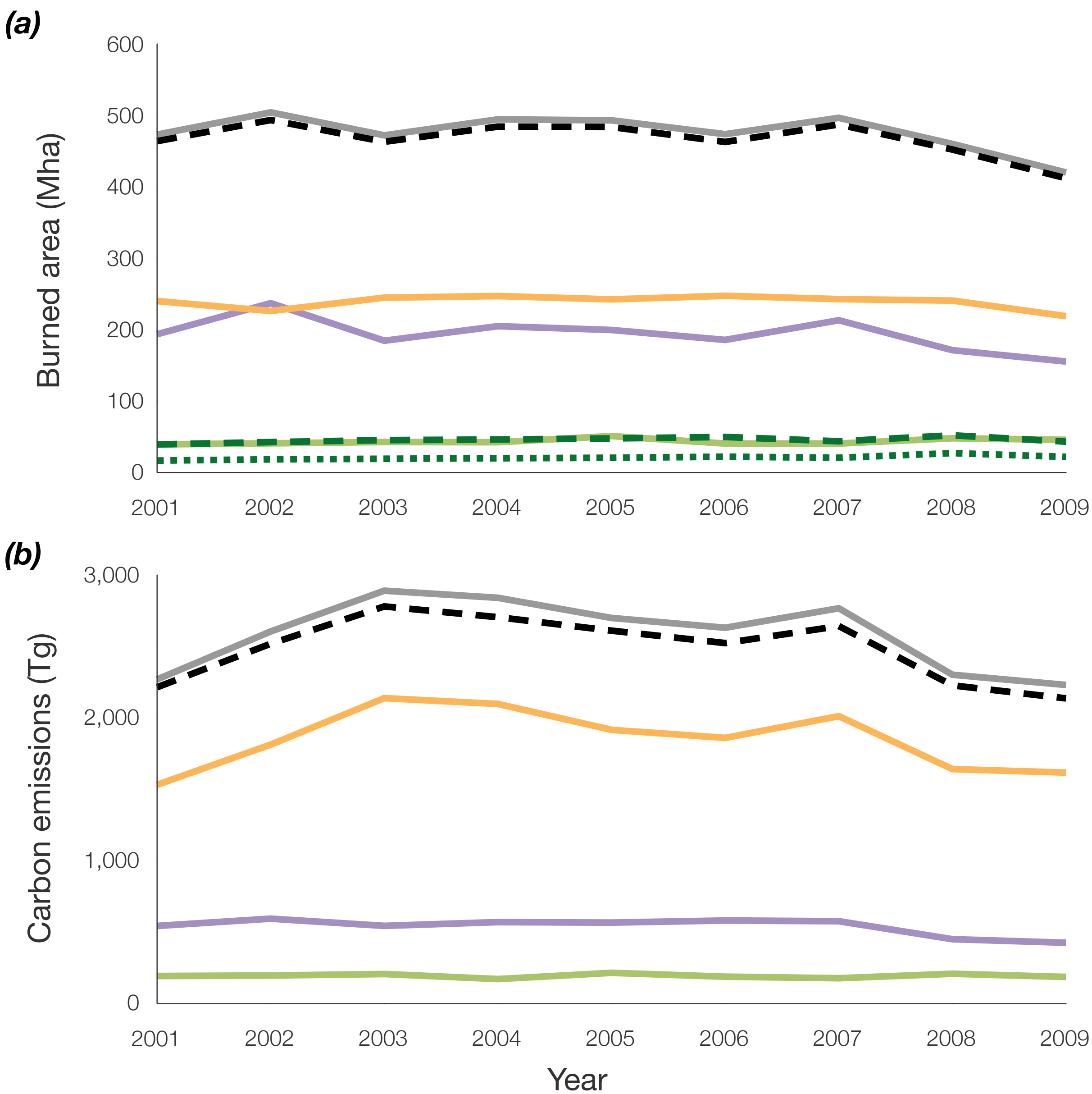

Figure A.2: Observed and estimated annual timeseries of net observed and estimated global burned area (**a**; Mha) and C emissions (**b**; Tg = Mt) from the constrained-$\widehat{F_k}$ analysis. Numbers in table represent annual means. "N.D." = no data; "Crop+" = cropland + cropland-natural mosaic. Corresponds to Fig. 1.2 in main text.

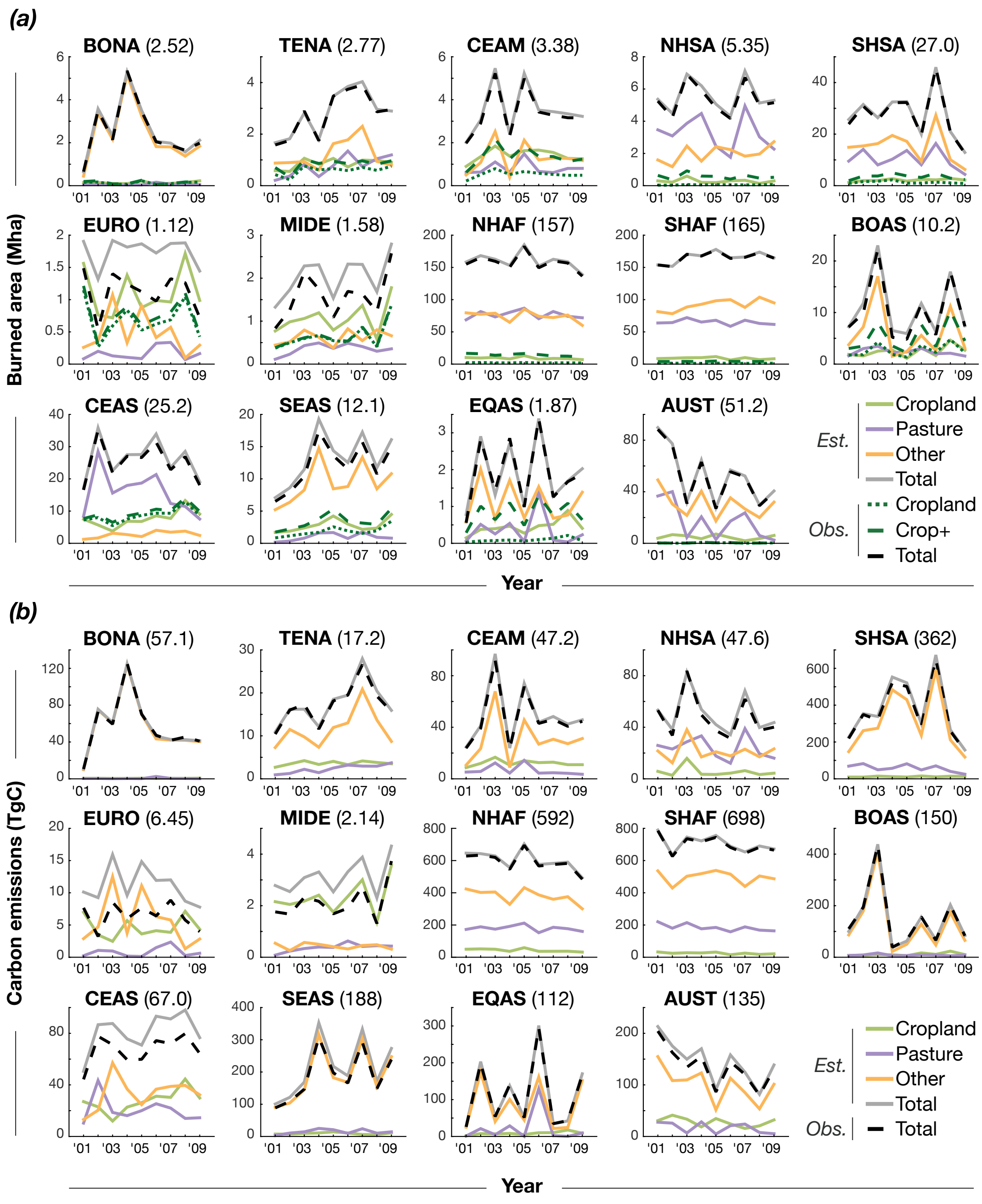

Figure A.3: Annual timeseries of different fire types in each GFED region based on constrained-$\widehat{F_k}$ analysis of burned area (**a**; Mha) and C emissions (**b**; TgC). Numbers in parentheses next to region names represent mean annual observed fire there (either burned area or C emissions). "Crop+" = cropland + cropland-natural mosaic. Corresponds to Fig. 1.3 in main text.

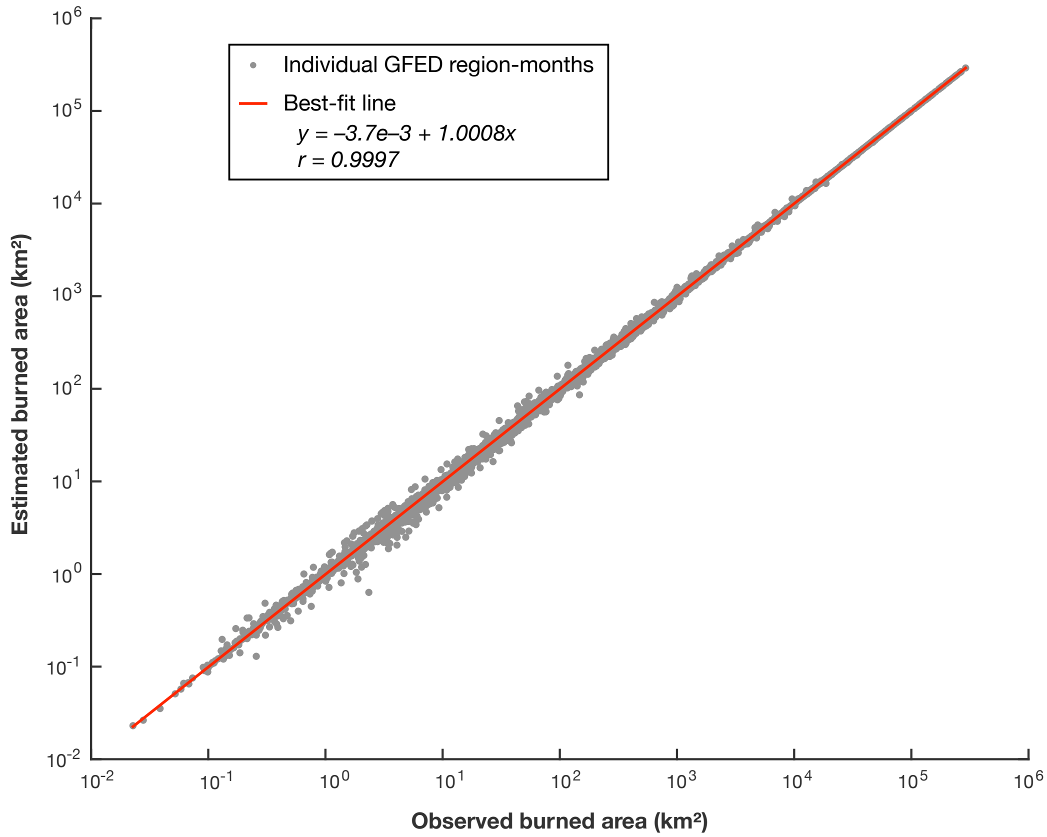

Figure A.4: Scatterplot comparing observed and estimated total burned area for each GFED region and month (gray points), with results of linear regression (red line).

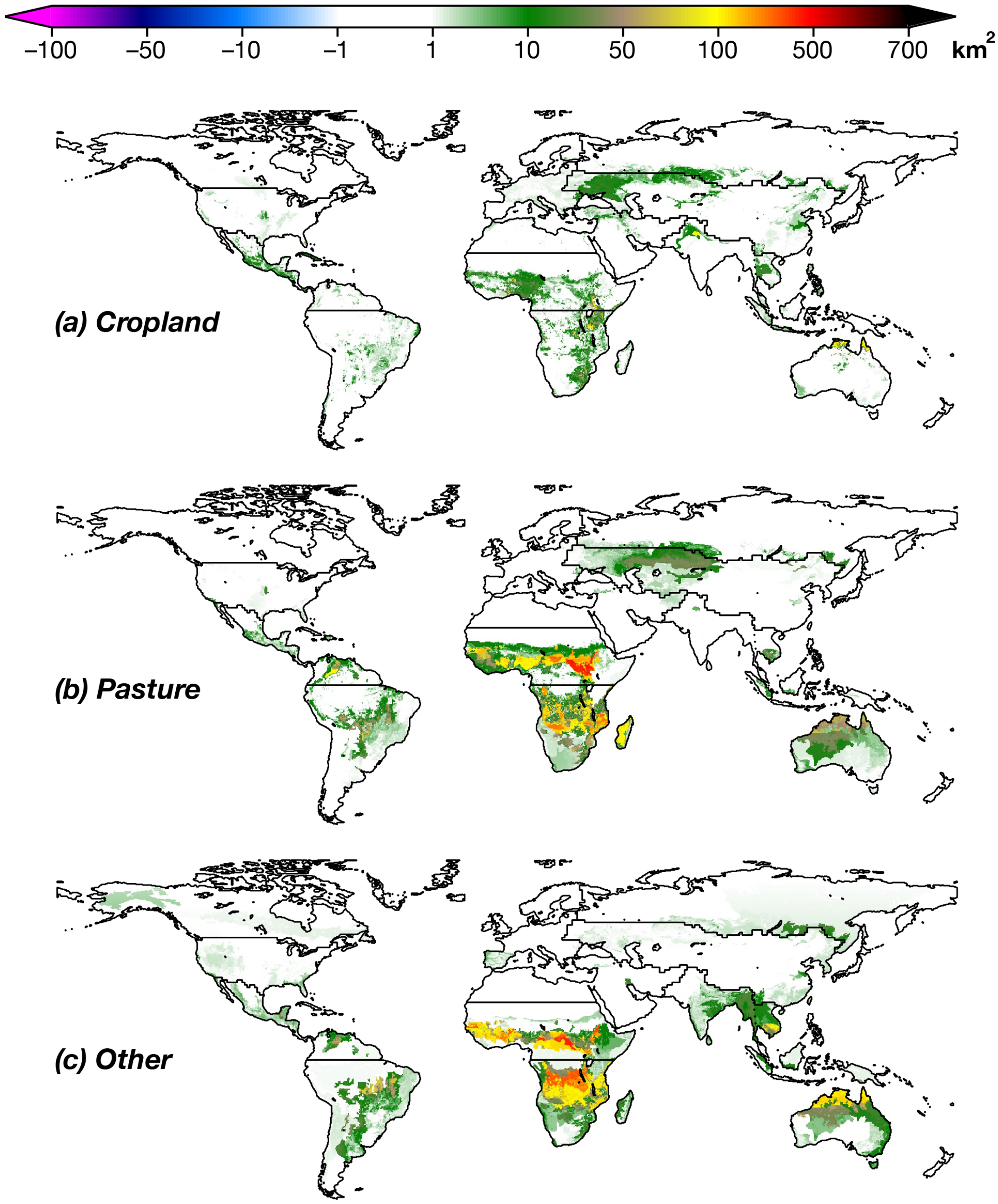

Figure A.5: Maps, from constrained-$\widehat{F_k}$ analysis, of mean annual burned area (km^2) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. These are calculated from monthly maps generated by the equation $B_i = \widehat{F_k} A_{k,i}$ for each month and region. The results can be interpreted as how much more (or less) fire would be expected if the area of the given land cover were to double (and the others remain the same). Corresponds to Fig. 1.4 in main text.

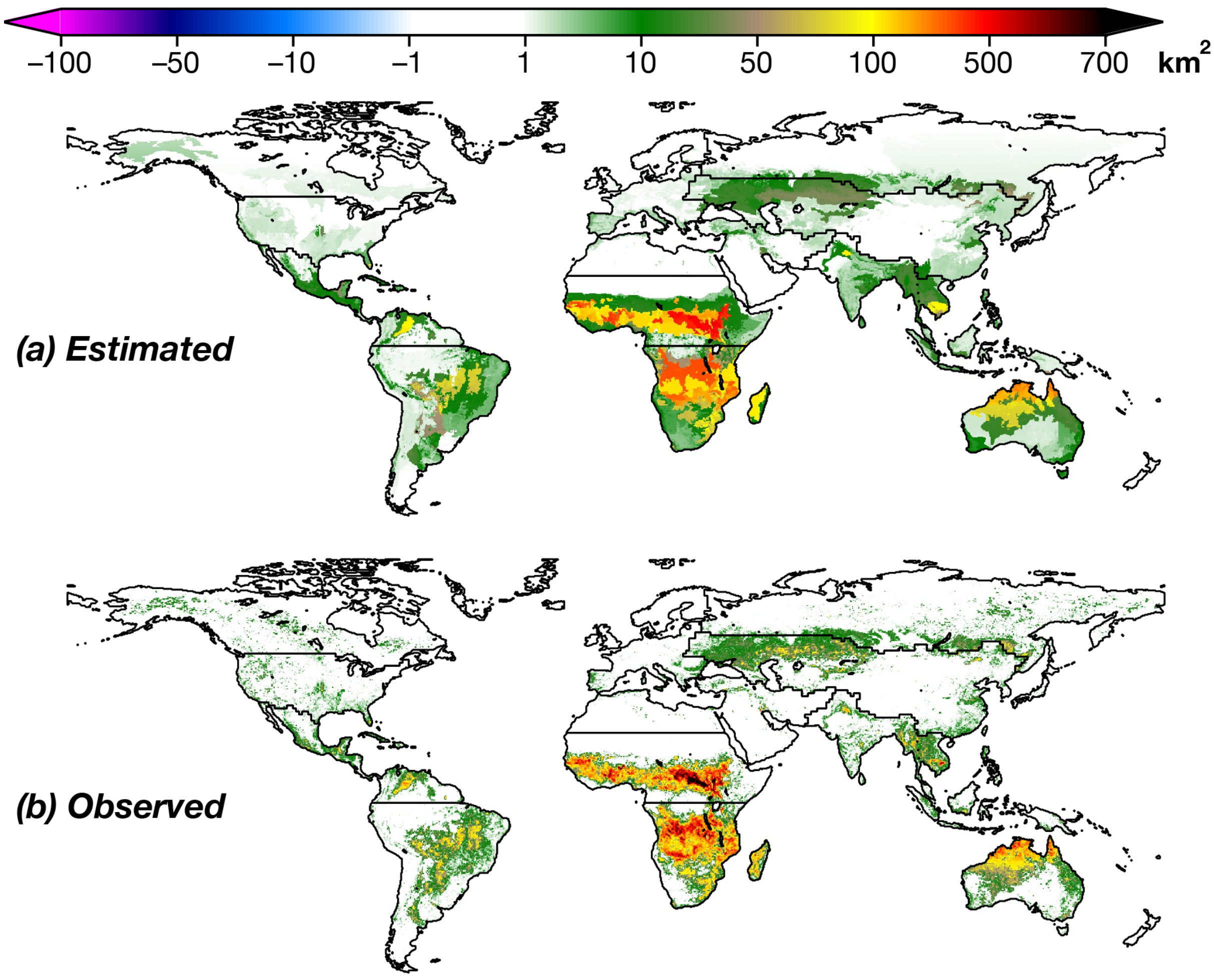

Figure A.6: Maps of mean annual total burned area (km^2): **(a)** Estimated by constrained-$\widehat{F_k}$ analysis. **(b)** Observed. Corresponds to Fig. 1.5 in main text.

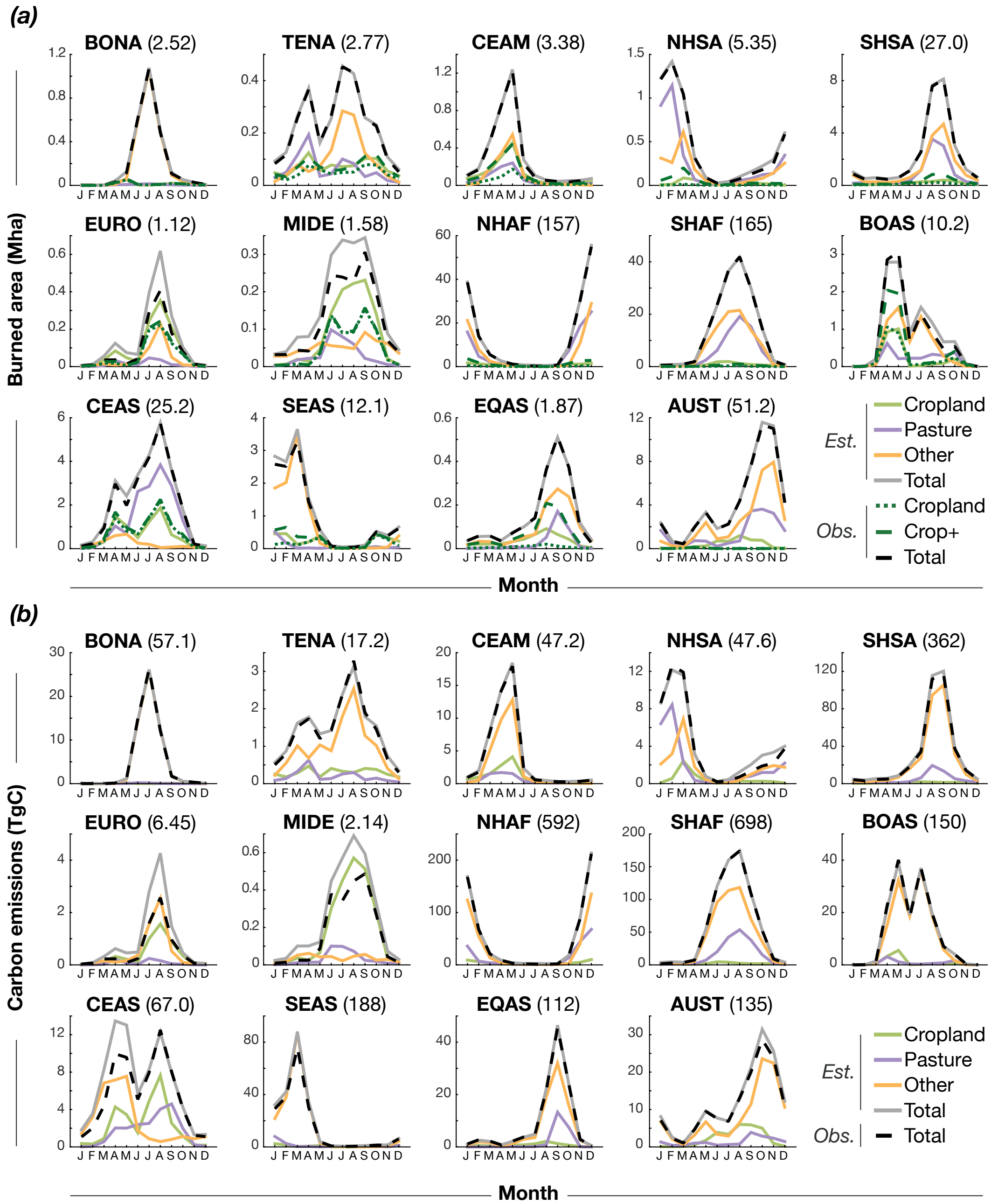

Figure A.7: Seasonality of different fire types in each GFED region based on constrained-$\widehat{F_k}$ analysis of burned area (**a**; Mha) and C emissions (**b**; TgC). Numbers in parentheses next to region names represent mean annual observed fire there (either burned area or C emissions). Corresponds to Figure 1.6 in main text.

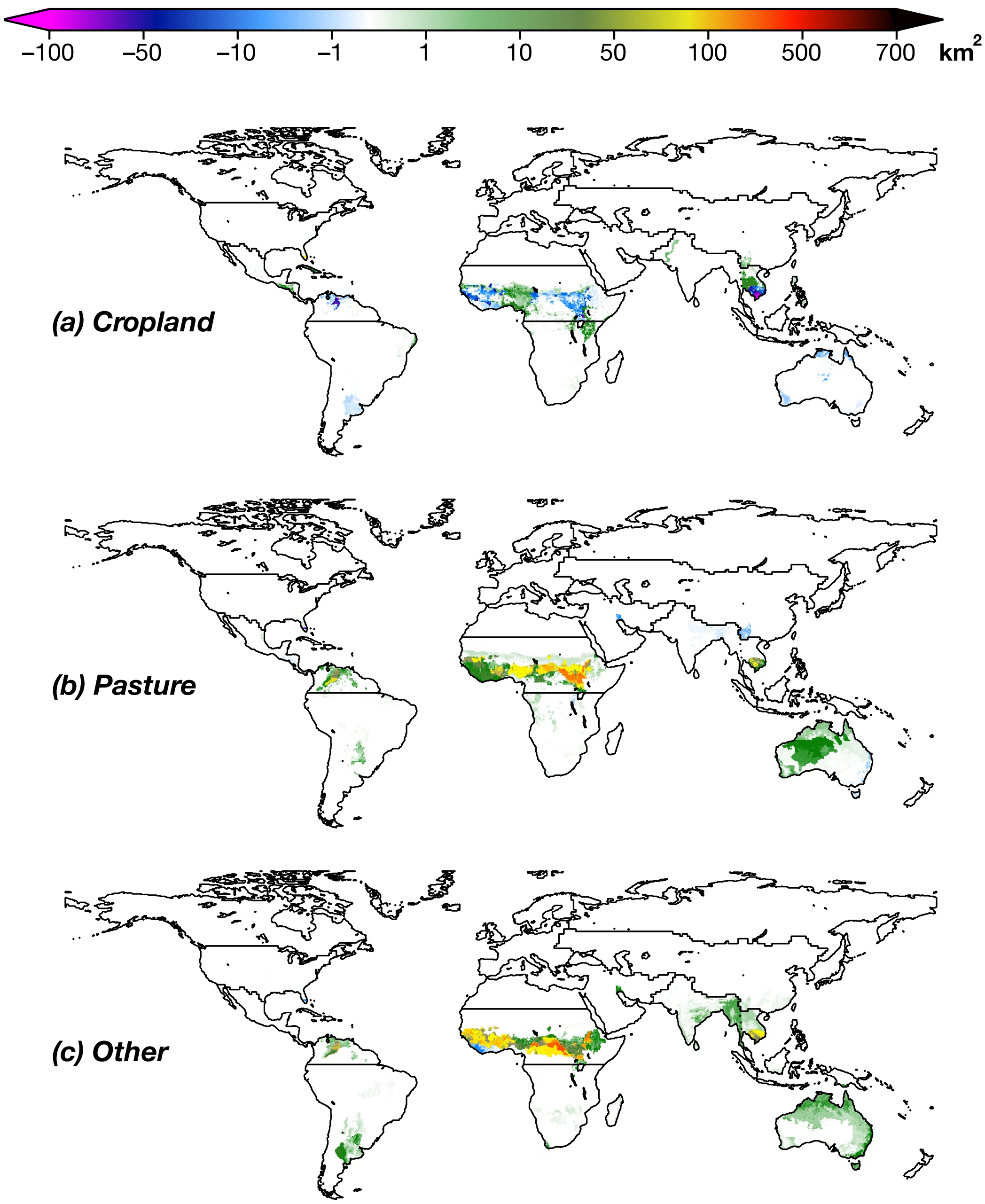

Figure A.8: Maps of mean burned area (km^2) during December, January, and February (DJF) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. Compare with annual means in Figure 1.4.

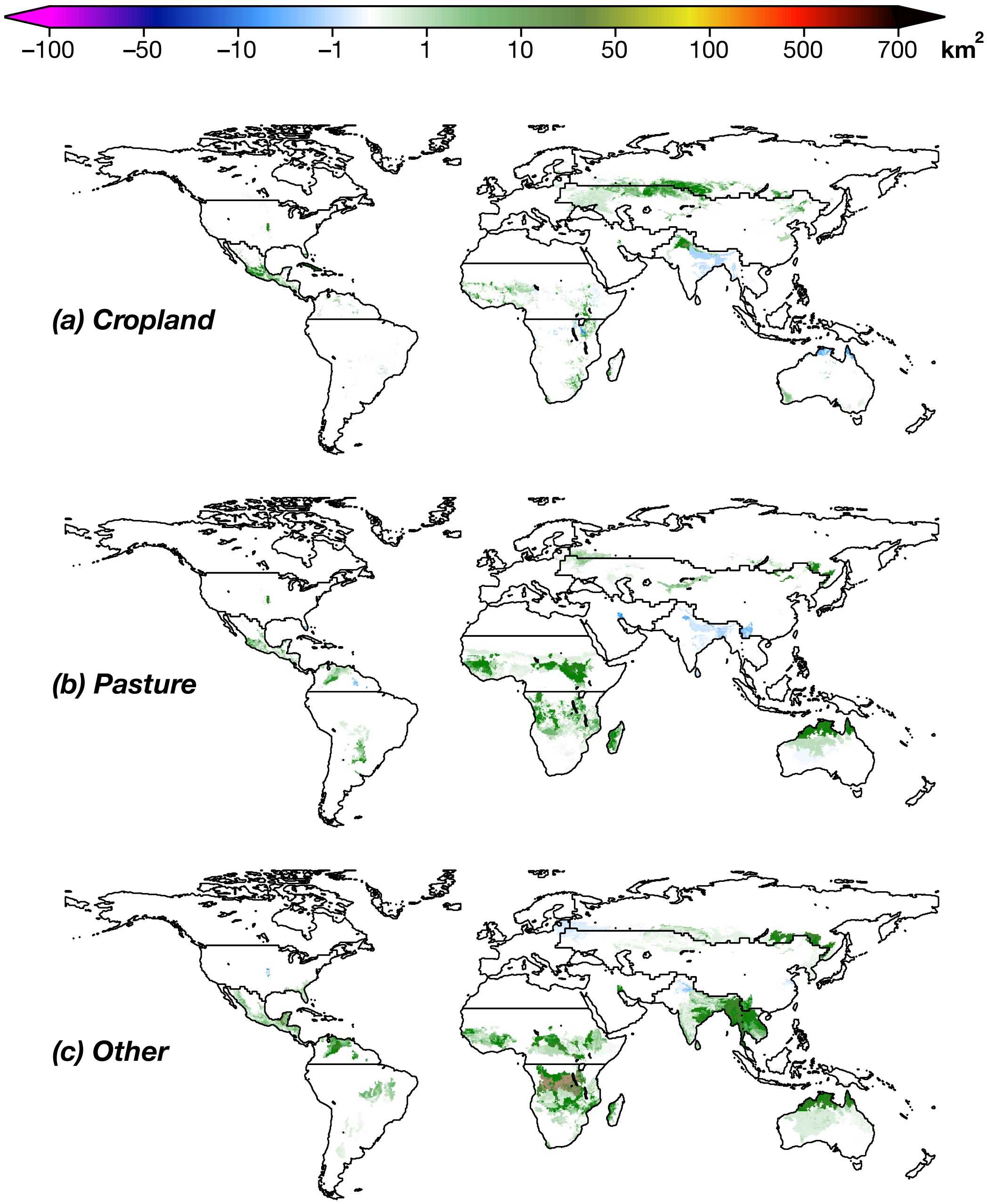

Figure A.9: Maps of mean burned area (km^2) during March, April, and May (MAM) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. Compare with annual means in Figure 1.4.

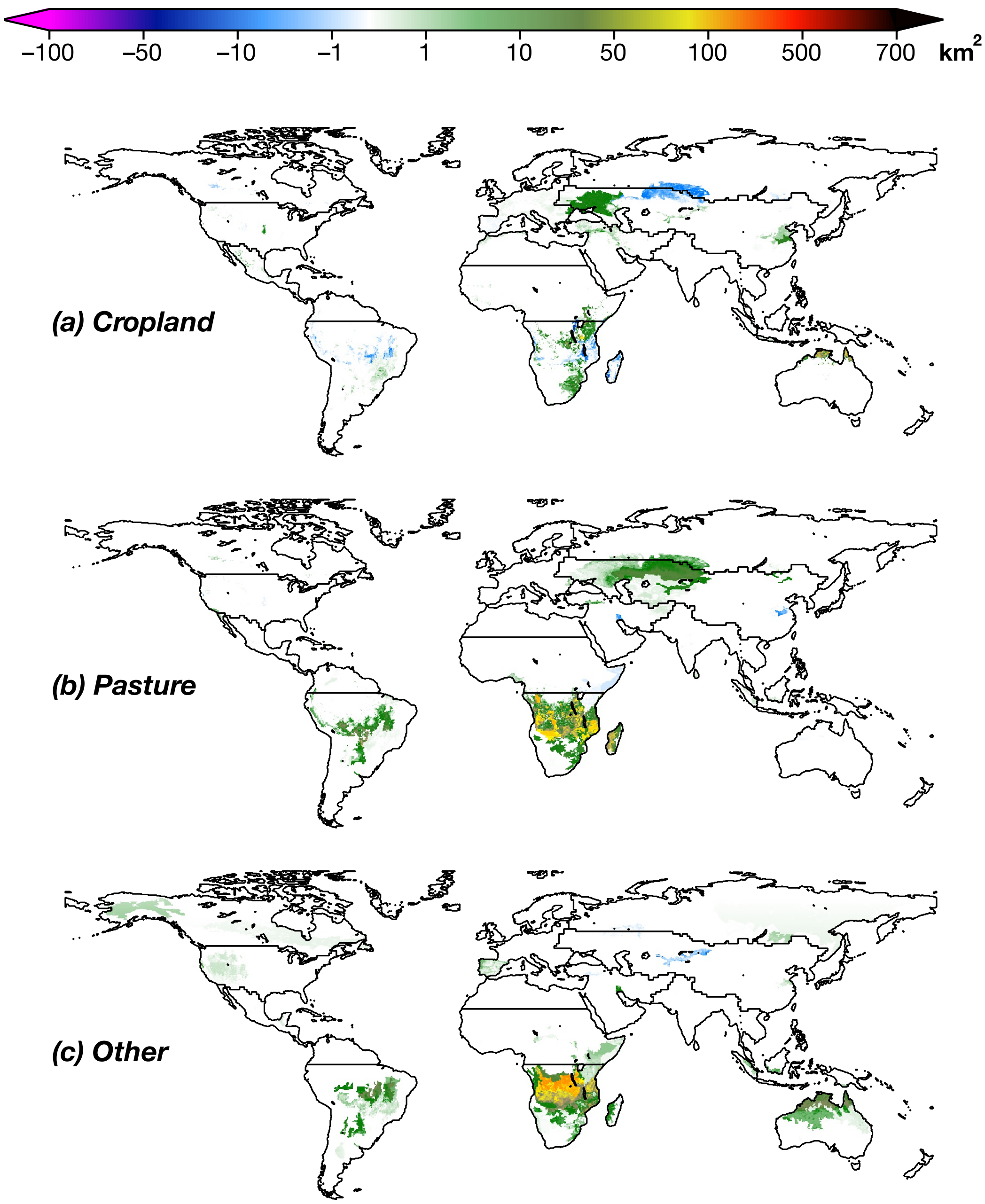

Figure A.10: Maps of mean burned area (km^2) during June, July, and August (JJA) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. Compare with annual means in Figure 1.4.

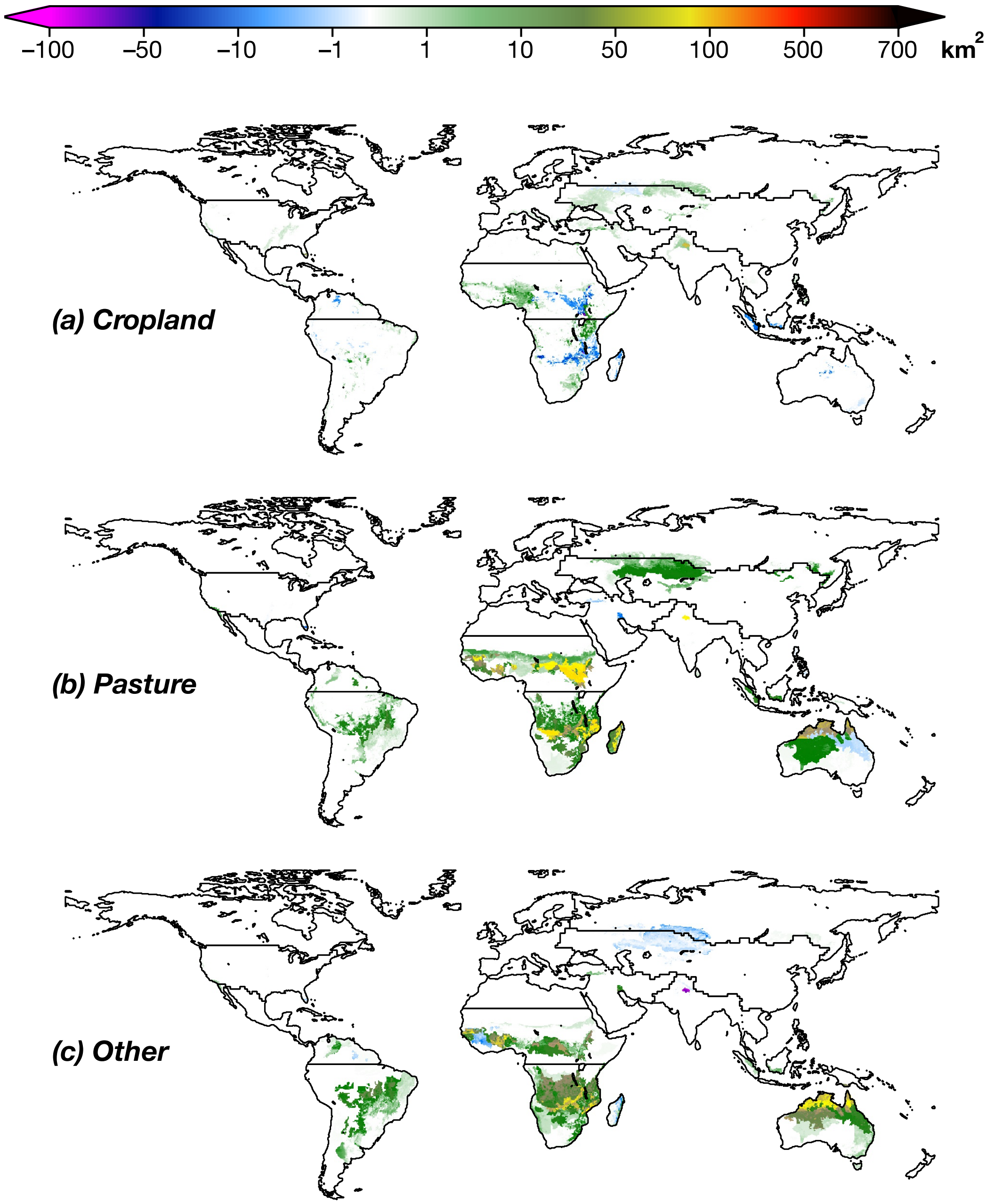

Figure A.11: Maps of mean burned area (km^2) during September, October, and November (SON) associated with **(a)** cropland, **(b)** pasture, and **(c)** other land. Compare with annual means in Figure 1.4.

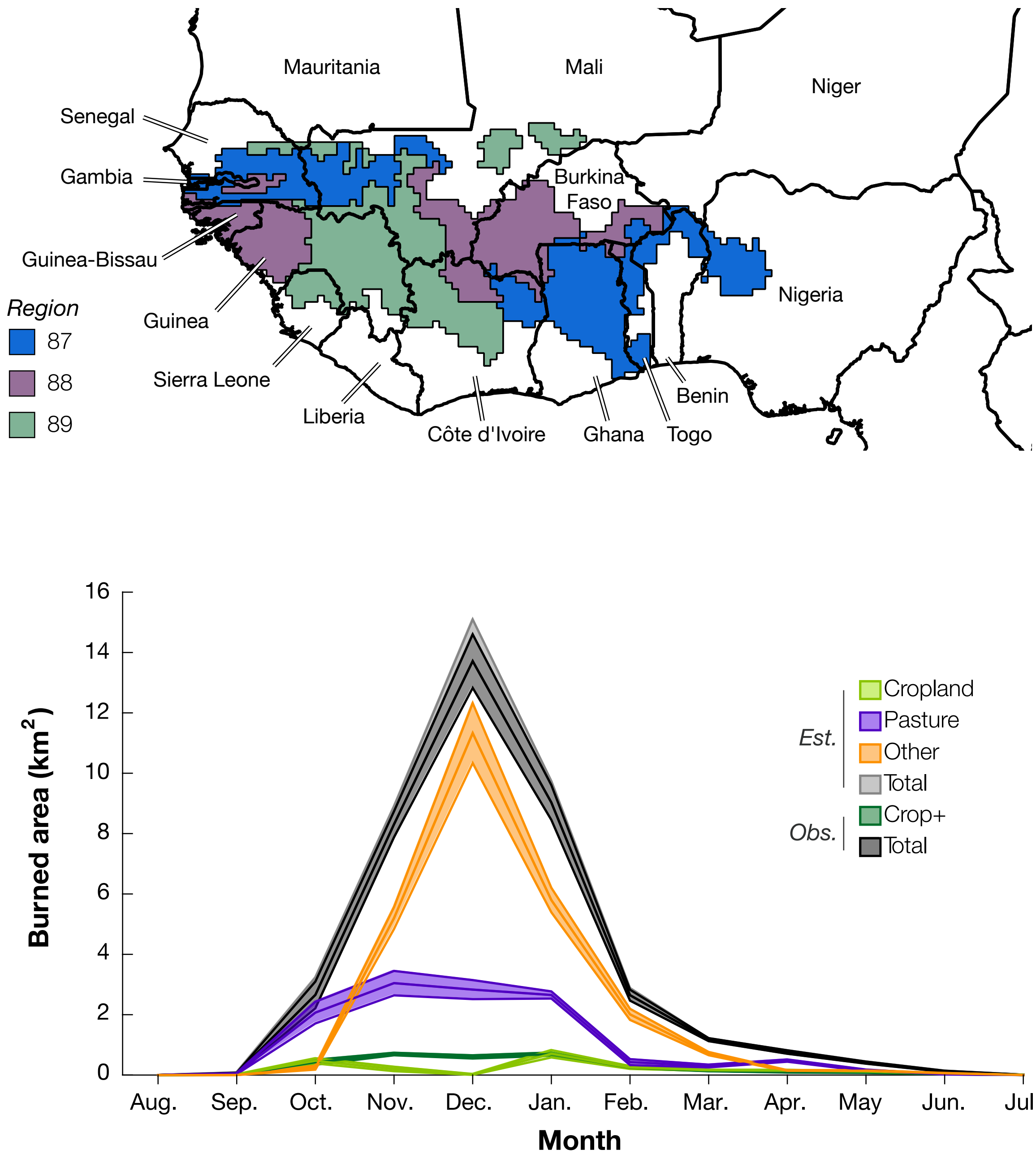

Figure A.12: **(a)** Area included in West African case study, color-coded by analysis region. **(b)** Mean seasonality of burned area in case study regions based on constrained-$\widehat{F_k}$ analysis. Shading represents interannual variability (± 1 SEM). Note that the X axis begins in August. Corresponds to Figure 1.7 in main text.

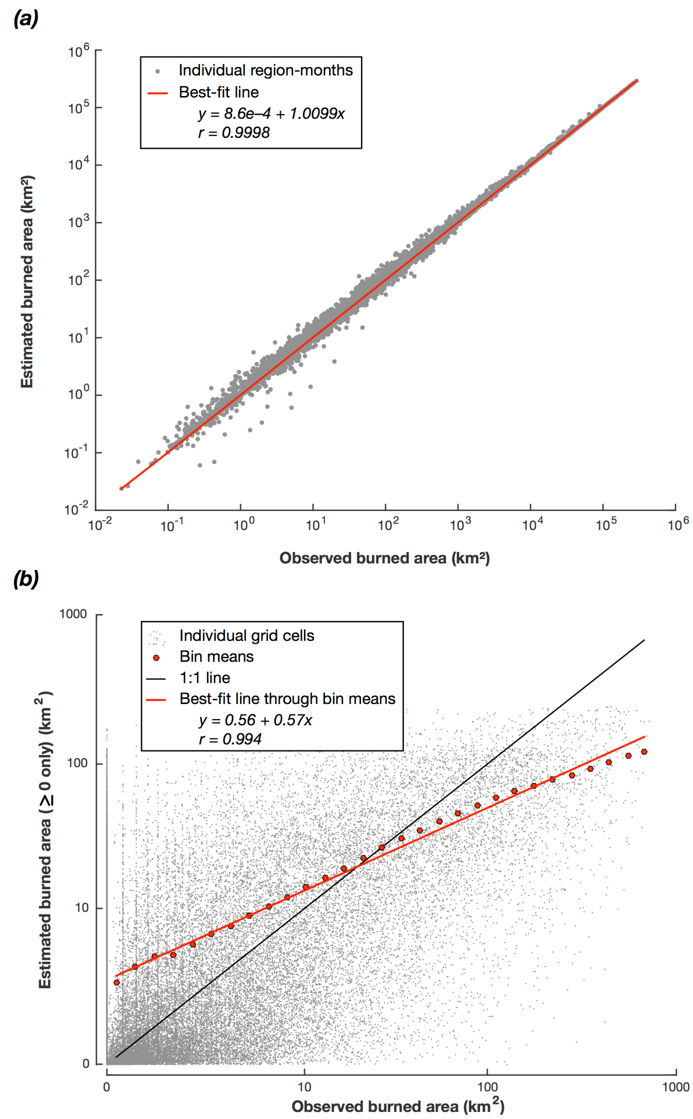

Figure A.13: Scatter plots comparing estimated burned area from constrained-$\widehat{F_k}$ analysis with observations. Gray points represent **(a)** each analysis region and month (region-month), or **(b)** individual gridcells ($\frac{1}{75}$ of cells chosen at random for plotting). Red lines represent the best-fit line from linear regression, with the regression in **(b)** fit to the red points, which represent mean observed and estimated values of gridcells in bins of observed burned area equally spaced along the X axis (with at least 100 gridcells required for a bin to be included). Values ≤ 0 not shown due to log-scale axes. Gridcells in region-months with no observed fire, where the analysis was not performed, were excluded from both plots and regressions. Corresponds to Figure 1.8 in main text.

Appendix B

Levenberg-Marquardt algorithm: Derivation

To understand how the Levenberg-Marquardt algorithm works, we can begin by walking through the Gauss-Newton method for a hypothetical problem. (Throughout this section, lowercase bold letters will indicate vectors and uppercase bold letters will indicate matrices.) Let's say we have some function that we are using to generate estimates of something:

$$f_i = f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right), \tag{B.1}$$

where f_i is the value estimated by f for instance i (a gridcell, for example) given input data $\boldsymbol{x_i}$ (a vector) and parameter set $\boldsymbol{\beta}$ (also a vector). The goal of optimization is to minimize the sum of squared differences (S) between each observed value y_i and each estimate f_i:

$$S(\mathbf{x}, \boldsymbol{\beta}) = \sum_{i=1}^{m} \left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right]\right)^2, \tag{B.2}$$

where $\mathbf{x}$ is the set of all input data for all instances. Taking a series of steps in parameter space towards a minimum of S, eventually settling on some locally optimal set of parameters $\boldsymbol{\beta^*}$. Each step can be thought of as going from one set of parameters, $\boldsymbol{\beta}$, to a new one, $\boldsymbol{\beta} + \boldsymbol{\delta}$. At each iteration, we want to find the set of step sizes $\boldsymbol{\delta}$ that minimize

$$S(\mathbf{x}, \boldsymbol{\beta} + \boldsymbol{\delta}) = \sum_{i=1}^{m} \left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta} + \boldsymbol{\delta}\right]\right)^2, \tag{B.3}$$

There is no way to know *a priori* what the value of f given some new parameter set (i.e., $f\left[\boldsymbol{x_i}, \boldsymbol{\beta}+\boldsymbol{\delta}\right]$) would be, so it must be approximated. This is accomplished via first-order Taylor expansion around the value of f given the current parameter set:

$$f\left(\boldsymbol{x_i}, \boldsymbol{\beta}+\boldsymbol{\delta}\right) \approx f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right) + \frac{\partial f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right)}{\partial \beta_1}\delta_1 + \cdots + \frac{\partial f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right)}{\partial \beta_n}\delta_n \tag{B.4}$$

This can be simplified by substituting in the Jacobian matrix $\mathbf{J}$, defined as

$$\mathbf{J} = \begin{pmatrix} \frac{\partial f(\boldsymbol{x_1}, \boldsymbol{\beta})}{\partial \beta_1} & \cdots & \frac{\partial f(\boldsymbol{x_1}, \boldsymbol{\beta})}{\partial \beta_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial f(\boldsymbol{x_m}, \boldsymbol{\beta})}{\partial \beta_1} & \cdots & \frac{\partial f(\boldsymbol{x_m}, \boldsymbol{\beta})}{\partial \beta_n} \end{pmatrix} = \begin{pmatrix} J_{11} & \cdots & J_{1n} \\ \vdots & \cdots & \vdots \\ J_{m1} & \cdots & J_{mn} \end{pmatrix} \tag{B.5}$$

Equation B.4 thus becomes

$$f\left(\boldsymbol{x_i}, \boldsymbol{\beta}+\boldsymbol{\delta}\right) \approx f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right) + J_{i1}\delta_1 + \cdots + J_{in}\delta_n \tag{B.6}$$

Substituting that in to Equation B.3 gives

$$S(\mathbf{x}, \boldsymbol{\beta}+\boldsymbol{\delta}) = \sum_{i=1}^{m}\left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right] - \sum_{j=1}^{n} J_{ij}\delta_j\right)^2. \tag{B.7}$$

We want to find the set of parameter steps $\boldsymbol{\delta}$ that minimizes $S(\mathbf{x}, \boldsymbol{\beta}+\boldsymbol{\delta})$, so we will take the derivative of S with respect to each parameter step δ_j, set it equal to zero, and solve for δ_j. We can see what this looks like for the first parameter step δ_1:

$$\frac{\partial S(\mathbf{x}, \boldsymbol{\beta}+\boldsymbol{\delta})}{\partial \delta_1} = 2 \times \sum_{i=1}^{m}\left(J_{i1} \times \left[y_1 - f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right) - \sum_{j=1}^{n} J_{ij}\delta_j\right]\right) = 0. \tag{B.8}$$

Then we can rearrange:

$$0 = \sum_{i=1}^{M}\left(J_{m1}\left[\sum_{j=1}^{N} J_{1j}\delta j\right] - J_{m1}\left[y_i - f\left(\boldsymbol{x_i}, \boldsymbol{\beta}\right)\right]\right) \tag{B.9}$$

$$\sum_{i=1}^{M} J_{m1}\left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right]\right) = \sum_{i=1}^{M} J_{m1}\left(\sum_{j=1}^{N} J_{1j}\delta j\right) \tag{B.10}$$

$$\sum_{i=1}^{M} J_{m1}\left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right]\right) = \sum_{j=1}^{N}\left(\delta_j\left[\sum_{i=1}^{M} J_{i1}J_{ij}\right]\right) \tag{B.11}$$

We can now do this for the rest of the parameters and write them all in matrix form:

$$\begin{pmatrix} \sum_{i=1}^{M} J_{m1}\left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right]\right) \\ \vdots \\ \sum_{i=1}^{M} J_{mn}\left(y_i - f\left[\boldsymbol{x_i}, \boldsymbol{\beta}\right]\right) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N}\left(\delta_j\left[\sum_{i=1}^{M} J_{i1}J_{ij}\right]\right) \\ \vdots \\ \sum_{j=1}^{N}\left(\delta_j\left[\sum_{i=1}^{M} J_{in}J_{ij}\right]\right) \end{pmatrix} \tag{B.12}$$

This is much more easily referred to and solved using matrix notation:

$$\mathbf{J^T}\left(\boldsymbol{y} - \boldsymbol{f}\left[\mathrm{x}, \boldsymbol{\beta}\right]\right) = \left(\mathbf{J^T J}\right)\boldsymbol{\delta} \tag{B.13}$$

$$\boldsymbol{\delta} = \left(\mathbf{J^T J}\right)\left(\mathbf{J^T}\left[\boldsymbol{y} - \boldsymbol{f}\left(\mathrm{x}, \boldsymbol{\beta}\right)\right]\right) \tag{B.14}$$

Thus, the Gauss-Newton algorithm generates a new parameter set guess based on the previous guess, the observations and estimates using the previous guess, and the derivative of the sum of squared errors with respect to each parameter at the previous guess.

The Gauss-Newton method should work well when the starting parameter set is near a local minimum. But far from a minimum, or where the curvature matrix $\mathbf{A} = \mathbf{J^T J}$ is poorly conditioned, this algorithm can take extremely large steps and have difficulty converging (Transtrum and Sethna, 2012). To avoid this problem, Levenberg (1944) introduced a version of the Gauss-Newton method that uses a "damped" version of **A**. Instead of Equation B.14, the next parameter guess is calculated as:

$$\boldsymbol{\delta} = \left(\mathbf{A} + \lambda\mathbf{I}\right)\left(\mathbf{J^T}\left[\boldsymbol{y} - \boldsymbol{f}\left(\mathrm{x}, \boldsymbol{\beta}\right)\right]\right) \tag{B.15}$$

where $\mathbf{A} + \lambda\mathbf{I}$ can also be denoted **A'**. If S is decreasing, the damping parameter λ can be decreased, with the algorithm approaching Gauss-Newton as $\lambda \to 0$. However, if the last parameter guess increased S relative to the most recently accepted guess, λ is increased. This can also be thought of as a trust region method: Far from a minimum, we don't trust the Gauss-Newton approximation of the Hessian very far from the current parameter set, so we should only take small steps.

Marquardt (1963) improved upon this concept by increasing the rate of descent specifically in the dimensions where the gradient is smaller (or more generally, change the rate of descent in each dimension based on how steep the gradient is in that dimension):

$$\boldsymbol{\delta} = \left(\mathbf{A} + \lambda \mathbf{I} \boldsymbol{A_{diag}}\right) \left(\mathbf{J^T} \left[\boldsymbol{y} - \boldsymbol{f}\left(\mathrm{x}, \boldsymbol{\beta}\right)\right]\right), \tag{B.16}$$

where $\boldsymbol{A_{diag}}$ is the column vector whose elements correspond to the diagonal elements of $\mathbf{A}$.

Marquardt (1963) describes the result as "a maximum neighborhood method which, in effect, performs an optimum interpolation between the Taylor series method [i.e., Gauss-Newton] and the gradient method, the interpolation being based upon the maximum neighborhood in which the truncated Taylor series gives us an adequate representation of the nonlinear model."

This general form is what's known now as the Levenberg-Marquardt method. However, the damping method described in Equation B.16 – i.e., $\mathbf{A'} = \mathbf{A} + \lambda \mathbf{I} \boldsymbol{A_{diag}}$ – is only one option, called additive damping. Multiplicative damping, where $\mathbf{A'} = \mathbf{A}\left(1 + \lambda \mathbf{I} \boldsymbol{A_{diag}}\right)$ instead, has the advantage of capturing how different parameters might be scaled differently in a poorly-posed problem. However, it also increases the algorithm's susceptibility to parameter evaporation, which is when a parameter ventures into space where the model becomes less sensitive to it (Transtrum and Sethna, 2012). The corresponding element of $\boldsymbol{A_{diag}}$ thus becomes small, increasing the step size in a positive feedback loop. In our implementation, we use multiplicative damping, but avoid parameter evaporation by setting a minimum value of $\boldsymbol{A_{diag}}$ for each parameter – in essence, limiting the maximum step size – after Transtrum and Sethna (2012).

Appendix C

Levenberg-Marquardt algorithm: Implementation

Our implementation of the Levenberg-Marquardt algorithm (Figure C.1) began with a Bash script that set up the files and directories necessary to run the fire model at the 241 points. These points would then be run for 1991–2009 in parallel. Once this first iteration was complete, a Python script calculated the sum of squared errors (S) over each gridcell (c), year (y), and month (m):

$$S = \sum_{c=1}^{241} \sum_{y=2001}^{2009} \sum_{m=1}^{12} \left(E_{c,y,m} - O_{c,y,m} \right)^2 . \tag{C.1}$$

Here, E refers to the model-estimated burned area, and O refers to an observation-based estimate of burned area. Specifically, we focused on non-agricultural lands, using as our "observations" estimates generated for each month and year by the method detailed in Chapter 1 but with $\widehat{F_k}$ estimates restricted to non-negative values. The Python script then generated a new parameter set guess based on the initial values of the parameters and saved a flag telling the Bash script to run the model again with the new guess.

After this and subsequent model runs, another Python script would calculate the associated value of the sum of squared errors (S_t) and compare it to the sum of squared errors from the most recently accepted guess (S^*). If $S_t < S^*$, the current parameter set guess ($\boldsymbol{\beta_t}$) would be "accepted" and become the new value of $\boldsymbol{\beta^*}$, and λ would be decreased. Otherwise, $\boldsymbol{\beta_t}$ would be "rejected," with $\boldsymbol{\beta^*}$ retaining its previous value, and λ being decreased. In either case, a new guess would then be generated based on $\boldsymbol{\beta^*}$ and the new value of λ, the model would be run again, and the process would repeat (Figure C.1).

The Python script we developed was based on a MATLAB routine for Levenberg-Marquardt solutions of nonlinear least squares problems called marquardt.m (Nielsen, 2001), further docu-

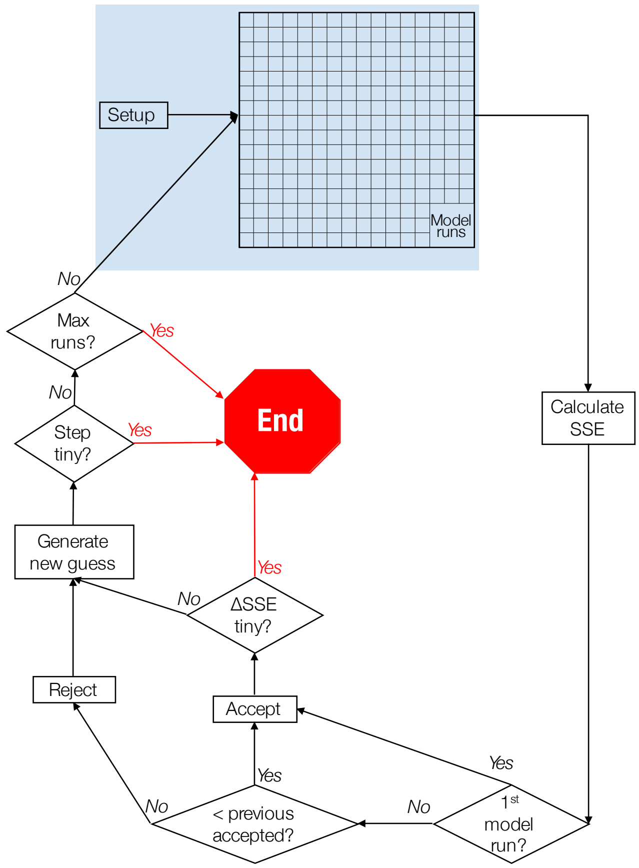

Figure C.1: Flowchart describing our implementation of the Levenberg-Marquardt algorithm. Blue shading indicates operations related to running the model; all other steps occur in Python.

mented in Nielsen (1999). Besides porting it to Python, we made a number of changes to the original code. Some restructuring was related to the fact that the new parameter sets could not be evaluated within Python. Others were to incorporate new features, such as the limited multiplicative damping based on work by Transtrum and Sethna (2012) described above.

Nielsen (2001) uses a somewhat complex method to update $\boldsymbol{\delta}$ after every each iteration (Figure C.2). If $S_t \geq S^*$, λ is multiplied by a value ν, whose initial value is 2 and is doubled after every rejected guess. If a guess is accepted ($S_t < S^*$), ν is reset to 2, and λ is decreased. We made some changes to the original code as a result of the aforementioned restructuring, with λ being reduced as:

$$\lambda = \lambda \times max\left(\frac{1}{3}, 1 - \left[\frac{S}{dL_{t-1}} - 1\right]^3\right) \tag{C.2}$$

where

$$dL_{t-1} = \boldsymbol{\delta_{t-1}} \times \left(\lambda \times \boldsymbol{\delta_{t-1}} - \mathbf{J^T} \times \boldsymbol{f}.\right) \tag{C.3}$$

Note that there have been many methods proposed over the years for updating the damping parameter in the Levenberg-Marquardt algorithm. These impact the size of the steps the algorithm takes while searching through parameter space, with implications for efficiency. However, the math by which the algorithm determines which direction on each dimension to move is unaffected.

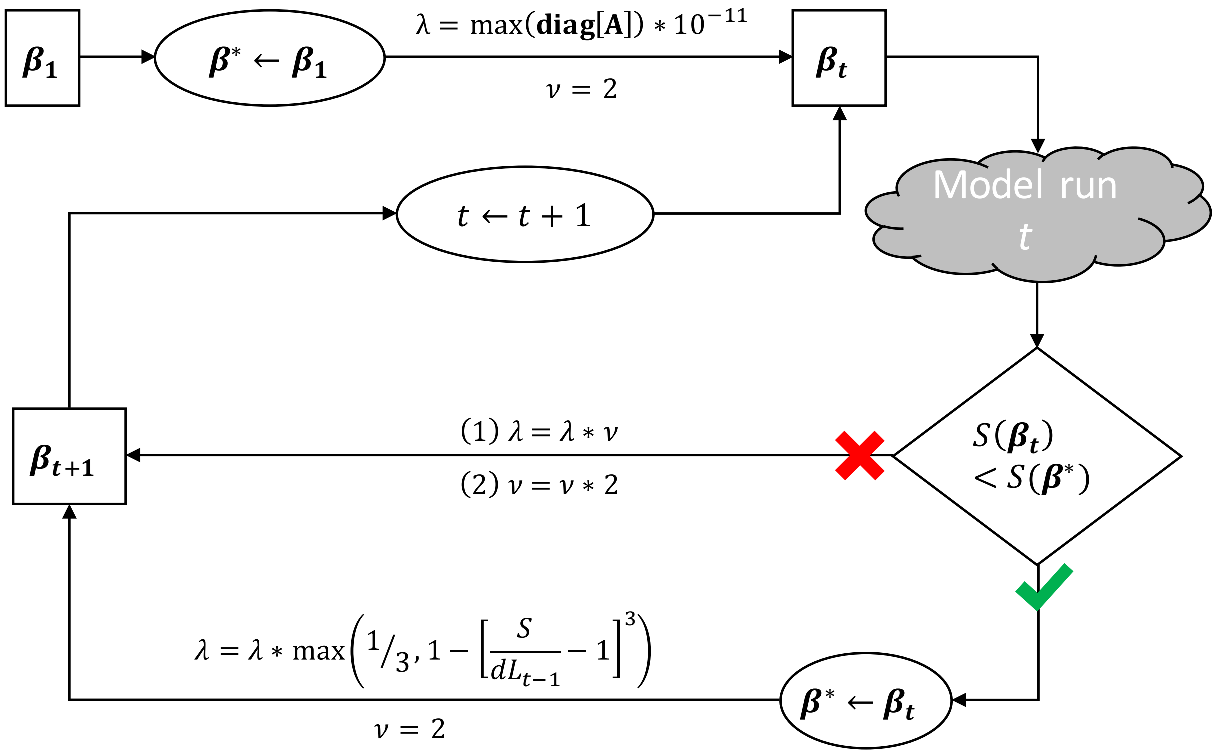

Figure C.2: Method for updating λ, after Nielsen (1999) and Nielsen (2001).

The algorithm has several possible stop conditions. We set a maximum of 300 iterations, which was never reached. The algorithm would also stop if the Python script detected that the gradient was decreasing very slowly:

$$||\mathbf{J^T} \times \boldsymbol{f}||_2 \leq 10^{-15}, \tag{C.4}$$

if the step size was very small:

$$||\boldsymbol{\delta_t}||_2 \leq 10^{-15} \times ||\boldsymbol{\beta^*}||_2, \tag{C.5}$$

or there was an issue of near-singularity in one of matrices involved in solving for the new parameter step:

$$||\boldsymbol{\delta_t}||_2 \geq \frac{||\boldsymbol{\beta^*}||_2}{\epsilon}, \tag{C.6}$$

where ϵ is the smallest number allowed by the numerical precision of the Python environment.

However, in practice, we usually ended up halting the algorithm manually. Each iteration took about two hours, and once we noticed neither the sum of squared errors nor any parameter changing by very much, we would stop the runs. This could have been avoided by choosing more appropriate threshold values for the stop conditions, but likely did not appreciably impact the results.

Appendix D

Chapter 3 supplementary tables and figures

Table D.1: Mean carbon pools and fluxes as simulated over 2001–2009 by `PASTASOTHR` (pasture burning using non-agricultural fire model; "old") and `PASTASPAST` (pasture burning using unpacked estimates of pasture burning; "new").

Flux/pool	TOTL old	TOTL new	OTHR old	OTHR new	AGRI old	AGRI new	NTRL old	NTRL new	SCND old	SCND new	CROP old	CROP new	PAST old	PAST new
gpp	224	221	104	103	47.7	47.7	90.2	89.2	48.8	47.8	38.5	38.6	46.1	45.8
npp	107	106	45	44.8	26.7	26.7	38.4	38	21.5	21.1	20.4	20.4	26.9	26.7
nep	32.6	33.5	4.88	4.99	14.4	14.8	2.91	3.15	4.14	4.26	8.1	8.22	17.5	17.9
resp	116	115	58.5	58.3	21	21	51.8	51.2	27.3	26.7	18.1	18.1	19.3	19.1
co2litter, leaf	26	25.3	14.7	14.5	3.47	3.37	12.6	12.4	7.39	7.11	3.24	3.22	2.73	2.6
co2litter, wood	21.7	21	13.3	13.2	2.61	2.39	12.3	12.1	4.64	4.5	2.07	2.03	2.68	2.37
co2soil	26.8	26.5	12.2	12.1	6.23	6.16	10.5	10.4	5.3	5.21	6.96	6.94	3.96	3.87
crop harv rate	8.19	8.22	5.41e-15	5.88e-15	4.4	4.42	0	0	1.17e-14	18.19e-14	8.19	8.22	1.07e-16	1.53e-16
past harv rate	18.6	18.3	0.214	0.206	10.3	10.2	0	0	0.473	0.452	0.0068	0.00662	18.1	17.8
wood harv rate	1.25	1.2	0.749	0.739	0.0583	0.0529	0	0	1.13	1.09	0.0678	0.0641	0.0518	0.0484
totl harv rate	28	27.7	0.964	0.945	14.8	14.6	0	0	1.6	1.54	8.26	8.29	18.1	17.9
burnCemitRate	0.935	2.34	0.264	0.472	0.242	0.59	0.243	0.558	0.285	0.774	0.298	0.297	0.109	0.712
bl	41.5	40.9	33.6	33.5	1.71	1.7	30.3	30	8.18	7.94	1.61	1.61	1.43	1.41
blv	7.01	6.73	4.34	4.29	0.553	0.516	3.39	3.3	2.72	2.59	0.359	0.36	0.543	0.483
bsw	296	291	250	249	3.07	2.93	226	223	64.8	62.9	2.91	2.92	2.73	2.47
bwood	363	355	275	273	4.96	4.51	278	273	76.8	74.7	0.93	0.927	7.45	6.67
br	17.2	17	10.9	10.9	2.42	2.42	9.77	9.66	3.13	3.06	2.33	2.33	1.94	1.93
agb	576	565	458	455	8.69	8.17	436	430	124	121	5.04	5.05	10.1	9.2
litter lf	51	50.1	37.1	36.9	3.89	3.8	32.7	32.3	11.4	11.1	3.97	3.93	2.9	2.8
litter cw	134	132	107	106	7.91	7.58	95.8	94.9	23	22.5	6.38	6.28	8.44	8.02
litter	185	182	144	143	11.8	11.4	129	127	34.5	33.5	10.3	10.2	11.3	10.8
rh	74.4	72.8	40.1	39.8	12.3	11.9	35.5	34.9	17.3	16.8	12.3	12.2	9.37	8.85
nbp	3.7	3.47	3.65	3.57	-0.593	-0.443	2.67	2.59	2.25	1.95	-0.462	-0.364	-0.76	-0.7

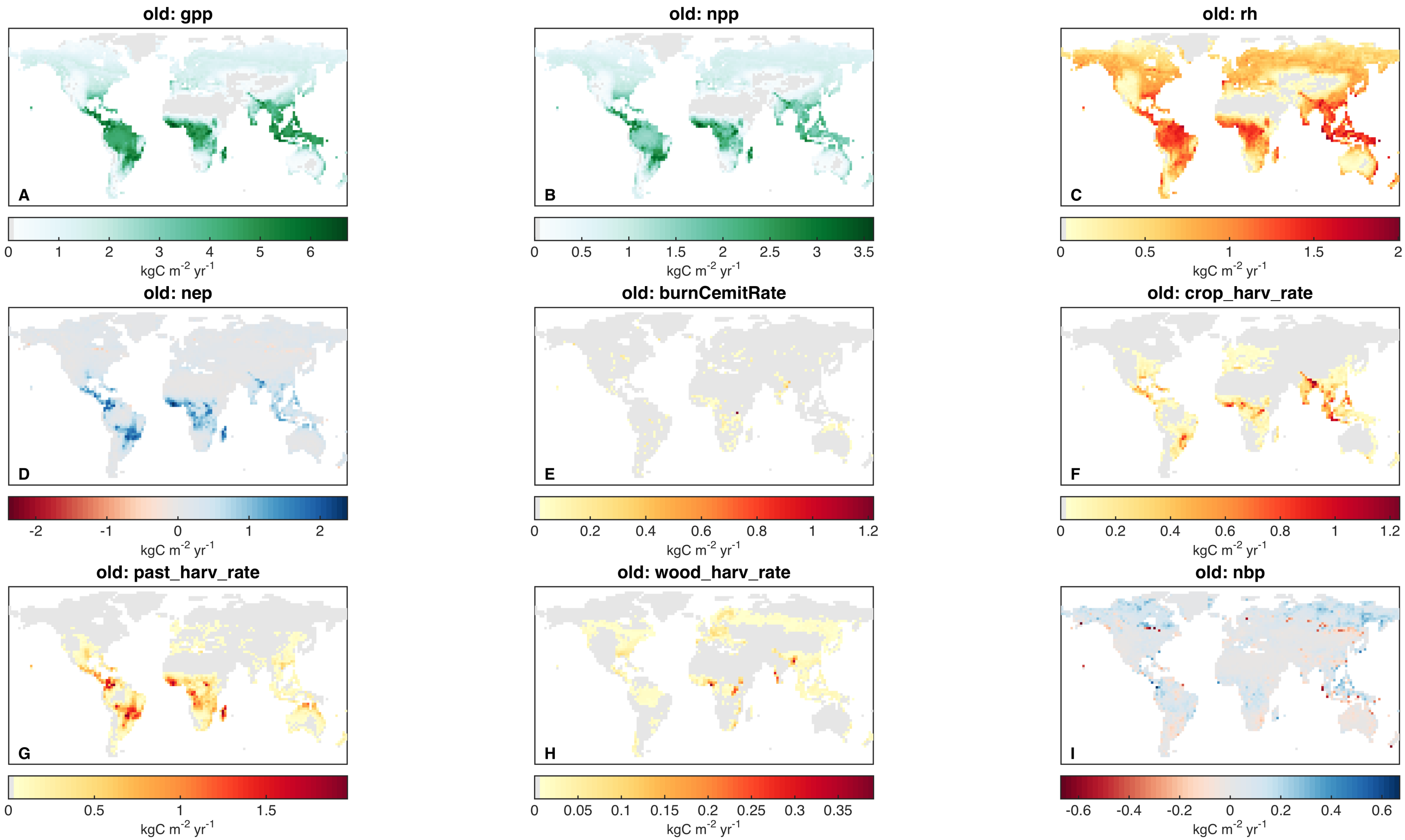

Figure D.1: Mean annual carbon fluxes simulated over 2001–2009, `PASTASOTHR`. Note that this is for all vegetation types, not just pasture.

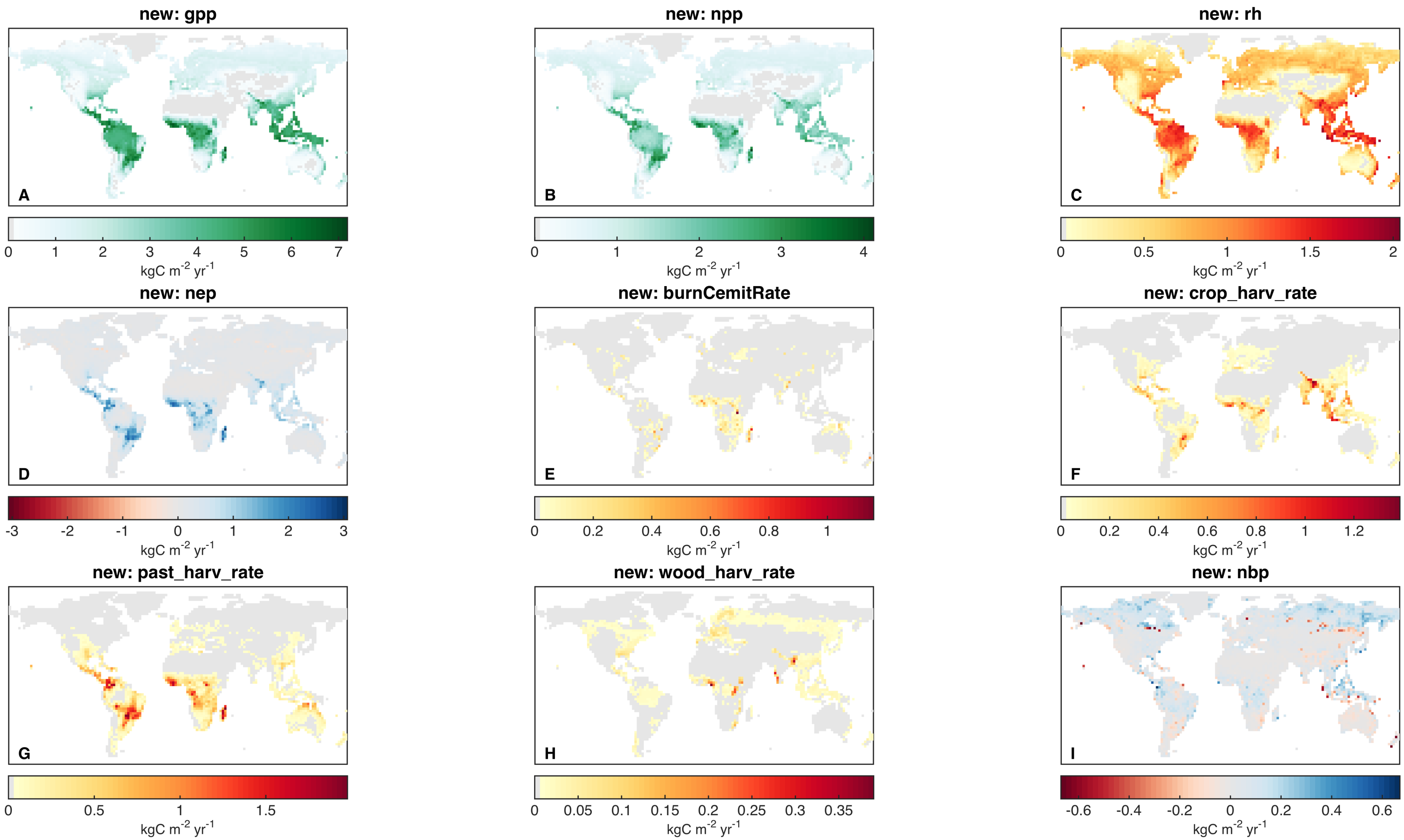

Figure D.2: Mean annual carbon fluxes simulated over 2001–2009, PASTASPAST. Note that this is for all vegetation types, not just pasture.

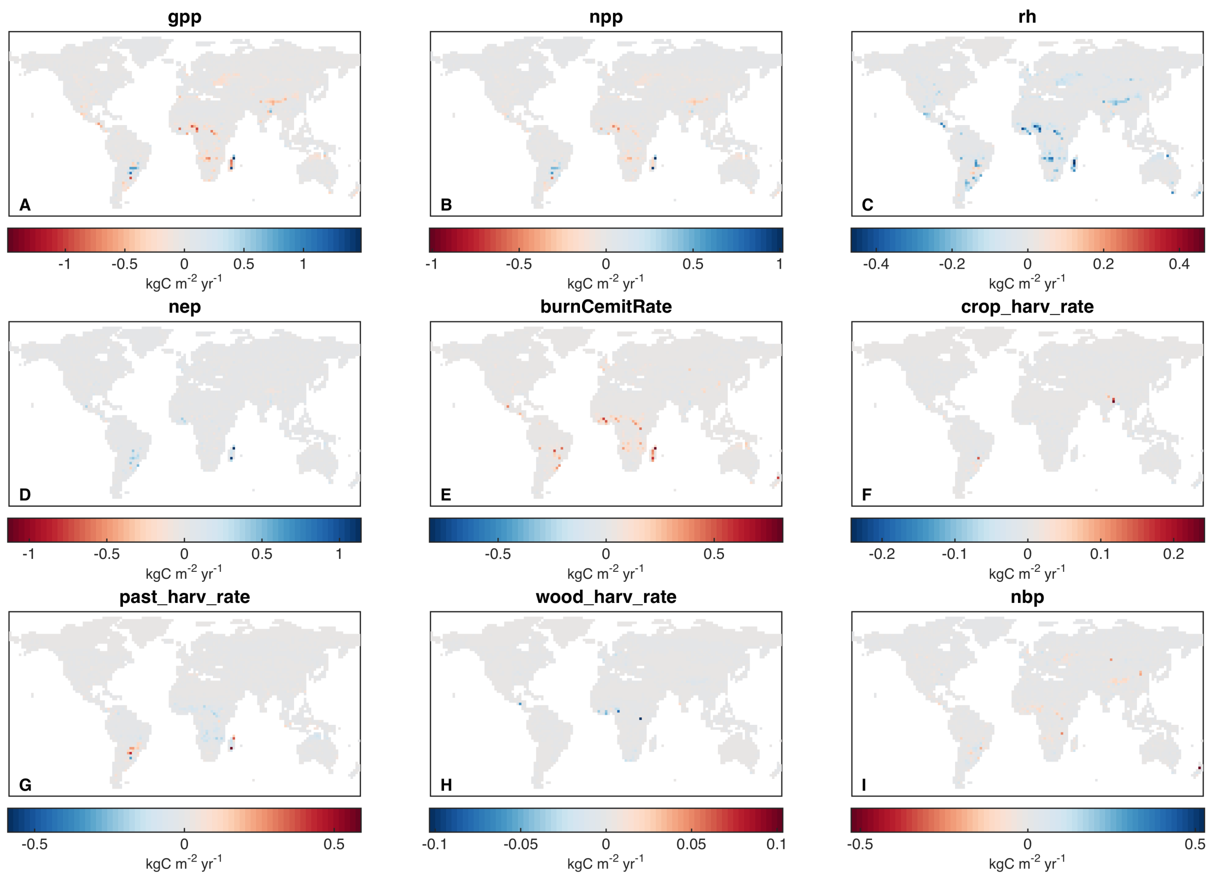

Figure D.3: Difference in selected mean annual carbon fluxes between `PASTASPAST` and `PASTASOTHR`, 2001–2009. Blue indicates increased ecosystem C retention in `PASTASPAST`; red, reduced. Note that this is for all vegetation, not just pasture. See Figures D.1–D.2 for separate maps for each run.

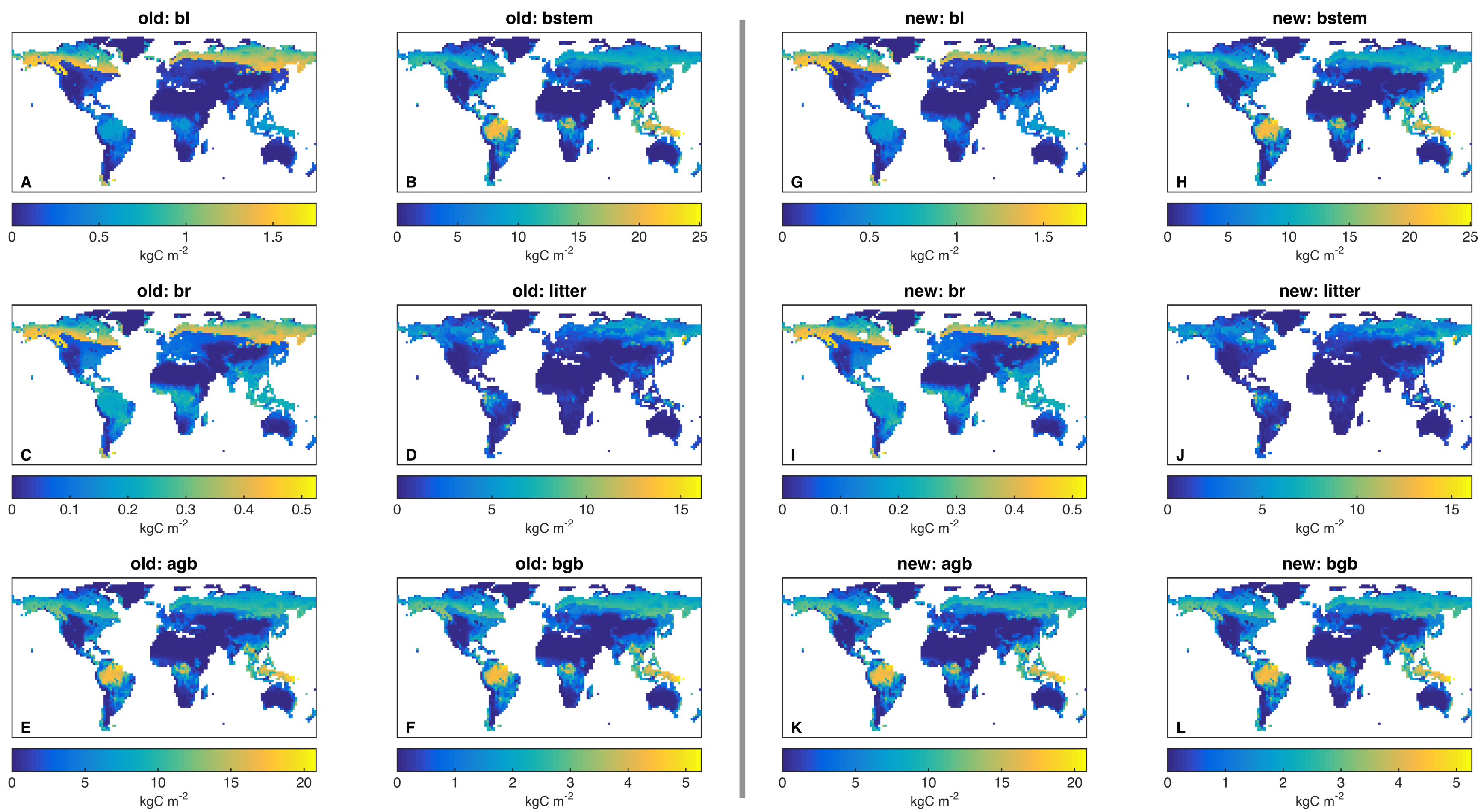

Figure D.4: Mean carbon pools simulated over 2001–2009 by **(a–f)** `PASTASOTHR` (pasture burning using non-agricultural fire model; "old") and **(g-l)** `PASTASPAST` (pasture burning using unpacked estimates of pasture burning; "new") runs. Note that this is for all vegetation types, not just pasture.

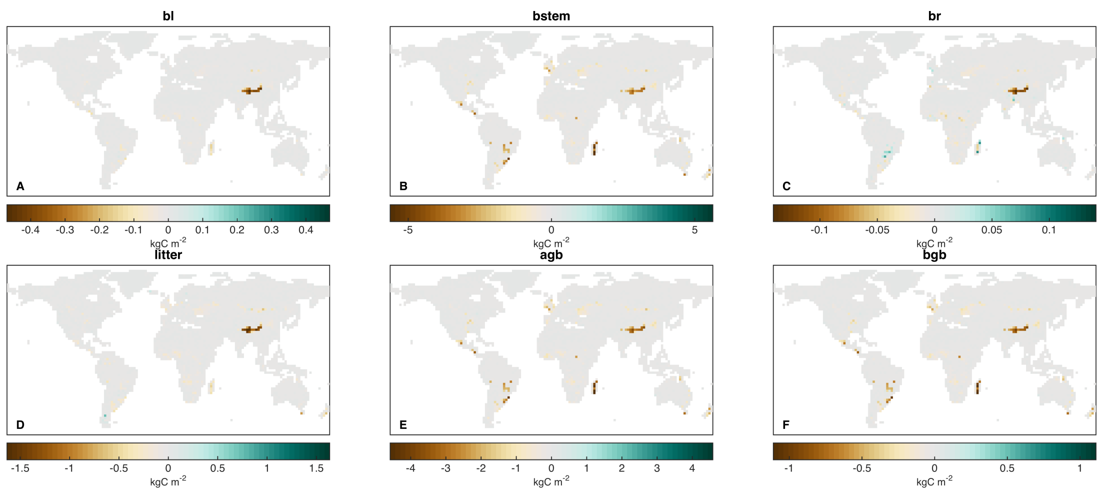

Figure D.5: Difference in selected mean carbon pools between `PASTASPAST` and `PASTASOTHR`, 2001–2009. Green indicates increased ecosystem C pools in `PASTASPAST`. Note that this is for all vegetation, not just pasture. See Figure D.4 for separate maps for each run.