Dissertation. ETH No. 27467

On the Biogeophysical Consequences of Forestation

A thesis submitted to attain the degree of
Doctor of Sciences of ETH Zurich
(Dr. sc. ETH Zurich)

presented by
Ronny Andreas Meier

Master degree in Environmental Sciences, ETH Zurich
born on 30.10.1991
citizen of
Seuzach ZH

supervised and revised by
Prof. Dr. Sonia I. Seneviratne
Dr. Edouard L. Davin
Dr. Sebastiaan Luyssaert

2021

Acknowledgments

My PhD would not have been possible in this form without the help of numerous people. First of all, I would like to thank my supervisors Edouard and Sonia for offering me the opportunity to conduct this PhD. Edouard, I must have walked into your office looking for good advice more than a hundred times (obviously not during the last year). And still, you managed to make me feel like I am always welcome. Sonia, I experienced you as a warm-hearted professor, who always tries to provide the best support for her PhD students and her group in general.

I thank Sebastiaan for acting as the external expert for my PhD defence. You provided a lot of insightful questions and comments, which advanced my learning process even at the very last stage of my PhD.

When writing this thesis, I frequently struggled when deciding whether to use the subject I or we. Eventually, I decided to use the subject I consistently, following examples of previous doctoral theses in our group. Yet, almost everything presented in this thesis is the effort of several people. I would therefore like to express my gratitude to all the people with which I worked during my PhD. First of all, I would like the thank the Post-Doc in the CLIMPULSE project, Jonas, for the nice collaboration. I think I never worked before with a person that is constantly in such a good mood and I could profit a lot from your knowledge. Next, I thank Quentin, who was the co-supervisor of my Master Thesis, for his advice and for helping me to have a good start in the group. Then, I would like to thank all my fellow Landclimers for the lively, serious and less serious discussions. We always had a relaxed atmosphere in our office, for which I would like to thank the office mates I had over the years: Kathrin, Ryan, Joel, Martin, and Iris. Also, thanks to Dominik and Ryan for the good times when installing the lysimeters out in the field. It always felt like a small vacation, even though it was oftentimes quite stressful. Further, I thank all the co-authors of my studies collectively. It was great to experience this vibe of mutual support, which many researchers share.

I would like to praise the Institute for Atmospheric and Climate Science as a whole. Many people have worked hard over the years to make this one of the top institutes in the field, which is able to offer great opportunities for its members. On several occasions, I was able to profit from the expertise that just happens to be around. In particular, I would like to highlight the awesome IT support from Urs and Mathias, which I will surely miss when I leave. In this context, I would also like to thank Matthieu Leclair and Eric Maisonnave for their technical support with the regional climate simulations.

Dann möchte ich mich auch bei all den Leuten bedanken die mich nicht nur wissenschaftlich auf meinem Weg durchs Leben begleitet haben. Zuoberst auf dieser Liste stehen natürlich meine Eltern. Mama und Papa, ich konnte immer mit euch sprechen, wenn mich etwas beschäftigt hat. Ich hatte stets den Eindruck, dass ich eure vollste Unterstützung geniesse, egal welchen Weg ich als nächstes einschlage. Auch bin ich dankbar, dass mein Bruder, Alex, in meinem Leben ist. Ich bin zuweilen immer noch überrascht, wie ähnlich wir denken. In der zweiten Hälfte meine PhDs konnte ich auf die Unterstützung von Anina zählen, die mir nicht nur emotional sondern zuweilen auch mit

wissenschaftlichen Ratschlägen zur Seite stand. Vielen Dank für die Geduld, wenn mich mein PhD mal ein bisschen mehr absorbiert hat, als es eigentlich gut ist. Zuletzt möchte ich mich auch bei all meinen Freunden und Bekannten dafür bedanken, dass ihr mich auch mal von meinem PhD abgelenkt habt, sei es mit Wanderungen, Konzerten, Filmabenden, Chillen am See, Gesellschaftsspielen oder schlicht und einfach einem Bier.

Abstract

Humans currently utilize 69-76 % of the ice-free land surface. The associated Land Use and Land Cover Change (LULCC) affects the local, regional, and global climate. The climate impact of LULCC comprises the release or sequestration of greenhouse gases, the biogeochemical effects, and the alteration of the local energy and water redistribution at the land surface, the biogeophysical effects. Observations and models often disagree on the size and even the sign of biogeophysical effects from LULCC, even though they indicate that those effects are relevant for the local and regional climate. Most scenarios that confine global warming levels to below 2 °C incorporate substantial alterations of human land use, often to sequester greenhouse gases. It is therefore crucial to obtain a thorough understanding of the biogeophysical effects of LULCC and reconcile them in models with observations.

Clearing of natural forests for agricultural food production has been a widespread LULCC in the past. This trend is now reverted in some of the developed countries. In addition, re- or afforestation is frequently proposed as a tool to mitigate anthropogenic greenhouse gas emissions. This thesis therefore investigates the biogeophysical effects of re- or afforesting grassland and cropland, which is subsequently called forestation. To this aim, I employ both climate model simulations and analysis of observational data. In particular, I evaluate and undertake targeted improvements to reconcile the local biogeophysical effect of forestation in the Community Land Model (CLM) with observations. Further, I investigate whether forestation might affect precipitation in Europe employing observational data sets.

In Chapter 2, I confront the local biogeophysical sensitivity of CLM to forestation with various observational constraints. It appears that CLM agrees reasonably with observations regarding the sensitivity of albedo, daily mean Land Surface Temperature (LST), and daily maximum LST. Nonetheless, the albedo decrease following forestation is more pronounced in CLM compared to remote sensing observations. The daily maximum LST is distinctly lower in forests than over grassland/cropland both in observations and in CLM with the exception of winters at higher latitudes. However, CLM exhibits a slight positive bias in the daily maximum LST difference of forest minus grassland/cropland. The latter bias appears to be linked to a pronounced underestimation of the increase in EvapoTranspiration (ET) following forestation compared to various observational constraints, which do however exhibit a substantial spread themselves. Subsequently, I propose various modifications of the model to improve its ET sensitivity to forestation, which also reduce the positive bias in the effect of forestation on daily maximum LST. The simulated daily minimum LST difference between forest and grassland/cropland by and large resembles the sensitivity of daily maximum LST, although somewhat weaker, while remote sensing observations indicate that the daily minimum LST of forests is often higher than the one of grassland/cropland. Overall, this study indicates that CLM can represent some aspects of the local biogeophysical sensitivity to forestation well, while further model development is required for other aspects to reconcile CLM with observations.

In the next chapter, I investigate whether the lack of Biomass Heat Storage (BHS) in CLM is responsible for the identified biases in the sensitivity of daily maximum and minimum LST to forestation. The cooling of daily maximum temperatures is marginal, as most of the energy uptake by the vegetation biomass is compensated by a reduction of the turbulent heat fluxes. On the other hand, this process results in a pronounced warming of nighttime temperatures in forests, because the stable structure of the surface layer at night inhibits the compensation of the energy release from the vegetation by the sensible heat flux. The resultant nighttime warming frequently exceeds 2 °C in forests, while BHS appears negligible for grassland and cropland, due their comparably small amount of biomass. Given this diurnal asymmetry, BHS warms daily mean temperatures in forested regions. CLM overestimates the diurnal temperature range in forests compared to remote sensing observations, which is improved substantially after including BHS in the model. Finally, I show that the inclusion of BHS alleviates the apparent deficiency of CLM related to the impact of forestation on the daily minimum LST, which emerged in Chapter 2. In summary, BHS strongly modulates nighttime temperatures in forests and is also relevant for the daily mean temperature, while its impact on daytime temperatures is only marginal.

In Chapter 4, I estimate alterations of precipitation from foresting agricultural land in Europe, a biogeophysical effect that has been largely disregarded in observational studies previously. This is done in two (almost independent) approaches: Firstly, I identify suitable site pairs in two rain gauge data collections that differ by at least 20 % in the agricultural land and forest fractions. Secondly, I model the climatology of a state-of-the-art spatially-continuous precipitation data set with Generalized Additive Models (GAMs) to link precipitation to land cover. In both approaches, forestation is estimated to increase precipitation locally, in particular during the winter months. The structure of the GAMs further allows to estimate precipitation changes downwind of the forestation locations. During winter, downwind precipitation increases in the southern and western parts of Europe, while the signal is near-neutral to negative in central and northern Europe. During summer, I find a downwind increase in precipitation due to forestation. The combined local and downwind effect from a realistic reforestation scenario are estimated to compensate a substantial fraction of the reduction in summertime precipitation, which is expected under RCP4.5 by the end of this century in an ensemble of regional climate models. While this study implies that forestation results in relevant alterations of precipitation in Europe, I would also like to highlight that this study is novel to the field and therefore more uncertain than the previous ones.

Forestation results in biogeophysical effects that are relevant for both the local and regional climate. Such effects should be considered before utilizing forestation as a tool to mitigate greenhouse gas emissions. Yet, many aspects regarding the biogeophysical effects of LULCC in observations and models are still uncertain or unknown. In this thesis, I demonstrate that BHS is relevant for the local climate in forests and should consequently be included in the next generation of earth system models that are used to assess the climate impact of LULCC. Further, I provide observational evidence of changes in precipitation following forestation in Europe. LULCC induces therefore not only temperature alterations, but also relevant modifications of the hydrological cycle, which need to be considered when assessing the climatic consequences of LULCC.

Zusammenfassung

Der Mensch nutzt momentan 69-76 % der eisfreien Landoberfläche. Die damit verbundenen Änderungen in der Landnutzung und –bedeckung (engl. LULCC für "Land Use Land Cover Change") wirken sich auf das lokale, regionale und globale Klima aus. Die klimatischen Folgen von LULCC lassen sich in biogeochemische und biogeophysikalische Effekte unterteilen. Während biogeochemische Effekte die Emission und die Sequestrierung von Treibhausgasen betreffen, beziehen sich die biogeophysikalischen Effekte auf die Änderung der lokalen Energie- und Wasserverteilung an der Landoberfläche. Beobachtungen und Modelle widersprechen sich häufig bezüglich der Grösse und sogar des Vorzeichens von biogeophysikalischen Effekten durch LULCC, obwohl beide zeigen, dass diese Effekte relevant für das lokale und regionale Klima sind. Die meisten Klimaszenarien, welche die globale Erwärmung auf 2 °C beschränken, beinhalten beachtliche Änderungen der menschlichen Landnutzung, meist um Treibhausgase zu sequestrieren. Es ist daher von entscheidender Bedeutung, ein gründliches Verständnis der biogeophysikalischen Auswirkungen von LULCC zu erlangen und diese in Modellen mit Beobachtungen in Einklang zu bringen.

Die Abholzung von Wäldern zur Nutzung als Landwirtschaftsflächen war in der Vergangenheit eine weitverbreitete LULCC. Dieser Trend wendet sich jetzt in einigen Industriestaaten. Ausserdem steht Aufforstung als Massnahme zur Sequestrierung von Treibhausgasen zur Debatte. Diese Arbeit untersucht daher die biogeophysikalischen Effekte durch Aufforstung von Gras- und Kulturflächen, was im Folgenden als Aufforstung bezeichnet wird. Hierfür habe ich sowohl Simulationen mit Klimamodellen als auch Beobachtungsdaten benutzt. Insbesondere habe ich die biogeophysikalischen Effekte von Aufforstung im Landoberflächenmodell "Community Land Model" (CLM) evaluiert und gezielte Verbesserungen am Modell vorgenommen, um diese Effekte mit Beobachtungen in Übereinstimmung zu bringen. Des Weiteren habe ich mit Hilfe von Beobachtungsdaten analysiert, wie sich Aufforstung auf den Niederschlag in Europa auswirkt.

In Kapitel 2 wird die lokale Sensitivität von CLM gegenüber Aufforstung mit mehreren beobachtungsbasierten Datensätzen verglichen. Dabei zeigt sich, dass die Sensitivität der Albedo, des Tagesdurchschnitts der Temperatur an der Landoberfläche (engl. LST für "Land Surface Temperature") und des Tagesmaximums der LST (im Nachfolgenden Tagesmaximum-Temperatur genannt) in CLM mit den Beobachtungsdaten hinreichend übereinstimmt. Allerdings ist die Verringerung der Albedo aufgrund von Aufforstung in CLM stärker ausgeprägt als in den Satellitendaten. Sowohl in den Beobachtungen als auch in CLM ist die Tagesmaximum-Temperatur in Wäldern geringer als auf Gras-/Kulturflächen, abgesehen von den höheren Breitengraden im Winter. Diese Temperaturdifferenz zwischen Wald und Gras-/Kulturland ist in CLM verglichen mit Beobachtungsdaten zu positiv, was sich auf eine starke Unterschätzung der Erhöhung von EvapoTranspiration (ET) durch Aufforstung zurückführen lässt. Die verschiedenen Beobachtungsdaten, die ich für diesen Vergleich hinzugezogen habe, weisen jedoch selbst eine grosse Spannweite auf. Darauffolgend schlage ich mehrere Modellmodifikationen vor, um die Sensitivität von ET gegenüber Aufforstung zu verbessern, was gleichzeitig

den positiven Fehler des Effekts von Aufforstung auf die Tagesmaximum-Temperatur verringert. Für das Tagesminimum der LST gleicht die modellierte Differenz zwischen Wäldern und Gras-/Kulturflächen im Grossen und Ganzen jener des Tagesmaximums, ist aber etwas schwächer ausgeprägt. Dies steht im klaren Widerspruch zu den Beobachtungsdaten, in denen das Tagesminimum der LST von Wäldern häufig höher ist als dasjenige von Gras-/Kulturflächen. Insgesamt weist diese Studie darauf hin, dass CLM einige Aspekte der lokalen biogeophysikalischen Sensitivität gegenüber Aufforstung gut wiedergeben kann, während für andere Aspekte weitere Modellentwicklungen von Nöten sind, um CLM mit Beobachtungsdaten in Übereinstimmung zu bringen.

Im nächsten Kapitel untersuche ich, ob sich die oben beschriebenen Defizite der Sensitivität des Tagesmaximums und des Tagesminimums der LST auf das Fehlen von Wärmespeicherung in der Biomasse (engl. BHS für "Biomass Heat Storage") in CLM zurückführen lassen. Der Kühlungseffekt von BHS auf die Tagesmaximum-Temperaturen ist marginal, da der Grossteil der Energieaufnahme der Biomasse durch eine Reduktion der turbulenten Wärmeflüsse kompensiert wird. Dagegen führt BHS zu einer deutlichen Erwärmung von Wäldern in der Nacht, weil die stabile Struktur der atmosphärischen Oberflächenschicht verhindert, dass die von der Biomasse freigegebene Energie durch Änderungen des sensiblen Wärmeflusses kompensiert wird. Die resultierende nächtliche Erwärmung übersteigt in Wäldern häufig 2 °C, wohingegen BHS für Gras- und Kulturflächen aufgrund derer vergleichsweise kleinen Biomasse vernachlässigbar zu sein scheint. Wegen der beschriebenen tageszeitlichen Asymmetrie wärmt BHS auch die Durchschnittstemperatur in Wäldern. CLM überschätzt den Tagesgang der LST in Wäldern verglichen mit Satellitendaten deutlich, was durch die Implementation von BHS im Modell merklich vermindert wird. Daneben wird auch die Abweichung des Effekts von Aufforstung auf das Tagesminimum der LST behoben, die ich in Kapitel 2 gefunden habe. Alles in allem verändert BHS den Tagesgang von Temperaturen in Wäldern stark und ist auch relevant für das Tagesmittel der Temperatur, während die Auswirkung auf Tagesmaximum-Temperaturen nur gering ist.

In Kapitel 4 habe ich die Veränderung des Niederschlags durch Aufforstung von Landwirtschaftsflächen untersucht; ein biogeophysikalischer Effekt, der in beobachtungsbasierten Studien bis anhin grösstenteils vernachlässigt wurde. Hierzu habe ich zwei (beinahe vollständig unabhängige) Ansätze verwendet: Erstens habe ich geeignete Paare von Regenmessstationen identifiziert, deren Umgebung sich um mindestens 20 % in den Anteilen an Wäldern und Landwirtschaftsflächen unterscheidet. Zweitens habe ich die Klimatologie eines modernen, räumlich kontinuierlichen Niederschlagsdatensatzes mit verallgemeinerten additiven Modellen (engl. GAMs für "Generalized Additive Models") modelliert, um einen Bezug zwischen der Landbedeckung und dem Niederschlag herzustellen. In beiden Ansätzen scheint Aufforstung den Niederschlag, insbesondere während des Winters, lokal zu erhöhen. Aufgrund der Struktur der GAMs lässt sich ausserdem die Veränderungen im Lee der aufgeforsteten Lokalitäten bestimmen. Im Winter erhöht sich der Niederschlag windabwärts in Süd- und Westeuropa, während der Effekt in Zentraleuropa nahezu neutral und in Nordeuropa negativ ist. Im Sommer erhöht sich der Niederschlag windabwärts von Aufforstung beinahe flächendeckend. Der kombinierte lokale und windabwärts gerichtete Effekt eines realistischen Aufforstungsszenarios kom-

pensiert einen beachtlichen Anteil der sommerlichen Niederschlagsabnahme, welche von einem Ensemble regionaler Klimamodelle gegen Ende dieses Jahrhunderts unter RCP4.5 prognostiziert wird. Während diese Studie darauf hinweist, dass Aufforstung relevante Veränderungen des Niederschlags in Europa hervorruft, muss berücksichtigt werden, dass sie im Forschungsfeld neuartig ist und daher grössere Unsicherheiten aufweist als die beiden vorangehenden Studien.

Zusammengefasst resultiert Aufforstung in biogeophysikalischen Effekten, welche für das lokale und das regionale Klima relevant sind. Solche Effekte sollten beachtet werden, bevor Aufforstung zur Kompensation von Treibhausgasemissionen genutzt wird. Dennoch sind viele Aspekte bezüglich der biogeophysikalischen Effekte durch LULCC in Modellen und Beobachtungen noch unsicher oder gar unbekannt. In dieser Arbeit zeige ich auf, dass BHS für das lokale Klima in Wäldern relevant ist und daher in der nächsten Generation von Erdsystemmodellen aufgenommen werden sollte, welche zur Bestimmung des Einflusses von LULCC auf das Klima herangezogen werden. Ausserdem lege ich mit Beobachtungen dar, dass Aufforstung den Niederschlag in Europa ändern kann. LULCC rufen daher nicht nur Temperaturveränderungen, sondern auch relevante Änderungen des Wasserkreislaufs hervor, welche bei der Abschätzung der klimatischen Folgen von LULCC berücksichtigt werden sollten.

Contents

1

Introduction

The services of land to humans are plentiful: Land not only provides a habitat to humans, but is also crucial for food, fiber, energy, and wood production and serves as a recreational space. It is therefore not surprising that an estimated 69-76 % of the land surfaces is currently directly affected by human land use (IPCC 2019a). This widespread use of land has not occurred without leaving marks at the land surface. More than half of the natural forests have been cleared since 1765, often to be used as cropland or pasture (Meiyappan and Jain, 2012). Land Use and Land Cover change (LULCC) has lead to an estimated release of 210±60 Gt C since pre-industrial times, accounting roughly a third of the total anthropogenic GreenHouse Gas (GHG) emissions (Friedlingstein et al., 2020). While land-use related alterations of the climate system have taken a somewhat secondary role behind the combustion of fossil fuels over the last century (IPCC 2014), land use is still a key component of climate change both on the problem and the solution side. Further, unintentional land degradation for example from soil erosion endangers the services provided to humans by land. 8-45 % of the ice-free land area have been degraded to date (Gibbs and Salmon, 2015), a problem which will most likely become exacerbated by climate change (IPCC 2019e).

Over the last three decades, awareness grew that LULCC), in particular deforestation, not only affects climate through the release or sequestration of greenhouse gases, but also through directly altering the energy and water redistribution at the land surface (e.g., Sud et al., 1996; Betts, 2001; Bonan, 2008; Pitman et al., 2009; Davin and de Noblet-Ducoudré, 2010; Pongratz et al., 2010; Devaraju et al., 2015; Ellison et al., 2017). This later category of consequences is commonly referred to as biogeophysical effects of LULCC, while direct alterations of the atmospheric GreenHouse Gas (GHG) content due to LULCC are called biogeochemical effects. Depending on the climate zone, the biogeophysical effect of deforestation may dominate the biogeochemical one (Claussen et al., 2001; Randerson et al., 2006; Bala et al., 2007; Bathiany et al., 2010). Biogeophysical effects are the result of the alteration of several properties of the land surface, whose importance varies in space and time. This poses a challenge to observational studies and climate models, which has not been fully solved to date (Perugini et al., 2017). This thesis is a contribution to

foster our understanding of the biogeophysical consequences of de-, re-, or afforestation. In particular, I investigate the biogeophysical effect of forestation, defined as afforestation or reforestation of grassland and crops.

In this first chapter, I provide a brief overview of the current state of the research on extent and occurrence of LULCC and its biogeophysical effects. Section 1.1 describes the extent of LULCC in the past and the potential occurrence of LULCC in the future. The next section makes a clear distinction between biogeochemical and biogeophysical effects of LULCC and highlights an important complication with the latter category of effects. Section 1.3 summarizes the most relevant observation-based and model-based approaches to estimate biogeophysical effects of LULCC. In the following section, I embrace the known biogeophysical consequences of forestation. And finally, I concretise the objectives of this thesis and introduce the studies it comprises.

1.1 Occurrence of LULCC and LCC

Before I continue to write about LULCC, I would like to properly define this expressions and some additional related expressions. LULCC encompasses alterations in both Land Use (LU) and Land Cover (LC), for which I follow the definitions of Pongratz et al. (2018). Land use refers to the purpose for which humans use a specific land area, such as forestry or cropping, and how they manage the land to serve this purpose (e.g., tillage or fertilization). On the other hand, LC describes the properties of the land surface and is often categorized into broad land cover classes, such as forest or cropland. Forestation of grassland and crops corresponds therefore always to a LC Change (LCC), but not necessarily to a LU change. As a biogeophysical effect always requires an alteration of the properties of the land surface, I will use the expression LCC in this context. The term LULCC is used for anthropogenic changes in LC or LU in this thesis, excluding LCC that occurs naturally or indirectly through human-induced climate change.

1.1.1 Extent of human LU

An estimated 15.9-18.8 Million km^2 and 22.8-32.8 Million km^2 were used as cropland and pasture in 2015, respectively, corresponding together to roughly a third of the total ice-free land area (IPCC 2019b). LU for food production has a long history. Already more than a millennium ago, humans have used a considerable fraction of the land area, to an uncertain extent however. Pongratz et al. (2008) estimate that 2.8 Million km^2 were used as agricultural land in 800 AD, while Kaplan et al. (2011) report that 18 Million km^2 were affected by LULCC in 950 AD, which corresponds to 12.6 % of the ice-free land. Rapid population growth has led to a strong expansion of agricultural land over the last three centuries (Ramankutty and Foley, 1999; Ma et al., 2020). This expansion has partly come at the expense of forests. Approximately 55 % of the natural forests disappeared between 1765 and 2005 (Meiyappan and Jain, 2012). Such net numbers hide that the LC may change several times. For example, agricultural land may be abandoned to grow a secondary forest (Hurtt et al., 2006; Hurtt et al., 2011; Hurtt et al., 2020). In addition, forests are frequently managed to provide wood as a fuel or construction material. Only

17 % of the European forests are estimated to be unmanaged currently (McGrath et al., 2015). As of today, merely 50 % of the ice-free land surface is classified as low human impact areas (Fig. 1.1; Jacobson et al., 2019) and only 24-31 % as unused land (IPCC 2019b). In summary, humans have altered the LC substantially in many regions of the world, which, as we know today, has affected the climate on earth.

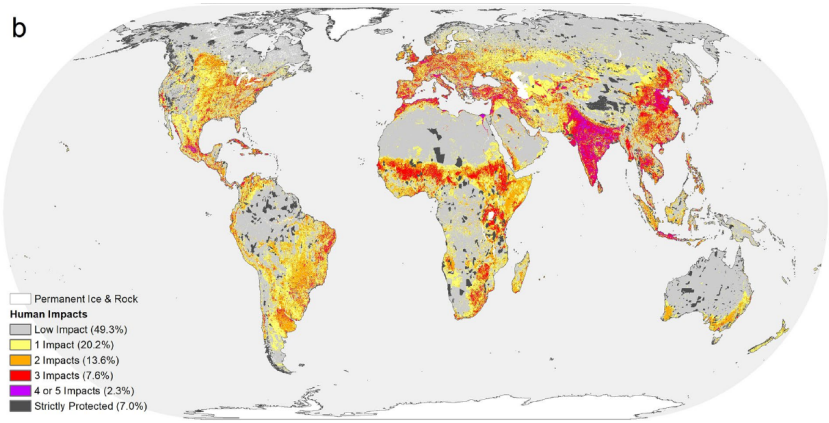


Figure 1.1: *Global map of the intensity of human impacts. The majority of impacted cells have more than one impact. Different impacts are related to the presence of human settlements, the presence of agricultural land, and anthropogenic LCCs at 1 km resolution. From Jacobson et al. (2019).*

1.1.2 Forest-related LCC over the last decades

The loss of primary forests has accelerated drastically over the last three centuries (Hurtt et al., 2020). From 2000 to 2012, forest loss amounted to 2.3 Million km^2 or 2.1 % of the ice-free land (Fig. 1.2; Hansen et al., 2013). Brazil alone lost more than 40'000 km^2 of forest in 2003, corresponding to roughly the area of Switzerland. While many tropical and sub-tropical regions have experienced fast rates of deforestation in the recent past, agricultural land is increasingly abandoned in industrialized countries, often to regrow forests. The deforestation mentioned above was therefore partly compensated by 0.8 Million km^2 forest gain from 2000 to 2012 (Fig. 1.2; Hansen et al., 2013). In the EU countries alone, the total forest area increased by 0.3 Million km^2 with roughly the same loss of cropland between 1950 and 2010 (Fuchs et al., 2013; Fuchs et al., 2014; Fuchs et al., 2015). China as another example has launched several afforestation programs (Zhang et al., 2016), which promoted a 0.3 Million km^2 increase in the forest area from 1999 to 2013 (Chen et al., 2019).

Forest loss can also result from natural disturbances, which become more likely due to anthropogenic climate change. Widespread disturbances by forest fires and bark beetle outbreaks were more frequent during the last two decades than previously (Safranyik et al., 2010; Ghimire et al., 2012; Hicke et al., 2015). Between 2000 and 2010, 0.18 Million

km^2 of forests were affected by bark beetles in western North America alone (Edburg et al., 2012), resulting in detectable alterations of surface albedo and evapotranspiration (Maness et al., 2013; Vanderhoof et al., 2013; Vanderhoof and Williams, 2015). Overall, forest-related LCCs were a wide spread phenomena over the last decades across the world, which will likely persist into the future.

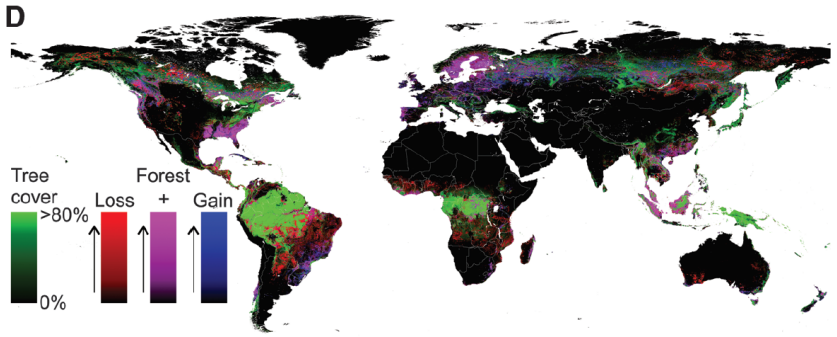


Figure 1.2: *Forested areas with no change in forest cover (green), forest loss (red), forest gain (blue), and concurrent loss and gain (purple) between 2000 and 2012. From Hansen et al. (2013).*

1.1.3 Future LULCC

The human population on Earth is growing rapidly at an annual growth rate of 1.2 % in 2015 and expected to reach 9.73 billion by 2050 (FAO 2017). The combination of a growing population and dietary shifts due to increasing wealth is expected to increase the demand for agricultural production by roughly 50 % in 2050 compared to 2013. Most likely, such an increase in food production cannot be met solely by intensification, creating a demand for more agricultural land, potentially at the expense of forests.

In 2016, representatives from 196 countries agreed to confine the global mean temperature to well below 2 K above pre-industrial levels by signing the Paris Agreement. As such, anthropogenic climate change has created a novel form of LU: land-based mitigation of greenhouse gas emissions. A number of land-based strategies to sequester GHGs were proposed, some of which exhibit a high potential (Fig. 1.3). However, the spread of the estimated potential for GHG mitigation is often substantial. Climate change scenarios that result in the lower range of global warming levels typically include some form of forestation (IPCC 2018; IPCC 2019e), as the forest biomass is a natural carbon reservoir. Griscom et al. (2017a) estimate that 10 $GtCO_2$ could be removed annually by reforestation, corresponding to roughly a quarter of the present-day GHG emissions, without harming security safeguards for food and fiber production. However, this value is at the higher range compared with the potential reported by the IPCC, which amounts to 0.5-10.2 $GtCO_2$/yr (Fig. 1.3; IPCC 2019d).

But land-based solutions go beyond simply reverting the deforestation of the past.

Bio-Energy with Carbon Capture and Storage (BECCS) is a component of many low GHG concentration scenarios (Azar et al., 2010; IPCC, 2018; IPCC, 2019c). The inclusion of trees on agricultural land in agroforestry systems is another promising tool to sequester carbon from the atmosphere (Zomer Robert J. et al., 2016; Swamy and Tewari, 2018). Besides the sequestration of GHGs, LU can also be optimized to mitigate climate change through biogeophysical mechanisms. For example, the albedo of the land surface could be increased to reduce the amount of absorbed shortwave radiation (Seneviratne et al., 2018). In summary, the way humans utilize land will be a central factor for the future development of the climate. It is therefore pivotal to understand all relevant climatic consequences of LULCC.

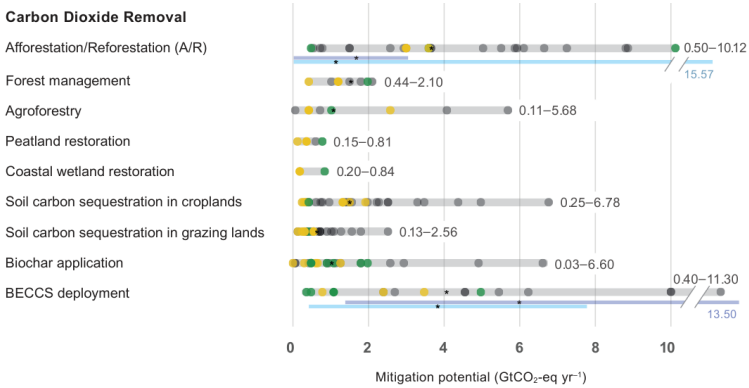


Figure 1.3: *Overview of estimated greenhouse gas mitigation potential of various land-based LULCC strategies in 2020–2050, measured in $GtCO_2$-eq/yr. Only studies that were published after 2010 were considered. Purple and blue bars show the range from integrated assessment models for the 1.5° and 2° scenarios, respectively. From IPCC (2019c).*

1.2 Definition and characterization of biogeophysical effects

The climatic consequences of LULCC are often divided into biogeochemical and biogeophysical effects. The first category of effects refers to the release or sequestration of greenhouse gases such as CO_2, CH_4, or N_2O by the land surface due to LULCC (I use the term LULCC here, because a specific LU, such as BECCS can affect atmospheric greenhouse gas concentrations without any LCC). Such gases are distributed relatively efficiently within the atmosphere and can therefore be directly added to other anthropogenic greenhouse gas emissions such as fossil fuel combustion. The climatic of the biogeochemical effects of LULCC can therefore be deduced from the overall sensitivity of the climate to anthropogenic GHG emissions. The challenge remaining to the scientific community is to quantify the release or sequestration of GHG by LULCC. This has been done through measuring fluxes of chemical species at the land surface and upscaling them (e.g., Jung et al., 2011), estimating the biomass within different biomes and therefrom the

amount of the chemical species stored within them (e.g. Tyukavina et al., 2015), or using earth system models, which represent chemical cycles (e.g., Bala et al., 2007; Pongratz et al., 2009; Pongratz et al., 2010).

The term biogeophysical refers to all naturally occurring objects and processes of biological, geological, or physical nature operating in an area of interest (USDA, 1976). In the context of LCC it refers to any alterations of the water and energy redistribution at the land surface due to LCC and their potential local and remote consequences. Prominent biogeophysical effects of LCC include changes in albedo, the evaporative fraction, and the surface roughness, which will be examined in more detail in Section 1.4. Fig. 1.4 (a) illustrates biogeophysical effects due to a LCC at one specific location. This LCC will affect the local conditions directly by altering the local energy and water redistribution at the land surface (e.g., a decrease in albedo will result in the absorption of more shortwave radiation and consequently in an increase of the LST). Such local biogeophysical effects are called ΔX^{l} throughout this thesis, where X can be any variable describing the local conditions and ΔX corresponds to a temporal change in X caused by LCC. ΔX^{l} depends on the atmospheric and radiative conditions, which change with location on earth and time (e.g., ΔX^{l} of a change in albedo depends strongly on the amount of incoming shortwave radiation). Thus the biogeophysical effect of one specific LCC varies spatially and temporally. Further, the LCC will alter the atmospheric conditions, which will in term affect the local climate, thereby potentially generating a feedback from the atmosphere. Such alterations of the atmospheric conditions can propagate for example through advection to locations in the proximity (i.e., locations where no LCC has occurred), which may affect the climate at those locations. For example, an albedo decrease due LCC might warm atmospheric temperatures above the location of LCC resulting both in locally increased convective activity and increased temperatures in proximity. Thus, a LCC can trigger remote biogeophysical effects, ΔX^{rem}, which should ideally be quantified alongside ΔX^{l}.

Even for the case of one isolated LCC, the quantification of ΔX^{l} and ΔX^{rem} based on modeling or observations is not trivial, as the climate system is evolving constantly due to other factors such as internal variability or climate change. In reality, LCCs occur concurrently at different locations on earth, making the scenario in Fig. 1.4 (a) somewhat synthetic. This evokes a further complication, which is exemplified in Fig. 1.4 (b) for two LCCs that occur at the same time. The effect at the location of LCC_1 is now influenced by LCC_2 and vice versa through remote effects. Hence, any LCC-induced changes in the local conditions at locations that underwent LCC is no more attributed unequivocally to the LCC at the respective location. In Fig. 1.4 (b) the overall effect is therefore denoted as ΔX^{tot} instead of ΔX^{l}. Similarly, ΔX^{rem} is now the result of both LCCs, complicating its attribution to one specific LCC. It is unsure to which extent the biogeophysical effects from different LCCs are additive (e.g., whether $\Delta X^{l} + \Delta X^{rem} = \Delta X^{tot}$ holds). The superposition of biogeophysical effects from various LCCs therefore further complicates their quantification.

An ideal method to estimate the biogeophysical effect from the LCC of interest would exhibit the following features: (1) It can distinguish biogeophysical effects due to this LCC from the background variability of the climate. This includes the distinction from

biogeophysical effects induced by LCCs at locations other than the location of interest. (2) It can estimate the biogeophysical effect of this LCC at all locations where it is relevant. (3) It quantifies both ΔX^{l} and ΔX^{rem}. (4) It captures temporal variations in the biogeophysical effect of this LCC. A step beyond a method that fulfills these criteria would be to investigate the superposition of the biogeophysical effects from various LCCs. Observation-based and model-based studies face different challenges and employ therefore also differing strategies to quantify biogeophysical effects of LCC, which will be described in the next section.

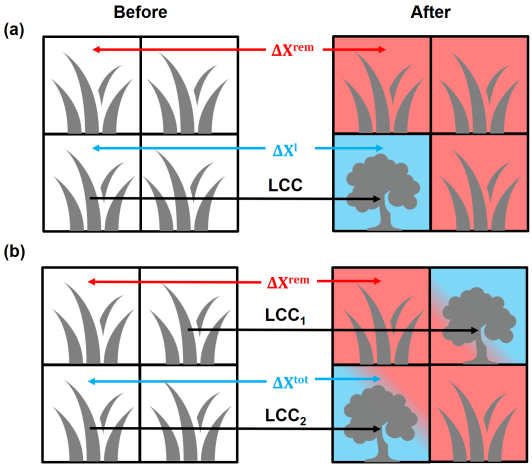


Figure 1.4: ***Graphical illustration of biogeophysical effects due to LCC (forestation here).*** *On the left, state prior to occurrence of LCC (world of four grassland locations) and to the right, state after LCC. Panel (a), local (ΔX^{l}) and remote (ΔX^{rem}) biogeophysical effect due to alteration of the LC at one location (lower left box). In red, boxes that are affected by remote biogeophysical effects and in blue, boxes that are affected by local biogeophysical effects. Panel (b), biogeophysical effect due to LCC at two locations (lower left and upper right boxes) at locations where LCC occurred (ΔX^{tot}) and were LCC did not occur (ΔX^{rem}). Note that there are now both local and remote effects at the locations with LCC. ΔX refers to a temporal change in any variable X that might be altered through biogeophysical processes (e.g. LST).*

1.3 Methodological approaches to estimate biogeophysical effects of LCC

As described in Section 1.1, LCC have not only been a widespread phenomena in the past, but are still ongoing and will most likely also occur in the future, be it as an adaptation or mitigation tool for climate change or just to satisfy the increasing demand of land for food production due to the growing world population and dietary shifts. The quantification of the biogeophysical effects of LCC is therefore required both by the scientific community

to understand the climate of the past and more importantly by decision makers to reach well-informed decisions. Ultimately, decision makers need to know ΔX^{tot} and ΔX^{rem} of the LCC(s) under consideration. Further, knowledge of ΔX^{l} and ΔX^{rem} of any relevant LCC at all relevant locations in isolation would be useful to develop LCC scenarios. Unfortunately, observation-based and model-based approaches both face difficulties to infer this information directly. This section provides an overview of the methodological approaches present in the scientific literature to date, including their advantages and disadvantages. The most important points are summarized in Table 1.1.

1.3.1 Observation-based approaches

Observation-based approaches to quantify biogeophysical effects of LCC have to overcome two main challenges. Firstly, the execution of a controlled experiment, where a land area of sufficient extent would be converted just for scientific purposes, is costly and hence in most cases unfeasible. This issue is aggravated by the fact that properties of vegetation and their biogeophysical effects change with climate. A large number of experiments would be required to properly quantify the biogeophysical effects of all relevant LCC at all locations of interest. Therefore, observation based studies have to rely either on existing distribution of the LC and/or the LCC that occur anyway. Secondly, the climate system evolves constantly through natural variability and anthropogenic drivers other than biogeophysical effects of LCC. This impedes the distinction of the biogeophysical response to a LCC from the noise of the background climate.

The so-called space-for-time analogy is an often-used strategy to alleviate both of the aforementioned problems. The basic principle of this approach is to compare the measured variables over two different LCs that are located in vicinity to each other, under the assumption that they receive the same atmospheric forcing. Given this assumption holds, the difference in any variable between the two LCs would correspond to the biogeophysical effect induced by a conversion between those two LCs, as any other temporal trend in atmospheric conditions should be the same at the two locations. A slight modification of this is to compare locations that underwent LCC to locations where no LCC occurred during the same period. This latter approach provides a clearer causal link between the occurring LCC and the change in climate, particularly if it can be demonstrated that the climate at two locations was comparable prior to the LCC. However, samples are more scarce than in a spatial comparison of different LCs, because continuous measurements are required at locations that undergo LCC.

Up to present, the space-for-time analogy was applied to two categories of measurements. The first studies providing an observational constraint of the biogeophysical effects of LCC compare in-situ measurements at close-by locations of differing LC. Typically, these studies employ measurements from FLUXNET sites (da Rocha et al., 2004; von Randow et al., 2004; Liu et al., 2005; Juang et al., 2007; Teuling et al., 2010; Lee et al., 2011; Luyssaert et al., 2014; Zhang et al., 2014; Vanden Broucke et al., 2015; Chen et al., 2018; Cherubini et al., 2018b; Tang et al., 2018; de Oliveira et al., 2019). FLUXNET sites provide a comprehensive list of measured variables, facilitating the physical understanding of the underlying mechanism of observed temperature differences. The study of Lee et al. (2011) is a prominent example of this category of studies in the

context of the biogeophysical effect of deforestation, investigating a collection of paired sites in North America rather than a single site pair (Fig. 1.5). An important limitation of in-situ studies is the limited amount of available site pairs. Noteworthy, Bright et al. (2017) upscaled FLUXNET measurements of specific LC types to global scale through empirical relationships to retrieve estimates of the biogeophysical effect of several LCCs with global coverage.

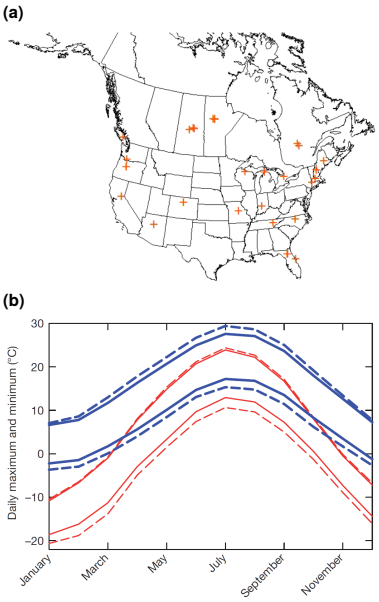


Figure 1.5: *Location of the forest/open land site pairs presented in Lee et al. (2011) (a). Panel (b) shows the mean across sites in daily maximum and minimum temperatures for the forest (solid lines) and the open land sites (dotted lines) located south of 45° N (blue lines) and north of 45° N 56° N (red lines). From Lee et al. (2011).*

The emergence of remote sensing measurements that cover several years has opened a second avenue of observation-based studies on the biogeophysical effects of LCC. These studies typically employ a moving window approach, where grid cells of differing LC or differing LC conversion rates are contrasted within moving windows of a certain size (Fig. 1.6; Peng et al., 2014; Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016b; Bright et al., 2017; Schultz et al., 2017; Duveiller et al., 2018b; Tang et al., 2018; Chen and Dirmeyer, 2019b; Forzieri et al., 2020; Zeppetello et al., 2020). However, there exist slight variations in the strategy to find spatial comparison samples (e.g., Ge et al., 2019). Since the employed remote sensing products typically cover the entire globe, the spatial coverage of these studies is much better than for the paired in-situ site studies (Fig. 1.7). However, the variety of variables that are measurable from space is limited. The Land Surface Temperature (LST) and the albedo can be estimated at relatively high certainty by measuring the radiation emitted by the land surface and the shortwave fluxes at the land surface, respectively. On the other hand, the 2 m air temperature (T2M), the sensible heat flux, and the latent heat flux can not be measured easily from space, as they are not directly linked to radiative fluxes. The inability to measure T2M could be problematic as modeling studies suggest that it might respond differently to deforestation than LST (Chen and Dirmeyer, 2019a; Winckler et al., 2019b; Breil et al., 2020). Further, the uncertainty in the turbulent heat flux estimates impedes the understanding of the processes behind the estimates of the LST response to LCC. The moving window approach results in typical window sizes of 0.25–0.5°, which can impede the assumption of the same atmospheric forcing within the windows in particular in regions with complex topography. More advanced statistical methods, such as generalized additive models, that allow to incorporate climatic drivers other than LC are a promising tool to alleviate this problem (e.g.; Schwaab et al., 2020), which will be used in Chapter 4.

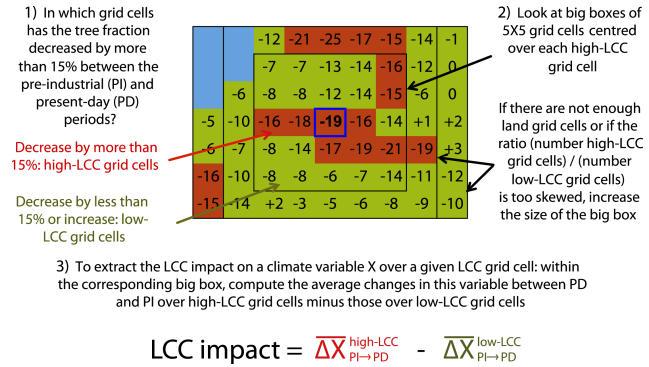


Figure 1.6: *Graphical illustration of the moving window approach to reconstruct the biogeophysical response to deforestation at the blue-edged grid cell. The numbers indicate the change in tree fraction between the preindustrial and present-day periods in each grid cell. Red grid cells are high-LCC grid cells, in which the tree fraction has decreased by more than 15 %, while green grid cells are low-LCC grid cells, in which the tree fraction has decreased by less than 15 % or increased. Light blue grid cells are ocean or lake grid cells. Note that the nomenclature in this figure is different from the rest of this thesis. "LCC impact" corresponds to ΔX^{l}, while $\overline{\Delta X}$ in this figure stands for a temporal change in the average of variable X rather than the biogeophysical response of variable X to LCC. From Lejeune et al. (2017).*

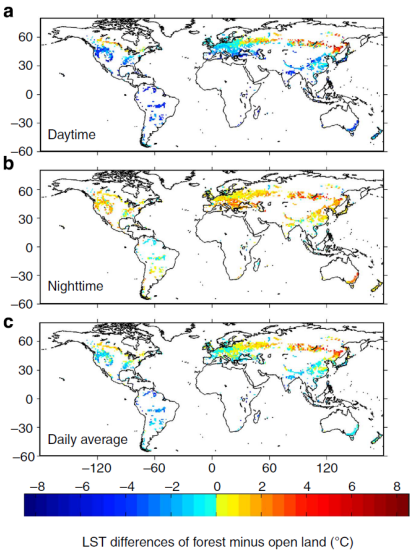


Figure 1.7: *Annual mean LST difference of forest minus open land (grassland and crops) at 13:30 (a), 1:30 (b), and daily average (c). From Li et al. (2015).*

A major caveat of the space-for-time analogy is its inability to properly disentangle direct local biogeophysical effects from the indirect ones that act through the alterations of atmospheric conditions (Chen and Dirmeyer, 2020). When comparing locations of differing LC or LC conversion rates, measurements at one location will be influenced to some extent by the LC at other locations (ΔX^{rem}). This approach therefore determines in fact ΔX^{tot} - ΔX^{rem}, which is likely a good approximation of ΔX^{l}. However, it is difficult to assess from observations alone whether this holds, as only few studies have attempted to quantify ΔX^{rem}. Cohn et al. (2019) showed that nearby forest loss increases daily maximum air temperatures over cropland in Brazil. Also, the amount of precipitation in tropical region increases when the respective air mass has overpassed more vegetation (Spracklen et al., 2012). These studies imply that ΔX^{rem} is indeed not negligible and should therefore be quantified alongside local effects.

1.3.2 Model-based approaches

Earth System Models (ESMs)* have two major advantages over the real world in terms of quantifying biogeophysical effects: First, the alteration of LC in the modeling world comes at virtually no cost (although model simulations can of course still require high amounts of computational resources). Second, the implied LCC occurs only in the modeling world and does therefore not affect the services provided by the land in the real world. For example, the biogeophysical effect of foresting all agricultural land on earth can be estimated in ESMs without financing an exorbitant re- and afforestation program and without starving the world population. However, the land surface is highly heterogeneous and a vast number of biological, physical, and chemical processes occur at the land surface at the same time. As a consequence, land surface models (LSMs) are not the numerical implementation of a well-defined set of physical laws, but rely heavily on parametrisations. Hence, the biogeophysical effect of LCC produced in ESMs needs to be verified carefully with observations before it can be trusted. Unsurprisingly, numerous studies have highlighted strong discrepancies among ESMs and/or between ESMs and observations (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Lejeune et al., 2017; Davin et al., 2020).

In total there exist four widely-used approaches to assess biogeophysical effects of LCC with models: (1) The factorial experiment, (2) the reconstruction method, (3) the sub-grid comparison, which will is employed in Chapters 2 and 3, and (4) the chessboard approach. The most intuitive and therefore first-emerging method is the factorial experiment, which compares two simulations of the same set up except for the LC (Oleson et al., 2004; Bala et al., 2007; Davin et al., 2007; Pongratz et al., 2010; de Noblet-Ducoudré et al., 2012; Tölle et al., 2014; Lejeune et al., 2015; Malyshev et al., 2015; Vanden Broucke et al., 2015; Devaraju et al., 2015; Li et al., 2016a; Naudts et al., 2016; Thiery et al., 2017; Cherubini et al., 2018a; Devaraju et al., 2018; Li et al., 2018; Luyssaert et al., 2018; Skinner et al., 2018; Tölle et al., 2018; Chen and Dirmeyer, 2019b; Ge et al., 2019; Hauser et al., 2019; Hu et al., 2019; Laguë et al., 2019; Davin et al., 2020; Devanand et al., 2020; Kim et al., 2020). The difference in any variable between the two simulations can then be attributed to the LC difference, assuming the simulations were of sufficient length to average out internal variability of the climate system. The retrieved signal corresponds to ΔX^{tot} at locations where LCC occurred and ΔX^{rem} elsewhere, the information required by decision makers. However, a separate simulation for each grid cell would be required to determine the biogeophysical effect of a given LCC at a specific location in isolation, which is not feasible. This impairs the comparison to observational studies, which only estimate ΔX^{l} in most cases. Such experiments are therefore typically used to assess the effect of larger-scale LCC, such as the total effect of deforestation since pre-industrial times. However, a large spread appeared in the first multi-model intercomparison projects that assessed the biogeophysical effect of historical LCC (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013). Therefore, new approaches to isolate ΔX^{l} were required to confront models with observations.

*For simplicity, I use the term earth system model in this thesis even though a general circulation model could in principle be sufficient to assess biogeophysical effects of LCC.

Kumar et al. (2013) introduced the moving window approach, which was subsequently applied frequently to remote sensing data, as described in the previous section. The first studies employing this method contrasted grid cells with LCC above a certain threshold to grid cells with LCC below this threshold in historical simulations (Kumar et al., 2013; Lejeune et al., 2017). Later, multiple linear regression models were used to derive relations between the LCC fraction (Lejeune et al., 2018) or the LC fraction (Lejeune et al., 2020) and response variables of interest. An obvious advantage of the moving window approach over factorial experiments is that it requires only one model simulation to isolate a signal at different locations on earth. Further, studies using this approach are comparable to observational studies that employ the same method. Naturally, such modeling studies suffer from similar disadvantages as their observational counterparts. The assumption of the same atmospheric forcing is undermined even more in modeling studies, as they typically have coarser resolutions of 1–2.5° than remote sensing data, resulting in moving window sizes of 10–20°. Also, they do not quantify remote effects in isolation, which could however be alleviated by combining the moving window approach with a factorial experiment: ΔX^{rem} corresponds to the difference in X between a simulation with LCC and a simulation with constant LC at locations where there was no LCC in both simulations. ΔX^{1} on the other hand can be estimated by applying the moving window approach to the simulation with LCC.

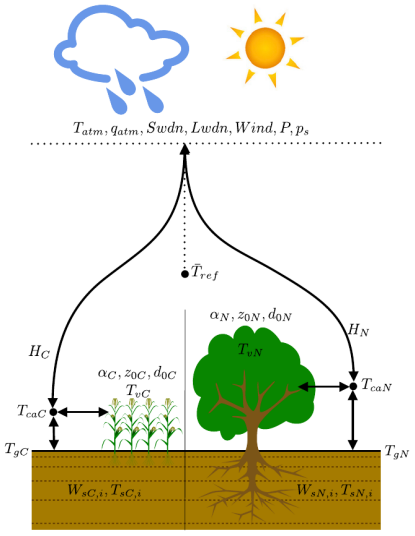


Figure 1.8: *Sketch of a cropland and a forest tile within one grid cell. Model output at tile level (e.g., T_{gC} and T_{gN}) can be used to compare different LC types under the same atmospheric forcing. From Malyshev et al. (2015).*

Due to the relatively high heterogeneity of the land surface in comparison to the atmosphere, a number of ESMs divide the land grid cells into tiles of different LCs, which are all forced by the same atmospheric grid cell above. Thus, model output at tile level allows to compare the different LCs within a grid cell under the same atmospheric forcing, which is called the sub-grid signal in this thesis (Fig. 1.8; Malyshev et al., 2015; Schultz et al., 2016). Since the different LCs within a grid cell share the same atmospheric forcing, such experiments can be conducted in land surface-only simulations, where the atmospheric forcing is prescribed from observations (sometimes referred to as offline simulations). A pseudo sub-grid signal can be retrieved by conducting factorial experiments in offline simulations, for models that do not use a tile approach (Fatichi et al., 2014; Duveiller et al., 2018a; Laguë et al., 2019). Model experiments that prescribe the atmospheric forcing are computationally much cheaper than experiments

that simulate the atmosphere. Also, the retrieved LCC signal at one location is not blurred by LCC in other grid cells and therefore comparable to observational studies that assume a comparison of different LCs under the same atmospheric forcing. However, this approach completely neglects any feedbacks from the atmosphere be it locally or remotely and therefore misses a potentially important part of the signal (see Section 1.4.5). On the other hand, the neglect of atmospheric feedbacks simplifies the process understanding somewhat. In Malyshev et al. (2015), the sub-grid signal of deforestation was somewhat stronger in most cases than the factorial signal, indicating that atmospheric feedbacks often form a negative feedback to the initial biogeophysical response of the land surface to LCC. The introduction of proxy LC classes that cover only a small fraction of each grid cell can be used further to distinguish changes in atmospheric conditions from the local signal of LCC (Chen and Dirmeyer, 2020). Since the proxy LC covers only a small fraction of a grid cell, it can be assumed that its presence has a negligible influence on the local climate. The sub-grid output from this LC class can therefore be used as a virtual measurement station in the modeling world that traces changes in atmospheric conditions.

Finally, Winckler et al. (2017) introduced the chessboard approach, which is able to determine all components of the biogeophysical effects of LCC (Robertson, 2019; Winckler et al., 2019c; Winckler et al., 2019b; Winckler et al., 2019a). This approach uses a factorial experiment, in which the LC is perturbed in a regular, chessboard-like pattern (Fig. 1.9). As a result, such an experiment retrieves a regular pattern of ΔX^{rem} at the grid cells with no LCC, which can be interpolated to retrieve a spatially continuous field of ΔX^{rem}. Correspondingly, the interpolated signal at the grid cells with LCC corresponds to ΔX^{tot}. Finally, ΔX^{l} is estimated as the difference between the fields of ΔX^{tot} and ΔX^{rem}. Unlike the methods presented previously, the chessboard approach determines all components of the biogeophysical response to LCC. However, ΔX^{rem} depends on the extent of the implied LCC (Winckler et al., 2017). The estimated ΔX^{rem} is therefore specific to the chosen LCC pattern, which is somewhat synthetic in the case of the chessboard approach. In addition, the method assumes additivity of ΔX^{rem} and ΔX^{l} to ΔX^{tot}, which is yet to be confirmed. Further, the spatial interpolation of ΔX^{tot} and ΔX^{rem} may introduce some errors, which appeared to be of non-negligible magnitude in particular for ΔX^{rem} in Winckler et al. (2017).

Overall, each model-based method has its strengths and limitations. Which method is appropriate therefore depends on the research question. The factorial experiment is best suited to provide policy relevant information about the biogeophysical implications of one specific LCC distribution. The reconstruction method and the sub-grid approach are best suited for comparison to observational studies. The chessboard approach provides the most comprehensive isolation of ΔX^{tot}, ΔX^{rem}, and ΔX^{l} and can also be used for comparison to observations (Winckler et al., 2017; Robertson, 2019). However, it is restricted to idealized experiments and estimates ΔX^{l} only from the difference in ΔX^{tot} and ΔX^{rem}.

Table 1.1: *Overview over the different observation-based and model-based approaches to estimate the biogeophysical effects of LCC. Third column indicates whether method provides information on the biogeophysical effect of LCC at one specific location.*

	Method	Components	Location specific	Restrictions	Other advantages	Other disadvantages
Observations	Paired in-situ	ΔX^{l}	Yes	Restricted to available site pairs	• Possible to estimate most of the relevant variables • Same climate assumption reasonable	• Small sample size
	Remote sensing	ΔX^{l}	Yes	Spatial heterogeneity of LC(C) required	• Decent spatial coverage	• Limited number of observable variables • Same climate assumption might be problematic
Models	Factorial	X^{tot} where LCC X^{rem} otherwise	No (computationally expensive)	Everything possible	• Provides information required by decision makers	• Not directly comparable to observations
	Reconstruction	ΔX^{l}	Yes	Spatial heterogeneity of LC(C) required	• Requires only one simulation • Directly comparable to observations	• Same climate assumptions likely harmed
	Sub-grid	ΔX^{l}	Yes	Everything possible	• Computationally cheap • Process understanding relatively easy	• Neglects atmospheric feedbacks completely
	Chessboard	ΔX^{tot}, ΔX^{rem}, and ΔX^{l}	ΔX^{l} yes, ΔX^{rem} and ΔX^{tot} no	Restricted to idealized experiments	• Estimates all components	• May introduce errors from interpolation and assumption of additivity

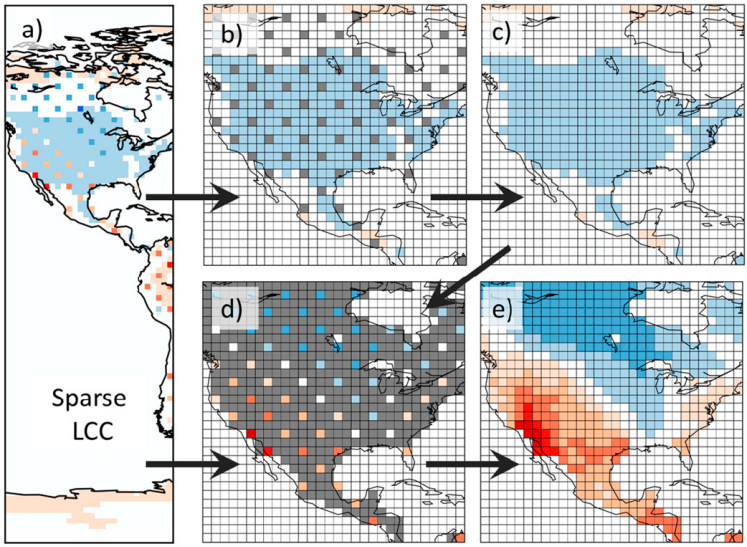


Figure 1.9: *Example of the chessboard approach. Panel (a) shows the response to deforesting one out of eight grid cells. The signal at grid cells with no deforestation (b) is interpolated to retrieve a spatially continuous field of ΔX^{rem} (c). The interpolation of the signal at grid cells with deforestation (d) provides ΔX^{tot} (e). ΔX^{l} is estimated as the difference between ΔX^{tot} and ΔX^{rem} (panel e minus panel c). From Winckler et al. (2017).*

1.4 Known biogeophysical effects of forestation

In this thesis, I focus on the biogeophysical effect of foresting open land, which I simply call forestation in the following, where open land corresponds to grassland and cropland. This section provides an overview of the known biogeophysical effect of this LCC. From here on, ΔX corresponds to the difference in variable X between forest minus open land, where X can be the albedo (α), EvapoTranspiration (ET), LST, T2M, or precipitation (P), with the superscripts 'l', 'rem', and 'tot' for the local, remote, and total difference in those variables, respectively. In this thesis, the LST is defined as the temperature at which the land surface as a whole (i.e., not distinguishing between the soil and vegetation) emits radiation according to the Stefan-Boltzmann law:

$$LST = \sqrt[4]{J/\epsilon\sigma} \tag{1.1}$$

,where J is the radiation emitted by the land surface, ϵ its emissivity, and σ the Stefan-Boltzmann constant (5.67×10^{-8} $\mathrm{Wm^{-2}K^{-4}}$). Forests differ in a number of physical properties from agricultural land. The nature of these differences as well as their importance varies in space and time. In consequence, the overall biogeophysical effect of forestation is the result of a complex superposition of various biogeophysical effects. A thorough scientific comprehension of the overall biogeophysical effect of LCC requires

to attribute parts of the effect to specific altered properties of the land surface. Modeling studies can achieve this by manipulating one property at the time (Davin and de Noblet-Ducoudré, 2010; Burakowski et al., 2018; Laguë et al., 2019; Winckler et al., 2019a). In observations, the difference in a specific property between different LCs can be measured at varying degrees of certainty. However, two kinds of LC rarely differ only by one property impeding a direct casual link of the contribution of individual properties to the overall biogeophysical effect. For the biogeophysical LST difference between LCs, (ΔLST^1) various decomposition methods have emerged to establish such a link. The energy balance decomposition attributes ΔLST^1 to differences in the energy fluxes at the land surface between the LCs of interest by solving the total derivative of the energy balance for the tendency in LST (Juang et al., 2007; Luyssaert et al., 2014; Vanden Broucke et al., 2015). While being able to complete the overall LST difference well, this method does still not attribute the LST difference to specific properties of the land surface, but to differences in various energy fluxes. Therefore, two other techniques have been proposed to attribute ΔLST^1 to differences in surface properties: The intrinsic biophysical mechanism method (Lee et al., 2011; Chen and Dirmeyer, 2016; Burakowski et al., 2018; Ge et al., 2019) and the two-resistance mechanism method (Rigden and Li, 2017; Liao et al., 2018; Wang et al., 2020). A description of those two methods is not provided here, but can be found for example in Liao et al. (2018).

1.4.1 Albedo effect

In most cases, forests tend to have a lower albedo than grassland or crops, which is probably the best constrained biogeophysical effect of forestation. In consequence, forests absorb more of the incoming radiation thereby warming the land surface. This biogeophysical effect was observed in numerous studies using in-situ observations (Juang et al., 2007; Luyssaert et al., 2014; Vanden Broucke et al., 2015) and satellite data (Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b). These studies revealed that the albedo difference between forests and open land ($\Delta\alpha^1$) is particularly distinct under the presence of snow, as forests partly cover the underlying snow surface while shorter vegetation is buried more easily by snow (Fig. 1.10; Harding et al., 2001). Most models are able to capture the albedo decrease following forestation (Bonan et al., 1992; de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Lejeune et al., 2017; Davin et al., 2020), which is not surprising given the qualitative robustness of this effect in observations. Still, models exhibit a large spread regarding the magnitude of the albedo change from forestation, which is relevant for the estimated radiative forcing due to LULCC of the past (Lejeune et al., 2020).

Figure 1.10: *Example of the snow masking by trees. Areas covered by forest appear much darker than open areas. Photo by Balazs Busznyak on Unsplash.*

1.4.2 Roughness effect

A forest consists of relatively large objects with an often considerable variability in height compared to open land. In the field, the amount of friction generated by a specific land surface is often characterized by the surface roughness. Following this terminology, forests have normally a higher surface roughness then open land. In consequence, they exert more friction at the land surface and induce more turbulence in the surface layer. The latter effect facilitates the turbulent exchange of energy, water vapour, and other chemical species between the land surface and the overlying atmosphere. Several modeling studies assessed the role of the higher surface roughness of forests for the biogeophysical effect of forestation by manipulating the surface roughness in isolation (Davin and de Noblet-Ducoudré, 2010; Laguë et al., 2019; Winckler et al., 2019a). In general, the increase in surface roughness cools the land surface, as the solar radiation absorbed by land can be transported away to the atmosphere through the sensible and latent heat fluxes more efficiently (Fig. 1.11). In observations, the relevance of the surface roughness for the observed LST change following forestation is more difficult to assess, as it cannot be manipulated in isolation. A few studies using in-situ observations applied either the intrinsic biophysical mechanism method and/or the two-resistance mechanism method, confirming the daily mean cooling due to increased surface roughness (Lee et al., 2011; Burakowski et al., 2018). Further, the higher surface roughness of forests increases nighttime LST, as the sensible heat flux is normally directed towards the land surface during night. Thus, the surface roughness effect of forestation exhibits a distinct diurnal asymmetry resulting in a dampening of the diurnal LST cycle in forests, which was found both in modeling studies (Winckler et al., 2019a) and observations (Chen and Dirmeyer, 2016; Schultz et al., 2017; Burakowski et al., 2018; Liao et al., 2018).

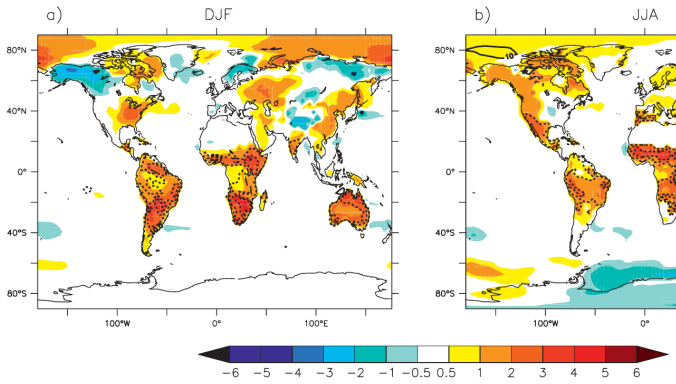


Figure 1.11: *Change in surface temperature [K] (colors) and in turbulent heat flux (sum of latent and sensible heat; contours every* $10\,\mathrm{W\,m^{-2}}$*, with dashed lines for negative values) due to the surface roughness change associated with converting forests to grassland in the IPSL model. From Davin and de Noblet-Ducoudré (2010).*

1.4.3 Surface resistance to evapotranspiration

Differences between forests and open land in physiological properties, such as the root distribution or stomatal regulation, influence how much water is transpired by those LC types and therefore how the available energy at the land surface is partitioned into sensible and latent heating. In consequence, the physiological properties of forests and open land exert a strong control on ΔET[1]. However, a difference in ET between forest and open land cannot be directly attributed to physiological differences, because ET is also affected by the amount of available energy and the atmospheric resistance to turbulent transport. As such, both the albedo effect and the roughness effect of forestation influence indirectly ΔET[1]. In addition, different plant species within the LC types forest and open land exhibit a wide variety of physiological properties and strategies to cope with potential soil moisture stress (e.g., an evergreen needleleaf tree in Scandinavia will react much differently to soil moisture stress than a rain green broadleaf tree in the Sahel region). Unsurprisingly, there are therefore no consistent results regarding the change in surface resistance to ET due to forestation.

Observations generally indicate that ET is higher over forests compared to open land in the tropics and during the warm season outside of the tropics (von Randow et al., 2004; Liu et al., 2005; Vanden Broucke et al., 2015; Peng et al., 2014; Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b). However, it is unsure how much of this difference emerges actually from differing water use strategies and how much is a side effect of the biogeophysical differences described in the previous two sections. Unlike nearby pasture, forests in the tropics were observed to sustain high amounts of ET during dry periods by extracting water from below 2 m through their deeper roots (Fig. 1.12; Jipp et al., 1998; von Randow et al., 2004). On the other hand, grasslands were observed to evaporate more during the early stages of a summer heat wave in Europe (Teuling et al., 2010; Van Heerwarden and Teuling, 2014). Also, Williams et al. (2012a) find that grasslands evaporate a larger fraction of the incoming precipitation than forested landscapes when applying the Budyko framework to the La Thuile FLUXNET data collection. The intrinsic biophysical mechanism and the two-resistance mechanism methods attribute LST temperature changes from forestation to changes in the bowen-ratio (ratio of sensible and latent heat flux) and the surface resistance, respectively, thereby providing a term more directly linked to the plant physiological control on transpiration. However, studies applying these methods in a comparison of forests and open land find generally no strong and consistent contribution of these terms

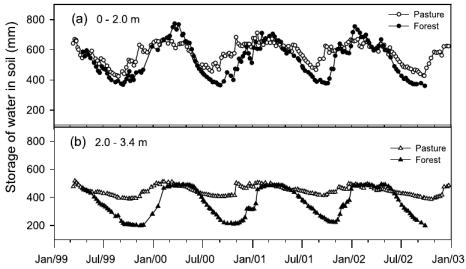


Figure 1.12: *Example of soil moisture access from deep soils at a forest site. Panel (a), water stored in soil up to a depth of 2 m (circles) and panel (b), water stored between 2 and 3.4 m depth (triangles) at a pasture site (empty symbols) and a forest site (filled symbols). Note that the water stored in the lower part of the soil decreases much stronger during the dry periods for the forest site than for the pasture site. From von Randow et al. (2004).*

to the overall LST difference, again highlighting the physiological diversity of different plant species (Lee et al., 2011; Burakowski et al., 2018; Liao et al., 2018). Some of the inconsistencies among observational studies could emerge from the underestimation of ET in eddy-covariance measurements (Foken, 2008; Teuling, 2018). Different models generally exhibit a wide spread of the ET response to forestation, with an increase in ET in the multi-model mean (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Lejeune et al., 2017). The inter-model spread in the evaporative fraction response to forestation was the dominant factor for the spread in the summer $\Delta T2M^{tot}$ in a multi-model experiment in Europe (Davin et al., 2020). This indicates that discrepancies in ΔET^{1} among models play a central role for the disagreement regarding the temperature response to forestation.

1.4.4 Overall effect of forestation on temperature

The overall biogeophysical effect of forestation on temperature varies in space and time, as the relevance and effect of the different biogeophysical properties altered by forestation vary with radiative and atmospheric conditions at the land surface. LST is the temperature variable most closely linked to the energy balance at the land surface. In tropical regions, the lower albedo and subsequent increase in absorbed solar radiation of forests is over-compensated by larger turbulent heat fluxes and a higher bowen-ratio of forests compared to open land, resulting in a negative ΔLST^{1} (Fig. 1.7 c; von Randow et al., 2004; Alkama and Cescatti, 2016; Li et al., 2015; Li et al., 2016b; Schultz et al., 2017; Duveiller et al., 2018b; de Oliveira et al., 2019; Zeppetello et al., 2020). Moving towards higher latitudes, $\Delta\alpha^{1}$ increases gradually, while ΔET^{1} decreases (Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b). In consequence, ΔLST^{1} becomes positive between roughly 40-50° N/S (Li et al., 2015; Ge et al., 2019). Beyond 30° N/S the sign of ΔLST^{1} varies seasonally, with positive values during winter, when $\Delta\alpha^{1}$ is intensified due to the snow-masking effect, and negative values during summer, when a positive ΔET^{1} dominates.

ΔLST^{1} varies not only seasonally, but also over the diurnal cycle. The (shortwave) albedo effect is unimportant during the night due to the absence of solar radiation. Similarly, the surface resistance to ET is less relevant at night due to the lack of solar radiation for photosynthesis, which causes plants to close their stomata to reduce water loss (Note however that nighttime ET still accounts for a non-negligible fraction of total ET; see Padrón et al., 2020). In addition, the higher surface roughness of forests has a warming effect when the surface layer is stable. Thus, the cooling effect of forests on LST emerges most distinct during the day, while ΔLST^{1} is only slightly negative during night in the tropics and even positive outside of the tropics (Fig. 1.7 b; Alkama and Cescatti, 2016; Li et al., 2015; Li et al., 2016b; Schultz et al., 2017; Duveiller et al., 2018b). Overall, forestation therefore dampens the diurnal temperature range in observations. Further, the T2M response to forestation in observations has generally the same sign as the LST response, but is smaller in magnitude (Vanden Broucke et al., 2015; Alkama and Cescatti, 2016).

Global and regional climate models exhibit a wide inter-model spread in the temperature response to deforestation (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Li et al., 2018; Davin et al., 2020), mainly caused by discrepancies in the simulated

ΔET^{l}. Also, models disagree with observations often even in sign, in particular regarding the effect of forestation on the diurnal temperature cycle (Fig. 1.13; Vanden Broucke et al., 2015; Lejeune et al., 2017). LST and T2M can respond in opposite direction to forestation in models especially during the day (Chen and Dirmeyer, 2019a; Winckler et al., 2019b; Breil et al., 2020). As such, part of the disagreement with observations may originate from comparing the wrong model variable to observations. In addition, observations neglect atmospheric feedbacks due to the space-for-time assumption and are restricted to clear-sky conditions in the case of remote sensing observations. In the case of CESM, the simulated LST response without atmospheric feedbacks conforms with observations, while the signal that includes atmospheric feedbacks differs considerably (Chen and Dirmeyer, 2020). ESMs suggest that the change in temperature extremes due to forestation differs from the change in the mean state. Lejeune et al. (2018) find a cooling in warm temperature extremes employing the reconstruction method on simulations of several ESMs. On the other hand, a multi-model factorial experiment finds an increase in warm temperature extremes (Li et al., 2018). This points towards the importance of ΔX^{rem} also for temperature extremes as the method of the first study ignores atmospheric feedbacks while they are accounted for in the second study.

1.4.5 Remote effects of forestation

As mentioned previously, some of the discrepancies between observed and modeled temperature responses to deforestation may originate from the neglect of remote effects associated with common observation-based approaches. Several modeling studies have revealed a considerable non-local warming due to forestation outside of the tropics (Fig. 1.14; note that this figure shows the effect of deforestation; Winckler et al., 2017; Devaraju et al., 2018; Winckler et al., 2019c; Winckler et al., 2019b; Chen and Dirmeyer, 2020). This warming originates from exporting the additional energy absorbed due to the negative $\Delta\alpha$ following forestation to downwind areas (Davin and de Noblet-Ducoudré, 2010; Winckler et al., 2019c). Unlike ΔLST^{l}, ΔLST^{rem} is strongly dependent on the extent of forestation (Fig. 1.14; Winckler et al., 2019c). As such, modeling studies likely exaggerate the relevance of remote effects, as they estimate the ΔLST^{rem} of unrealistic extreme forestation scenarios in most cases (e.g., forestation of the entire land area). Observation-based studies on ΔLST^{rem} are scarce, because the extent of forestation in reality is comparably small resulting in weaker remote effects that are difficult to detect. Only one study has estimated the non-local effect of forestation in the tropics, finding a cooling effect on daily maximum temperatures (Cohn et al., 2019). Therefore, it is difficult to assess the realism of the simulated ΔLST^{rem}, which can differ substantially among different ESMs (Devaraju et al., 2018; Winckler et al., 2019c).

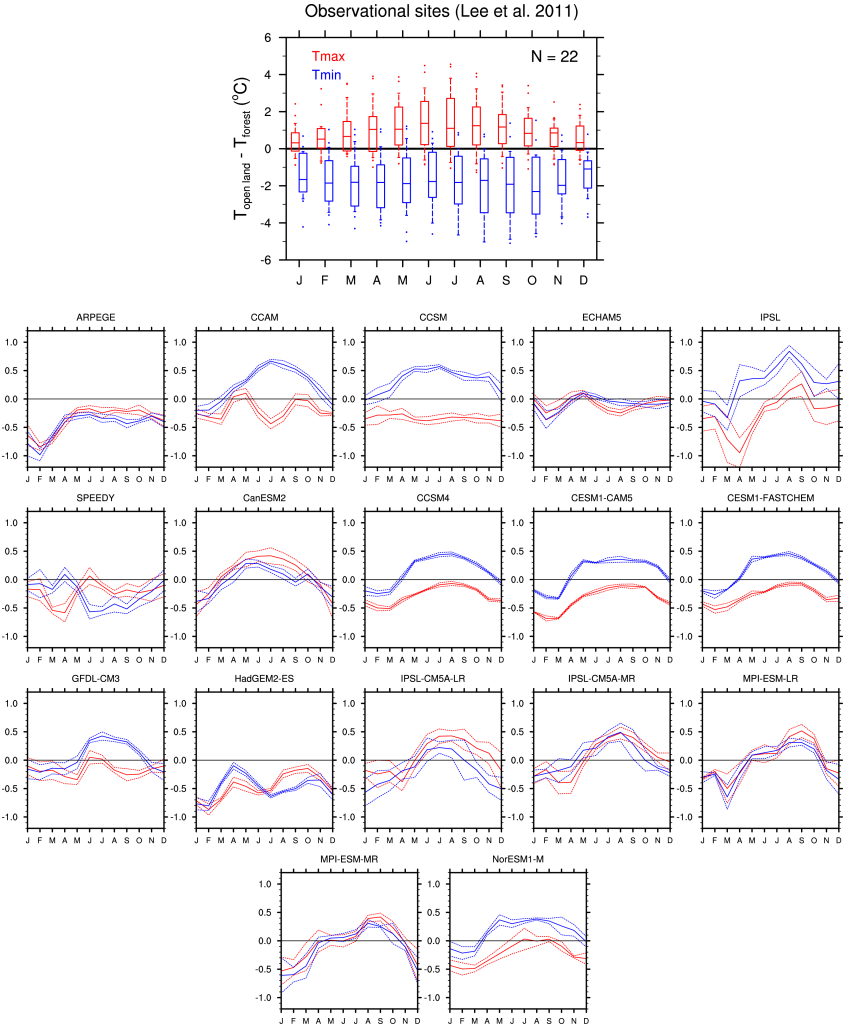


Figure 1.13: *At the top, seasonal cycle of the mean observed difference in daily maximum (red) and minimum (blue) air temperatures between open land and forest over 22 selected site pairs from Lee et al. (2011). The boxes indicate the interquartile range, while the whiskers show the range between the first and ninth deciles. Below, the seasonal cycle of the reconstructed $\Delta T2M^{l}$ for six LUCID models and 11 CMIP5 models over North America. The full lines indicate the results for the ensemble mean, while the dashed lines represent the spread between ensemble simulations (two standard deviations). Note the different y-axis scale between the topmost panel and the others. From Lejeune et al. (2017).*

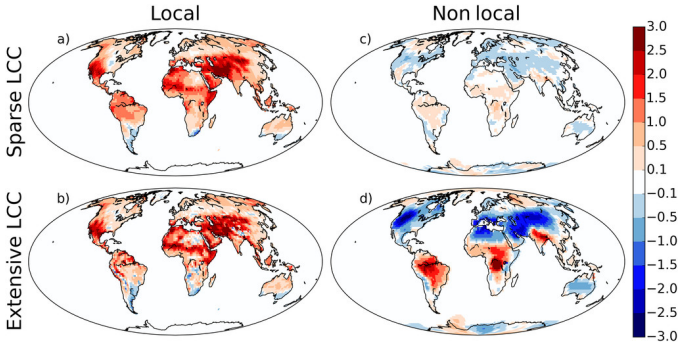


Figure 1.14: *Boreal summer (JJA) change in mean LST [K] due to sparse deforestation (a, c; one out of eight grid cells) and extensive deforestation (b, d; seven out of eight grid cells). Panels (a), (b), local effects and panels (c), (d) remote effects. From Winckler et al. (2017).*

1.4.6 Effect of forestation on the hydrological cycle

The aforementioned changes in biogeophysical properties due to forestation may not only alter temperatures but also the hydrological cycle. Modeling studies find an increase in precipitation due to forestation in most cases, in particular in tropical regions (Lejeune et al., 2015; Perugini et al., 2017; Winckler et al., 2017; Cherubini et al., 2018a). However, precipitation may be increased locally above cleared forest patches in the tropics (Khanna et al., 2017).

On the observational side, only few studies exist on the impact of forestation on the water cycle besides the estimates of ΔET described in Section 1.4.3. In tropical regions, forestation appears to increase precipitation (Spracklen et al., 2012; Kumagai et al., 2013), likely owed to the high ET of forests. Teuling et al. (2017) find enhanced cloud coverage over two large forest patches in Europe compared to surrounding areas during summer (Fig. 1.15), which might be linked to a high sensible heat flux over forests (Bosman et al., 2019). More clouds over forests could be connected to more precipitation. Further, increasing the forest coverage is estimated to reduce runoff, hinting towards higher ET of forests compared

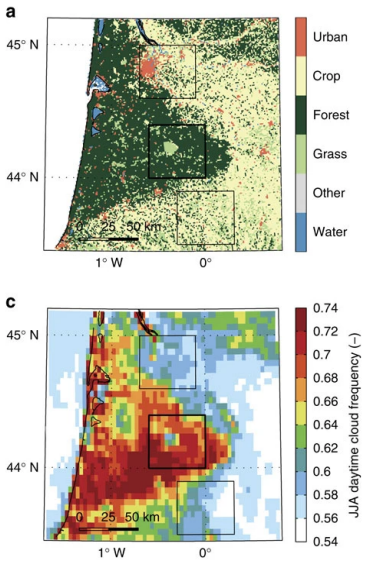


Figure 1.15: *Panel (a), land cover map for the Landes region in south-western France. Panel (c), mean JJA cloud frequency (2004-2008, 6-18UTC) for the MSG-CPP product. From Teuling et al. (2017).*

to open land (Trabucco et al., 2008; Buendia et al., 2016; Teuling et al., 2019). Overall, observational studies on the consequences of forestation on the hydrological cycle exhibit high uncertainties and are sometimes contradictory, which impedes the evaluation of the representation of such biogeophysical effects in ESMs.

1.5 Objectives and outline

In this thesis, I intend to foster our understanding of the biogeophysical consequences of converting open land, defined as grassland and crops. In the present chapter, I have given an overview of the current state of research on the biogeophysical effects of forestation. The following chapters consists of three published articles, each forming one chapter of the thesis (Meier et al., 2018; Meier et al., 2019; Meier et al., 2020). The final chapter synthesizes my results and provides an overview of potential directions for future research. Specifically, my thesis is a contribution to answer the following three research questions:

- Question 1: What are the biogeophysical effects of forestation on temperature and precipitation?
- Question 2: Which are the underlying mechanisms of those biogeophysical effects?
- Question 3: How well are the biogeophysical effects represented in the Community Land Model (CLM) and how could their representation be improved?

I address these research questions in the following three studies:

Chapter 2 *"Evaluating and improving the Community Land Model's sensitivity to land cover". Meier et al. (2018), Biogeosciences.* The emergence of global observation-based data sets on the local biogeophysical effect of forestation enables to assess their representation in LSMs. In this study, I confront the Community Land Model 4.5 (CLM4.5) with several observation-based data sets. I evaluate the representation of $\Delta\alpha^{l}$, ΔET^{l}, and ΔLST^{l} in CLM4.5 and propose several modifications of this model to alleviate some of the identified discrepancies with the observation-based data sets. This chapter therefore focuses on Questions 1 and 3.

Chapter 3 *"Biomass heat storage dampens diurnal temperature variations in forests". Meier et al. (2019), Environmental Research Letters.* Energy is stored in the vegetation biomass during the course of the day. Several in-situ studies have estimated that biomass heat storage is of relevant magnitude in particular in forests. However, most LSMs, including CLM, disregard this energy storage term. In this chapter, I test the effect of including biomass heat storage in CLM5.0 on the simulated diurnal temperature variability and on the biogeophysical temperature effect of forestation. As such, this chapter addresses both Questions 2 and 3.

Chapter 4 *"Empirical estimate of forestation-induced precipitation changes in Europe". Meier et al. (2020), Nature Geoscience.* To date, no observation-based study on precipitation changes due to forestation has been published for regions outside of the tropics, despite evidence from several modeling studies that they could be relevant. In this chapter, I estimate the local and remote precipitation change due to forestation in Europe employing a site-pair analysis as well as a statistical model. This study contributes therefore to answering Question 1 and provides some first insights into the underlying mechanisms (Question 2).

2

Evaluating and improving the Community Land Model's sensitivity to land cover

*Biogeosciences, 15, 4731-4757, https://doi.org/10.5194/bg-15-4731-2018**, Ronny Meier[1], Edouard L. Davin[1], Quentin Lejeune[2], Mathias Hauser[1], Yan Li[3], Brecht Martens[4], Natalie M. Schultz[5], Shannon Sterling[6], and Wim Thiery[7]

Abstract Modelling studies have shown the importance of biogeophysical effects of deforestation on local climate conditions, but have also highlighted the lack of agreement across different models. Recently, remote sensing observations have been used to assess the contrast in albedo, EvapoTranspiration (ET), and Land Surface Temperature (LST) between forest and nearby open land on a global scale. These observations provide an unprecedented opportunity to evaluate the ability of land surface models to simulate the biogeophysical effects of forests. Here, we evaluate the representation of the difference of forest minus open land (i.e., grassland and cropland) in albedo, ET, and LST in the Community Land Model version 4.5 (CLM4.5) using various remote sensing and in-situ data sources. To extract the local sensitivity to land cover we analyse plant functional type level output from global CLM4.5 simulations, using a model configuration that attributes a separate soil column to each plant functional type. Using the separated soil column configuration, CLM4.5 is able to realistically reproduce the biogeophysical contrast between forest and open land in terms of albedo, daily mean LST, and daily

*This publication was slightly changed from its original version to ensure consistency throughout this thesis.

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]Climate Analytics, Berlin, Germany
[3]Faculty of Geographical Science, Beijing Normal University, Beijing, China
[4]Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium
[5]School of Forestry and Environmental Studies, Yale University, New Haven, USA
[6]Department of Earth Sciences, Dalhousie University, Halifax, Canada
[7]Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium

maximum LST, while the effect on daily minimum LST is not well captured by the model. Furthermore, we identify that the ET contrast between forests and open land is underestimated in CLM4.5 compared to observation-based products and even reversed in sign for some regions, even when considering uncertainties in these products. We then show that these biases can be partly alleviated by modifying several model parameters, such as the root distribution, the formulation of plant water uptake, the light limitation of photosynthesis, and the maximum rate of carboxylation. Furthermore, the ET contrast between forest and open land needs to be better constrained by observations in order to foster convergence amongst different land surface models on the biogeophysical effects of forests. Overall, this study demonstrates the potential of comparing sub-grid model output to local observations to improve current land surface models' ability to simulate land cover change effects, which is a promising approach to reduce uncertainties in future assessments of land use impacts on climate.

2.1 Introduction

While the forested area has stabilized or is even increasing over Europe and North America, deforestation is still ongoing at a fast pace in some areas of South America, Africa, and Southeast Asia (Huang et al., 2009; Hansen et al., 2013; Margono et al., 2014; McGrath et al., 2015). In addition, carbon sequestration by re- or afforestation has been proposed as a strategy to mitigate anthropogenic climate change (Brown et al., 1996; Sonntag et al., 2016), making forest loss or gain likely an essential component of future climate change. Changes in forest coverage impact climate by altering both the carbon cycle (Ciais et al., 2013) and various biogeophysical properties of the land surface such as albedo, evaporative fraction and roughness length (Bonan, 2008; Pitman et al., 2009; Davin and de Noblet-Ducoudré, 2010; Li et al., 2015). However, there exist considerable discrepancies in the representation of biogeophysical effects amongst land surface models, thus generating a need for a thorough evaluation of the representation of these effects in individual models.

Model simulations indicate that the biogeophysical effects of historical deforestation have been rather small on a global scale (Davin et al., 2007; Findell et al., 2007; Davin and de Noblet-Ducoudré, 2010; de Noblet-Ducoudré et al., 2012; Malyshev et al., 2015). However, they have likely been significant on regional and local scales, especially over areas which experienced intense deforestation rates (Pongratz et al., 2010; de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Malyshev et al., 2015; Lejeune et al., 2017; Lejeune et al., 2018). Similarly, present-day observational data, either based on in-situ (Juang et al., 2007; Lee et al., 2011; Zhang et al., 2014; Bright et al., 2017) or remote-sensing measurements (Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016b; Duveiller et al., 2018b) show that biogeophysical effects of forests can strongly influence local climate conditions. Among the different biophysical effects, the increased surface albedo (cooling effect), the alteration of the evaporative fraction (warming or cooling effect, depending on the region and season), and the lower surface roughness causing a reduction of the turbulent heat fluxes (warming effect) have been identified as the three main drivers of the climate impact of deforestation (Bonan, 2008; Pitman

et al., 2009; Davin and de Noblet-Ducoudré, 2010; Li et al., 2015). However, some of these biogeophysical processes are not well represented in current land surface models. The model intercomparison projects LUCID (Land-Use and Climate, IDentification of robust impacts) and CMIP5 (Coupled Model Intercomparison Project Phase 5) exposed the lack of model agreement concerning the biogeophysical impacts of historical Land Cover Change (LCC), especially regarding the impact on EvapoTranspiration (ET) and temperature during the warm season over the mid-latitudes of the northern hemisphere (de Noblet-Ducoudré et al., 2012; Kumar et al., 2013; Lejeune et al., 2017). In addition, distinct discrepancies between present-day temperature observations and the simulated historical effects of LCC over North America were identified (Lejeune et al., 2017). This highlights the need for systematic evaluation and improvement of the representation of biogeophysical processes in land surface models.

Observing the local climatic impact of LCC is not straightforward. When temporally comparing observational data over an area undergoing LCC, it is difficult to disentangle the effect of the LCC forcing from other climatic forcings (e.g., greenhouse gas forcing). To overcome this difficulty, observational studies often spatially compare nearby sites of differing Land Cover (LC), assuming that they receive the same atmospheric forcing (e.g., von Randow et al., 2004; Lee et al., 2011). Hence, the sensitivity of land surface models to LC can be evaluated best with observational data by spatially comparing different LC types in models. Recently, Malyshev et al. (2015) employed a new approach to assess the local impacts of LCC in land surface models by comparing climate variables over tiles corresponding to different Plant Functional Types (PFTs) located within the same grid cell. Since PFT tiles within the same grid cell experience exactly the same atmospheric forcing, the resulting sub-grid LC signal extracted by this method achieves good comparability to local observations which contrast neighbouring forest and open land sites (Lee et al., 2011; Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016b). Here we aim to evaluate and improve the sensitivity of the Community Land Model 4.5 (CLM4.5) to LC, using observational data of the local contrast between forest and open land (i.e., grassland and cropland). In Section 2.3.1 of this study, we systematically analyse the representation of the local difference of forest minus open land in albedo, ET and land surface temperature (LST) in CLM4.5 against the newly released observational remote-sensing-based products of Li et al. (2015). The forest signal in CLM4.5 is extracted by comparing tiles corresponding to forest and open land, similar to Malyshev et al. (2015). Given the uncertainties in observation-based ET estimates, we further extend our evaluation by including data from the Global Land Evaporation Amsterdam Model (GLEAM) version 3.1a (Miralles et al., 2011; Martens et al., 2017) and the Global ET Assembly (GETA) 2.0 (Ambrose and Sterling, 2014), which are based on remote-sensing and in-situ observations, respectively. Finally, a sensitivity experiment is presented in Section 2.3.2, in order to explore the possibilities to better represent the ET impact of forests in CLM4.5. This configuration of CLM4.5 incorporates modifications in root distribution, plant water uptake, light limitation of photosynthesis, and maximum rates of carboxylation.

2.2 Methods and data

2.2.1 Model description and set up

CLM is the land surface component of the Community Earth System Model (CESM), a state-of-the-art earth system model widely applied in the climate science community (Hurrell et al., 2013). CLM represents the interaction of the terrestrial ecosystem with the atmosphere by simulating fluxes of energy, water and a number of chemical species at the interface between the land and the atmosphere. The represented biogeophysical processes include absorption and reflection of both diffuse and direct solar radiation by the vegetation and soil surface, emission and absorption of longwave radiation, latent and sensible heat fluxes from the soil and canopy, and heat transfer into the snow and soil. Sub-grid heterogeneity is taken into account in CLM by the subdivision of each land grid cell in five land units (glacier, wetland, vegetated, lake, and urban). The vegetated land unit is further divided into 16 tiles representing different PFTs (including bare soil).

We run CLM version 4.5 at 0.5° resolution for the period 1997-2010. A five-year (1997-2001) spin-up period is excluded from the analysis to minimize the impact of the model initialization. The analysis of CLM4.5 therefore covers the period of 2002 to 2010 which matches well with the observation period of 2002 to 2012 of Li et al. (2015). Assuming that the feedback of the land surface to the atmosphere is of minor importance for the sub-grid contrast between forest and open land tiles, simulations are performed in offline mode using atmospheric forcing from the CRUNCEP v4 reanalysis product (Vivoy, 2009; Harris et al., 2014). The LC map and vegetation state data are prescribed based on MOD - Li15 observations (Lawrence and Chase, 2007, Fig. A.1). The LC map from the year 2000 is kept static during the entire simulation period, since no LCC is required to retrieve a spatial contrast between forest and open land. The optional carbon and nitrogen module of CLM4.5 as well as the crop- and irrigation modules are kept inactive in our simulations.

By default, all PFTs within a grid cell in CLM4.5 share a single soil column (Oleson et al., 2013), implying that all PFTs experience the same soil temperature and soil moisture. Further, the surface energy balance at PFT level is closed using the ground heat flux (i.e., it is calculated as the residual of the other energy fluxes). Hence, the soil warms in case of an energy excess at the land surface and vice versa. Warmer (cooler) soil in turn will result in increased (decreased) sensible and latent heat fluxes away from the ground and/or increased (decreased) emitted longwave radiation, thereby counteracting the initial energy imbalance. Consequently, this model architecture eventually results in near-zero daily mean ground heat flux, once the soil temperature has adjusted to an equilibrium state with a near-zero energy imbalance. On shared soil columns, however, ground heat fluxes can reach unrealistically high values for individual PFTs (Fig. A.2 a and c), because a common soil temperature is artificially maintained for all PFTs, which differs from their individual equilibrium states. This assumption leads to a net ground heat flux into the soil over open land PFTs and out of the soil over forest PFTs for the majority of the locations across the globe, implying a lateral subsurface heat transport from open land towards forests (Schultz et al., 2016). To resolve this issue, Schultz et al. (2016) proposed a modification of CLM4.5 which attributes a separate soil column to each PFT. This

modification allows the soil of individual PFTs to equilibrate to a different temperature (Fig. A.3) and suppresses these unrealistically high (lateral) ground heat fluxes (Fig. A.2 b and d). Here, we present results from a simulation on separated soil columns, called CLM - BASE, unless stated otherwise (Table A.4). We also performed a simulation on shared soil columns named CLM - DFLT.

Further, we present a sensitivity experiment, named CLM - PLUS in Section 2.3.2, in which we try to alleviate detected biases in ET. Besides the separated soil columns, four aspects in the parametrisation of vegetation transpiration are modified in this sensitivity experiment:

- Shallower root distribution for grass- and cropland PFTs. CLM4.5 accounts for soil moisture stress on transpiration through a stress function β_t, which ranges from zero (when soil moisture limitation completely suppresses vegetation transpiration) to one (corresponding to no soil moisture limitation of vegetation transpiration). Forests for the most part experience higher soil moisture stress than open land in CLM - DFLT except in the northern high-latitude winter (Fig. A.4), partly caused by the similar root distribution for all PFTs but evergreen broadleaf trees (Fig. A.5). In reality, observed maximum rooting depths are considerably higher for forests than for grassland and cropland (Canadell et al., 1996; Fan et al., 2017). Likewise, in-situ observations in the tropics show that grassland ET decreases during dry periods, because grasses have only limited access to water reservoirs located below a depth of 2 m (von Randow et al., 2004). Hence, we aim to increase soil moisture stress of open land PFTs and reduce their ability to extract water from the lower part of the soil, by introducing a shallower root distribution for these PFTs (Fig. A.5). This root distribution was not fitted to a particular observed root distribution. However, the new root distribution agrees better with the average rooting depth of annual grass reported by Fan et al. (2017).

- Dynamic plant water uptake. Tropical forests are often observed to exhibit increased ET during dry periods, due to increased incoming shortwave radiation (da Rocha et al., 2004; Huete et al., 2006; Saleska et al., 2007). That is, despite the upper soil being dry, tropical trees still have sufficient access to water from deeper soil layers (Jipp et al., 1998; von Randow et al., 2004). We aim to allow a similar behaviour in CLM4.5 by introducing a dynamic plant water uptake, where plants only extract water from the 10 % of the roots with best access to soil moisture (example in Fig. A.6).

- Light limitation reduction for all C_3 PFTs and enhancement for C_4 PFTs. In CLM - BASE ET of boreal PFTs is underestimated compared to GETA 2.0 (Fig. 2.4 f). Since vegetation transpiration of these PFTs is only weakly affected by soil moisture stress, light limitation for C_3 plants is reduced. On the other hand, C_4 grass shows a considerable positive bias in ET, which we try to alleviate by increasing the light limitation of this PFT.

- Modified maximum rates of carboxylation (V_{cmax}; Table A.1). This PFT-specific parameter is suitable to tune vegetation transpiration, since it is not well constrained

from observations and vegetation transpiration in models is highly sensitive to this parameter (Bonan et al., 2011). The new values were chosen with the aim to alleviate biases relative to GETA 2.0 (Fig. 2.4 f) and still lie well within the range of observations collected in the TRY plant trait database (Boenisch and Kattge, 2017). Additionally, the minimum stomatal conductance of C_4 plants, which is by default four times larger than that of C_3 plants, is reduced.

A technical description of these modifications as well as a discussion of the effect on ET by each individual modification is provided in Appendix A.1.

2.2.2 Observational data

The data published in Li et al. (2015) are used to evaluate the effects of forests on local climate variables in CLM4.5. This data set was created by applying a window searching algorithm to remote-sensing LST, albedo, and ET products from the MODerate resolution Imaging Spectroradioameter (MODIS), in order to systematically compare these variables over forest and open land on a global scale. The data of this study, hereafter referred to as MOD - Li15, cover the period of 2002 to 2012 and were aggregated from the initial window size of $0.45° \times 0.25°$ to $0.5° \times 0.5°$ spatial resolution. Hence, the similar spatial scale of the MOD - Li15 data and the CLM4.5 simulations allows for good comparability between these two data sources.

We also use two additional observation-based datasets of ET to consider uncertainties in present-day ET estimates. A number of different global ET products are available which, however, exhibit substantial discrepancies (Mueller et al., 2011; Wang and Dickinson, 2012; Mueller et al., 2013; Michel et al., 2016; Miralles et al., 2016). In particular, the algorithm from Mu et al. (2011) used to retrieve the MODIS ET product was found to systematically underestimate ET compared to in-situ and catchment-scale observations (Michel et al., 2016; Miralles et al., 2016). In addition, algorithms used to infer ET from remote-sensing observations make assumptions on how the LC type influences ET, preventing an independent identification of the influence of LCC on ET. We therefore complement our evaluation of the ET impact of forest in CLM4.5 with two additional data sets: GLEAM version 3.1a and GETA 2.0. GLEAM was introduced in 2011 (Miralles et al., 2011) and revised twice, resulting in the current version 3.1 (Martens et al., 2017). It provides estimates of potential ET for tall canopy, bare soil, and low vegetation after Priestly and Taylor (1972). Potential ET of vegetated land surfaces is converted into actual ET using vegetation-dependent parametrisations of evaporative stress. Canopy interception evaporation is calculated separately using the parametrisation of Gash and Stewart (1979). GLEAM uses surface radiation, near-surface air temperature, surface soil moisture, precipitation, snow water equivalent, and vegetation optical depth observations to estimate ET globally at 0.25° resolution. To maximize spatial and temporal overlap with the MOD - Li15 observations, we choose GLEAM version 3.1 a (hereafter referred to as GLEAM), which incorporates reanalysis input besides satellite observations. We compare the ET estimates for tall canopy and low vegetation to model output for forests and open land, respectively. Since interception loss is only estimated for tall canopy, it was fully attributed to ET from forests.

GETA 2.0 (Ambrose and Sterling, 2014) is a suite of global-scale fields of actual ET for 16 separate LC types derived from a collection of in-situ measurements between 1850 and 2010. Using a linear mixed effect model with air temperature, precipitation, and incoming shortwave radiation as predictors, yearly ET estimates for each of these 16 different LC types have been obtained with a global coverage and 1° spatial resolution. We then use the same LC map employed for the CLM4.5 simulations to weigh the different LC types in this data set and retrieve an ET value for forest and open land (see Section 2.2.3 for more details). Since our CLM4.5 simulations were conducted without irrigation, we did not include the GETA 2.0 irrigation layer. We refer to this data set as GETA in this study.

2.2.3 Model evaluation

The forest signal in CLM4.5 is extracted by comparing the area-weighted mean of the variables of interest over all forest tiles to its corresponding values over open land tiles (i.e., grassland and cropland), similar to Malyshev et al. (2015). As such, it becomes possible to infer a forest signal for every model grid cell containing any forest and any open land PFT, no matter how small the fraction of the grid cell covered by these PFTs. The different PFT tiles within a $0.5° \times 0.5°$ grid cell in our CLM4.5 simulations are subject to the exact same atmospheric forcing and are hence comparable to the almost local effect of forests retrieved at a resolution of $0.45° \times 0.25°$ in MOD - Li15. It needs to be noted that the MODIS observations can only be retrieved under clear-sky conditions, thereby potentially impairing the comparability to our CLM4.5 data which are not filtered for clear-sky days. Nevertheless, it was decided to include cloudy days for the analysis of the CLM4.5 simulations, to preserve the comparability to studies which do not distinguish between cloudy and clear-sky days (e.g. GLEAM; GETA; da Rocha et al., 2004; von Randow et al., 2004; Liu et al., 2005).

Twelve of the 16 PFTs of CLM4.5 are attributed to either the forest or the open land class as described in Table A.2. Consistent with Li et al. (2015), open land was considered the combination of grassland and cropland. Hence, bare soil as well as shrubland are excluded from our analysis. Forest and open land ET of GETA was aggregated similarly using the same LC map as in the CLM4.5 simulations, with the LC types of GETA attributed to the different CLM4.5 PFTs as listed in Table A.3. To ensure a consistent comparison with the LST data from MODIS, we derive a radiative temperature (T_{rad}) from the emitted longwave radiation output (LW_{up}) in CLM4.5 according to Stefan-Boltzmann's law (assuming that emissivity is 1 as in Eq. 4.10 of Oleson et al., 2013):

$$T_{rad} = \sqrt[4]{\frac{LW_{up}}{\sigma}} \tag{2.1}$$

with σ being the Stefan-Boltzmann constant ($5.67 \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$). Hereafter T_{rad} will be referred to as LST. For the local difference of forest minus open land in albedo, ET, daily mean LST, daily maximum LST, and daily minimum LST we will use the symbols $\Delta\alpha^{l}$, $\Delta\mathrm{ET}^{l}$, $\Delta\mathrm{LST}^{l}_{\mathrm{avg}}$, $\Delta\mathrm{LST}^{l}_{\mathrm{max}}$, and $\Delta\mathrm{LST}^{l}_{\mathrm{min}}$, respectively.

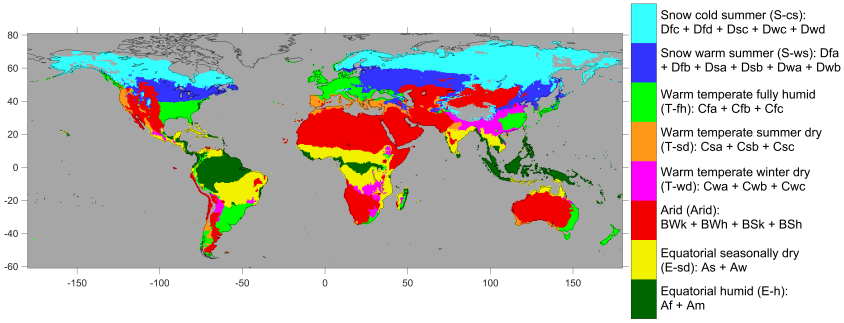


Figure 2.1: *The Köppen-Geiger climate zones (Kottek et al., 2006) used for the analysis.*

In order to evaluate the different CLM4.5 simulations objectively, three different metrics are calculated over the following eight Köppen-Geiger climate zones (Kottek et al., 2006): Equatorial humid (E-h), Equatorial seasonally dry (E-sd), arid (Arid), warm Temperate winter dry (T-wd), warm Temperate summer dry (T-sd), warm Temperate fully humid (T-fh), Snow warm summer (S-ws), and Snow cold summer (S-cs) (Fig. 2.1). As a first metric, the area-weighted mean for a given variable over these climate zones ($\overline{\Delta X_i}$) is calculated as follows:

$$\overline{\Delta X_i} = \frac{\sum_i A_i \Delta X_i}{\sum_i A_i} \tag{2.2}$$

where ΔX_i are the differences of forest minus open land in variable X of all the grid cells i belonging to the respective climate zone and A_i their areas. Secondly, the CLM4.5 simulations are compared in terms of the area-weighted Root Mean Squared Deviation (RMSD) to the observation-based data sources:

$$RMSD(\Delta X_i^l) = \sqrt{\frac{\sum_i A_i \left(\Delta X_{i,sim}^l - \Delta X_{i,obs}^l\right)^2}{\sum_i A_i}} \tag{2.3}$$

where $\Delta X_{i,sim}^l$ and $\Delta X_{i,obs}^l$ are the simulated and observed differences of forest minus open land in variable X. RMSD for a Köppen-Geiger climate zone is calculated from a data pool collecting all monthly values with data in CLM4.5 and the given observational data which lie within the respective climate zone (except when comparing to GETA for which only long-term annual means are available). Lastly, the Index of Agreement (IA; Duveiller et al., 2016) was calculated for the same data pools as RMSD. This dimensionless metric describes the agreement between two data sets, with 0 indicating no agreement and 1 indicating perfect agreement. By definition, this metric is set to 0 if the two compared data sets exhibit a negative Pearson correlation. Since results of this metric generally support those of RMSD, they are shown in the Appendix (Fig. A.8).

2.3 Results

2.3.1 Evaluation of the local effect of forests in CLM4.5

Albedo

The MODIS satellite observations and CLM - BASE agree on a generally negative $\Delta\alpha^1$ (Fig. 2.2). Effectively, MOD - Li15 observations show slightly positive $\Delta\alpha^1$ for some latitude-month combinations concentrated in the tropics and sub-tropics (Fig. 2.2); however, these differences are mostly insignificant and must be considered in the light of uncertainties in the MOD - Li15 observations, which are more sparse over these regions due to frequent cloud coverage (Li et al., 2015). The negative albedo difference is amplified towards the poles and in wintertime due to the snow masking effect (Harding et al., 2001). Among the non-snow climate zones, the albedo contrast between forest and open land is strongest in the Arid and the T-sd climate zones (Fig. 2.4 a). This could be related to the occurrence of dry periods in these climate zones during which open land dries out more easily than forests due to their shallower root profiles (Canadell et al., 1996; Fan et al., 2017). As green leaves have lower albedo than dry leaves and the soil, the albedo contrast between the still-green forest and the dried-out open land would be intensified in such a scenario (Dorman and Sellers, 1989). $\Delta\alpha^1$ tends to be more negative in CLM - BASE than in the satellite observations in all Köppen-Geiger climate zones, especially in the snow climate zones. RMSD values over the climate zones exhibit similar tendencies as the magnitudes of mean $\Delta\alpha^1$ and have roughly the same magnitude of mean $\Delta\alpha^1$ (Fig. 2.3 a). The exception to this are the tropical climate zones where the magnitude of RMSD is considerably higher than the mean values of $\Delta\alpha^1$. This is likely related to the fact that MOD - Li15 observes only a weak albedo signal of forests in these climate zones.

Evapotranspiration

All of the considered observation-based ET products indicate that annual mean ΔET^1 is positive in every climate zone, despite considerable variations in the magnitude of this difference (Fig. 2.4 e). GLEAM suggests a near zero ΔET^1 in the Arid climate zone most likely because it uses surface soil moisture data as an input to estimate ET. Also, GLEAM exhibits positive ΔET^1 throughout the year in the mid-latitudes, unlike MOD - Li15 which has a negative ΔET^1 during winter (Fig. 2.6). Paired-site FLUXNET studies offer an additional opportunity to compare ET over forest and over open land on a point scale. Overall, they report higher ET for tropical forests (Jipp et al., 1998; von Randow et al., 2004; Wolf et al., 2011). In the mid- and high-latitudes a number of FLUXNET studies observe a positive ΔET^1 during summer, and a near-zero negative ΔET^1 during winter, similar to MOD - Li15 (Fig. 2.6; Liu et al., 2005; Stoy et al., 2006; Juang et al., 2007; Baldocchi and Ma, 2013; Vanden Broucke et al., 2015; Chen et al., 2018). On the other hand, negative ΔET^1 have been observed at some paired FLUXNET sites in the tropics (Van der Molen et al., 2006) and in the mid-latitudes during summer (Teuling et al., 2010). The considered global ET data sets however consistently exhibit higher ET over forests in most regions (Fig. 2.5). This agreement across the different independent global

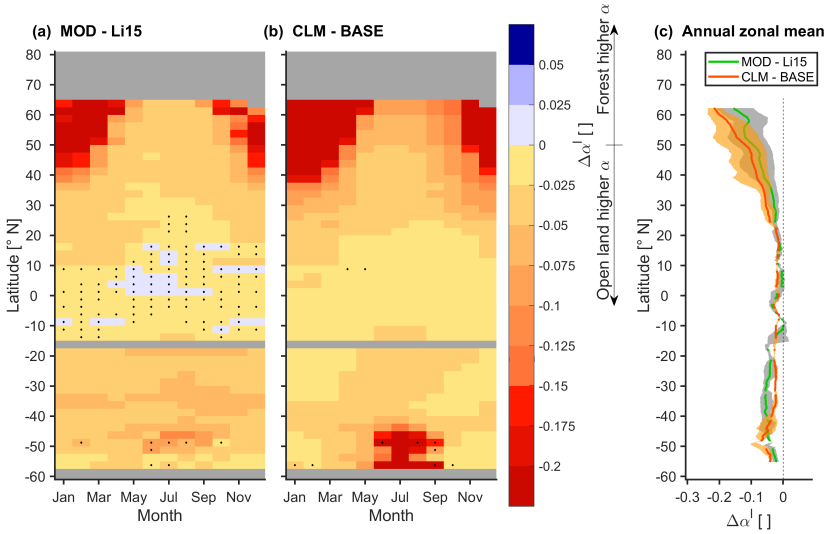


Figure 2.2: *Seasonal and latitudinal variations of $\Delta\alpha^1$ in (a) the MOD - Li15 observations and (b) CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. Panel (c) shows the zonal annual mean of both MOD - Li15 (in green along with its interquartile range in grey) and CLM - BASE (in red, interquartile range in orange). Note that on this subfigure results have been smoothed with a 4° latitudinally-running mean. Only grid cells containing valid data in the MOD - Li15 observations were considered for the analysis of CLM - BASE.*

data sources gives some confidence in the fact that ET is generally higher over forests. Nevertheless, it needs to be noted that ΔET^1 GETA shows fundamentally different results when considering the data over irrigated crops instead of data over rainfed (resulting in negative ΔET^1 at many locations). Therefore, distinguishing irrigated from rainfed crops in future evaluations would be essential, but remains beyond the scope of this study.

CLM - BASE exhibits considerable discrepancies in ΔET^1 to the observation-based data sets both for the annual mean values (Fig. 2.5) and the seasonal cycle (Fig. 2.6). ΔET^1 in CLM - BASE is near zero in all climate zones (Fig. 2.4 e), and even negative in the E-sd climate zone, unlike the global ET datasets which clearly suggest positive values. The large bias of ΔET^1 in CLM - BASE is also apparent in the RMSD values, which tend to be slightly larger than the observed mean signal (compare Figs. 2.4 e and 2.3 e). A comparison of the absolute ET values of each PFT in CLM - BASE versus the GETA data reveals that CLM - BASE generally exhibits similar ET averages for needleleaf PFTs, lower ET averages for broadleaf deciduous trees as well as crops, and higher ET averages for non-arctic grasses and broadleaf evergreen trees (Fig. 2.4 f). Notably, evergreen and deciduous tropical broadleaf trees as well as C_4 grass have a bias larger than 0.2 mm/day relative to GETA. The biases of these PFTs can have a large effect on the overall ΔET^1 as

they cover a large proportion of the land surface (9.5 %, 8.0 %, and 8.0 %, respectively). Similarly, CLM - BASE overestimates ET compared to in-situ measurements conducted over a pasture site in the Amazon by von Randow et al. (2004) and underestimates ET compared to the two forest sites in Alaska reported in the study of Liu et al. (2005) (Table 2.1).

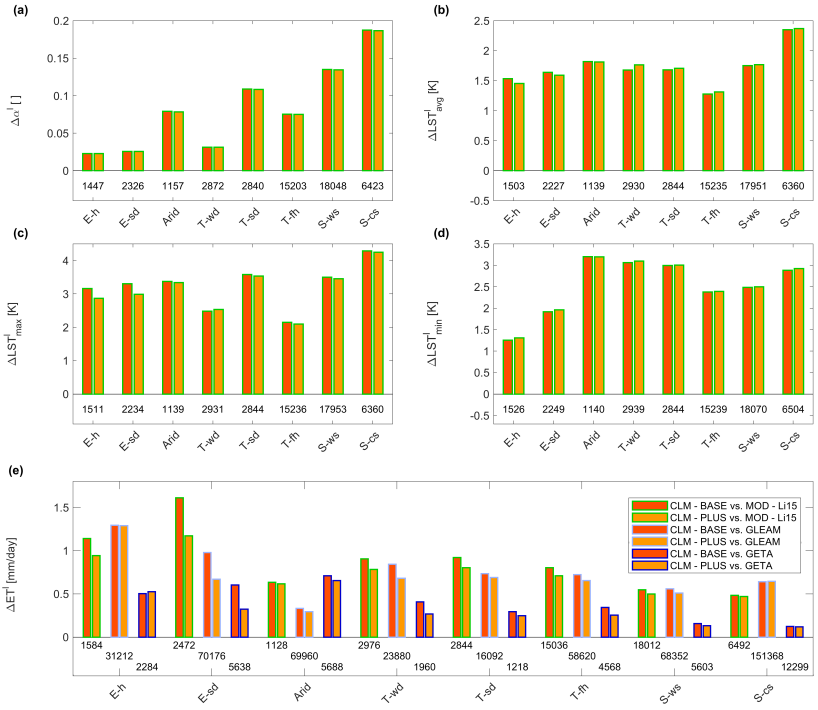


Figure 2.3: *RMSD of CLM - BASE (red), and CLM - PLUS (orange) against MOD - Li15 observations over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 2.1) of monthly (a) $\Delta\alpha^l$, (b) ΔLST^l_{avg}, (c) ΔLST^l_{max}, and (d) ΔLST^l_{min}. Panel (e) shows the RMSD over the Köppen-Geiger climate zone of ΔET^l of CLM - BASE (red), and CLM - PLUS (orange) against MOD - Li15 (green edge), GLEAM (light blue edge), GETA (dark blue edge). The numbers indicate the size of the data samples used for the calculation of RMSD.*

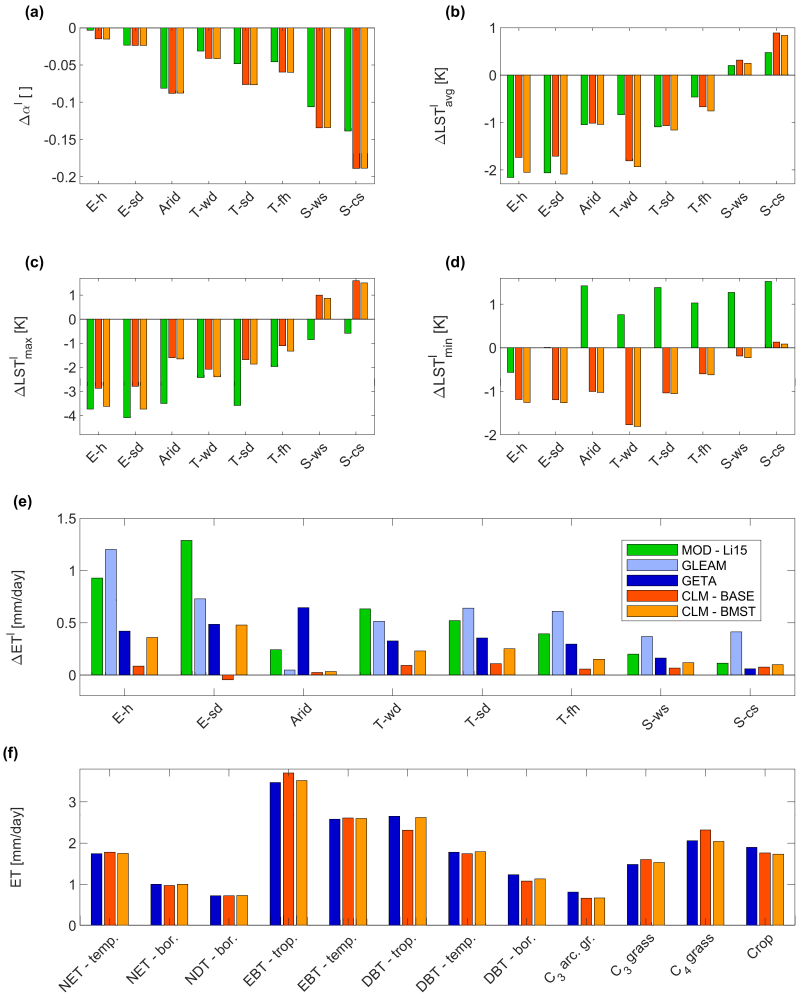


Figure 2.4: *Area-weighted annual mean over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 2.1) of (a) $\Delta\alpha^l$, (b) ΔLST^l_{avg}, (c) ΔLST^l_{max}, and (d) ΔLST^l_{min} in MOD - Li15 (green), CLM - BASE (red), and CLM - PLUS (orange). Only grid cells containing valid data in the MOD - Li15 observations were considered for analysis of CLM4.5. Panel (e) shows the area-weighted mean over the Köppen-Geiger climate zone of ΔET^l in MOD - Li15 (green), GLEAM (light blue), GETA (dark blue), CLM - BASE (red), and CLM - PLUS (orange) and panel (f) the area-weighted mean ET for each PFT analyzed in this study according to the GETA (dark blue), CLM - BASE (red), and CLM - PLUS (orange). The acronyms of the PFTs are defined in Table A.2.*

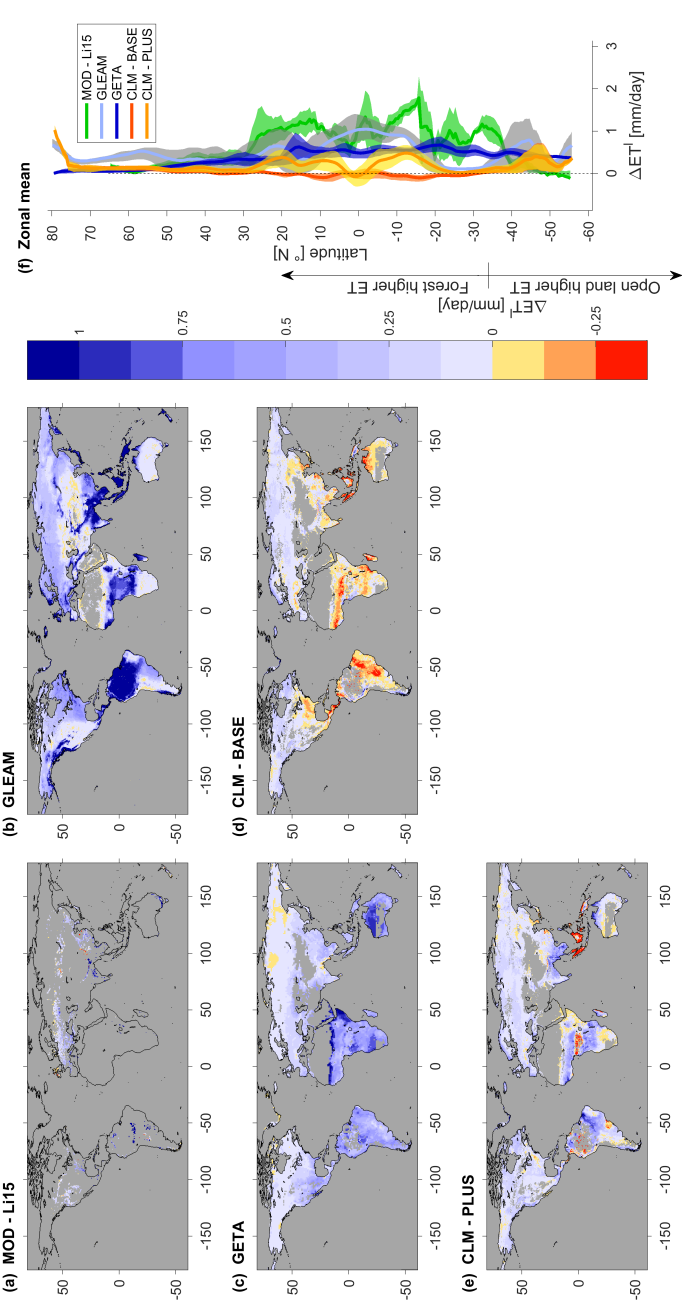


Figure 2.5: *Annual mean $\Delta ET'$ in (a) MOD - Li15, (b) GLEAM, (c) GETA, (d) CLM - BASE, and (e) CLM - PLUS. Panel (f) shows the zonal mean (thick line) and interquartile range (shading) of MOD - Li15 (green), GLEAM (light blue, grey shading), GETA (dark blue), CLM - BASE (red), and CLM - PLUS (orange). Note that on this subfigure results have been smoothed with a 4° latitudinally-running mean.*

Table 2.1: *ET and latent heat flux in-situ observations from various studies and the values in CLM - BASE and CLM - PLUS at the respective locations.*

Study	Region	PFTs	Unit	Season	Obs.	CLM - BASE	CLM - PLUS
da Rocha et al. (2004)	Amazon	EBT	mm/day	Dry	3.96	3.49	3.48
				Wet	3.18	3.57	3.37
				All	3.51	3.54	3.40
von Randow et al. (2004)	Amazon	EBT	$W\,m^{-2}$	Dry	108.6	82.9	90.8
				Wet	104.5	113.9	108.9
		Grass		Dry	63.9	81.2	64.7
				Wet	83.0	113.9	100.1
Liu et al. (2005)	Alaska	Grass	$W\,m^{-2}$	All	16.1	16.4	16.8
		DBT		All	22.5	13.7	14.1
		ENF		All	23.9	18.0	18.4

Interestingly, deciduous trees are mostly responsible for this discrepancy in ΔET^1 at latitudes below 30° (Fig. A.7). In the mid-latitudes, on the other hand, both deciduous and evergreen trees show lower ET than open land during summer and higher ET during winter, which is inconsistent with GLEAM and, even more so, inconsistent with the seasonally-varying ΔET^1 in MOD - Li15. Another noteworthy result is that the separated soil column configuration (i.e., CLM - BASE) appears to impair the agreement on ΔET^1 between CLM4.5 and the observations (Fig. 2.6). CLM - DFLT exhibits a positive ΔET^1 throughout the year except for the tropical dry season which is caused by deciduous broadleaf trees exhibiting lower ET than open land (Fig. A.7 a, b, and c). There are two potential reasons for the negative bias in ΔET^1 introduced by separated soil columns. First, the implicit lateral ground heat flux from open land towards forest which occurs in CLM - DFLT (Fig. A.2) provides additional energy over forests for turbulent heat fluxes. This energy source (sink) for forests (open land) is disabled by separated soil columns. Second, the lower soil temperature of forests in CLM - BASE (Fig. A.3) reduces the specific humidity gradient between the soil surface and the atmosphere and hence also the absolute soil evaporation. It needs to be noted that the weaker agreement with observational data of ΔET^1 in CLM - BASE than in CLM - DFLT does not necessarily imply a worse representation of the evaporative processes in CLM - BASE, but could also originate from the fact that CLM4.5 was tuned to retrieve realistic ET values on shared soil columns.

To shed light on the origin of the ΔET^1 bias in CLM4.5, we separately analyze the three components of ET in CLM4.5: soil evaporation (including sublimation/evaporation from the snow- and water-covered fraction of the soil), canopy interception evaporation, and vegetation transpiration. As seen in Fig. 2.7 (b), there is a distinct band around the equator where soil evaporation is considerably lower in forests than in open land. Interestingly, both the study of Chen et al. (2018) and ours show that the lower soil evaporation signal only arises for the configuration with separated soil columns (data of CLM - DFLT are not presented here). Thus, lower soil evaporation around the equator in CLM - BASE is likely related to the diminution of the soil temperature and of the available energy mentioned earlier in this section. It appears reasonable that, in comparison with open land, forests have lower soil evaporation since (1) the forest soil surface receives

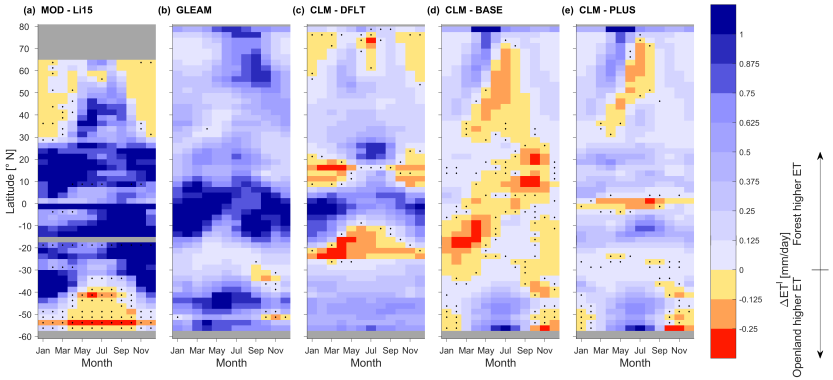


Figure 2.6: *Seasonal and latitudinal variations of ΔET^1 in (a) the MOD - Li15 and (b) GLEAM observations, (c) CLM - DFLT, (d) CLM - BASE, and (e) CLM - PLUS. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test.*

less incoming solar radiation, (2) more of the incoming precipitation is intercepted by the canopy, and (3) the water vapour concentrations within the canopy are higher. Yet soil evaporation and canopy interception evaporation contribute a larger proportion to total ET in CLM4.5 (31 % and 19 %) compared to GLEAM (14 % and 10 %; Martens and Miralles, 2017). It is thus possible that the strength of this effect is too large in CLM4.5. However, most ET measurement techniques cannot distinguish among the different components of ET, making it difficult to assess which partitioning is more realistic. Overall, negative ΔET^1 values in CLM - BASE typically coincide with negative differences in vegetation transpiration, in particular during the wet season in the tropics and sub-tropics and during summer at higher latitudes (Fig. 2.7 a and d), whereas negative values in the soil evaporation difference are partly compensated by positive values in interception evaporation (Fig. 2.7 b and c). It is therefore likely, that vegetation transpiration is the main driver behind the ΔET^1 bias even though the contribution of the individual ET components to the total signal cannot be evaluated. For this reason, the modifications in the CLM - PLUS sensitivity experiment are targeted at altering vegetation transpiration.

In summary, ΔET^1 in CLM4.5 exhibits considerable discrepancies to the considered global ET datasets and in-situ observations. The separated soil column configuration amplifies these discrepancies, which are typically driven by the difference in vegetation transpiration of forest minus open land.

Land surface temperature

The overall local temperature impact of forests is the result of several biogeophysical properties acting simultaneously. They include lower albedo of forests (warming effect), higher surface roughness (cooling effect if land surface is warmer than boundary layer), and alteration of the evaporative fraction (Bonan, 2008; Pitman et al., 2009; Davin and

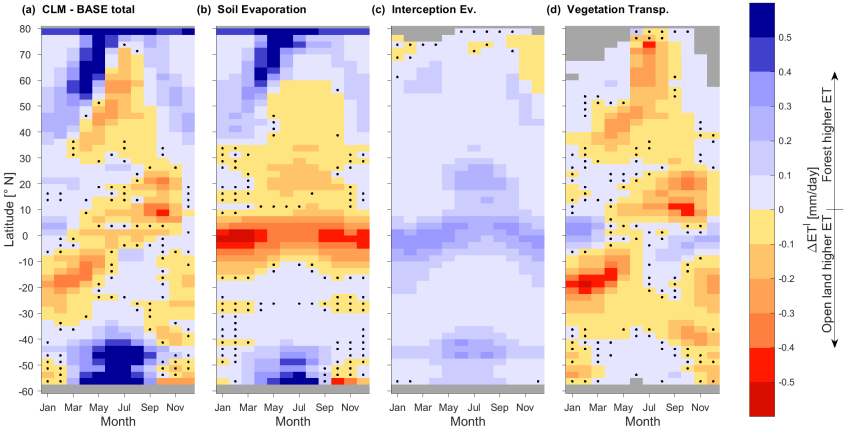


Figure 2.7: *Seasonal and latitudinal variations of difference of forest minus open land in (a) total ET, (b) soil evaporation, (c) canopy interception evaporation, and (d) vegetation transpiration in CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. Note that the colorbar in this figure is different from the one in Figs. 2.5 and 2.6*

de Noblet-Ducoudré, 2010; Li et al., 2015). For daily mean LST forests exhibit a cooling effect in MOD - Li15 except for the winter months at latitudes exceeding 30° (Fig. 2.9 a). This implies that the cooling effects of higher surface roughness and generally higher evaporative fraction over forests are stronger than the warming effect due to their lower albedo. ΔLST^{l}_{avg} and ΔLST^{l}_{max} are positive only under the presence of snow, as $\Delta\alpha^{l}$ is amplified due to the snow masking effect (moreover sensible heat fluxes are often directed towards the land surface during winter at high latitudes, resulting in warmer forests due to their higher surface roughness inducing stronger turbulent heat fluxes; Liu et al., 2005). The observed magnitude of ΔLST^{l}_{max} tends to be larger than that of ΔLST^{l}_{avg} likely due to the fact that the observed daytime effect is partly compensated by an opposing nighttime effect (Fig. 2.4 b, c, and d). MOD - Li15 exhibits an overall cooling effect of forests on daily mean LST in all climate zones, including the snow climate zone where the sign of the difference changes seasonally (Fig. 2.9 d). Further, this dataset shows a slightly negative ΔLST^{l}_{min} in tropical and sub-tropical regions and even a positive ΔLST^{l}_{min} in the mid-latitudes (Fig. 2.9 g). This nighttime signal in the mid-latitudes is observed in several observational studies but its source is not yet fully determined (Lee et al., 2011; Vanden Broucke et al., 2015; Li et al., 2015).

CLM - BASE generally captures the sign and magnitude of ΔLST^{l}_{avg} and ΔLST^{l}_{max} compared to MOD - Li15 (Fig. 2.9). The separated soil columns used in CLM - BASE allow for larger LST differences between forest and open land than the default version of CLM4.5 (CLM - DFLT) on shared soil columns, resulting in a better agreement with the observed magnitudes. This is due to the fact that the ground heat flux on shared soil columns counteracts the soil temperature difference and thereby also the LST difference between forest and open land. Nevertheless, there are still some discrepancies between the

LST signal in CLM - BASE and the MOD - Li15 observations. It appears that ΔLST^{l}_{avg} in CLM - BASE has a positive bias in the equatorial, the arid, and the snow climate zones and a negative bias in the T-wd and T-fh climate zones (Fig. 2.4 b). ΔLST^{l}_{max} in CLM - BASE appears qualitatively similar to the MOD - Li15 observations (Fig. 2.9 d, e, and f) but is biased positively in all climate zones (Fig. 2.4 c). In contrast, daily minimum LST shows much larger discrepancies between CLM - BASE and MOD - Li15 (Fig. 2.9 g, h, and i). In CLM - BASE, ΔLST^{l}_{min} is similar to ΔLST^{l}_{avg} and ΔLST^{l}_{max}, i.e. forests have an overall nighttime cooling effect in all climate zones except for the neutral signal in the snow climate zones, whereas MOD - Li15 exhibits an only weak nighttime cooling effect in the tropical climate zones and a clear nighttime warming effect in all other climate zones (Fig. 2.4 d). The weak performance of CLM - BASE in terms of ΔLST^{l}_{min} is also visible in the RMSD values which are considerably larger than the mean ΔLST^{l}_{min} signal (compare Figs. 2.4 d and 2.3 d)

Interestingly, and in contrast to LST, CLM4.5 simulates a small year-round warming effect of forests on daily maximum 2 m air temperature (T2M, Fig. 2.8). This contradicts a number of observational studies which show that the T2M difference of forest minus open land ($\Delta T2M^{l}$) has the same sign, but is attenuated compared to ΔLST^{l} (Li et al., 2015; Vanden Broucke et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016b). The fact we use offline simulations in our experiments might explain this behaviour, because some land-atmosphere feedbacks are not represented. However, Lejeune et al. (2017) report similar discrepancies of $\Delta T2M^{l}$ in CLM with observational data for coupled simulations, suggesting that the behaviour of $\Delta T2M^{l}$ in our simulations may not be related to the lack of atmospheric feedbacks.

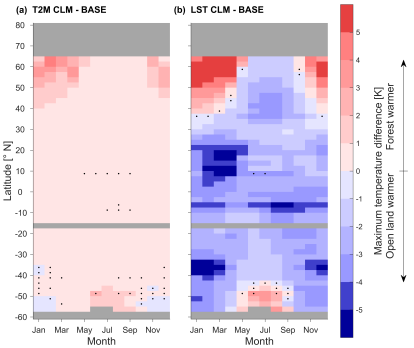
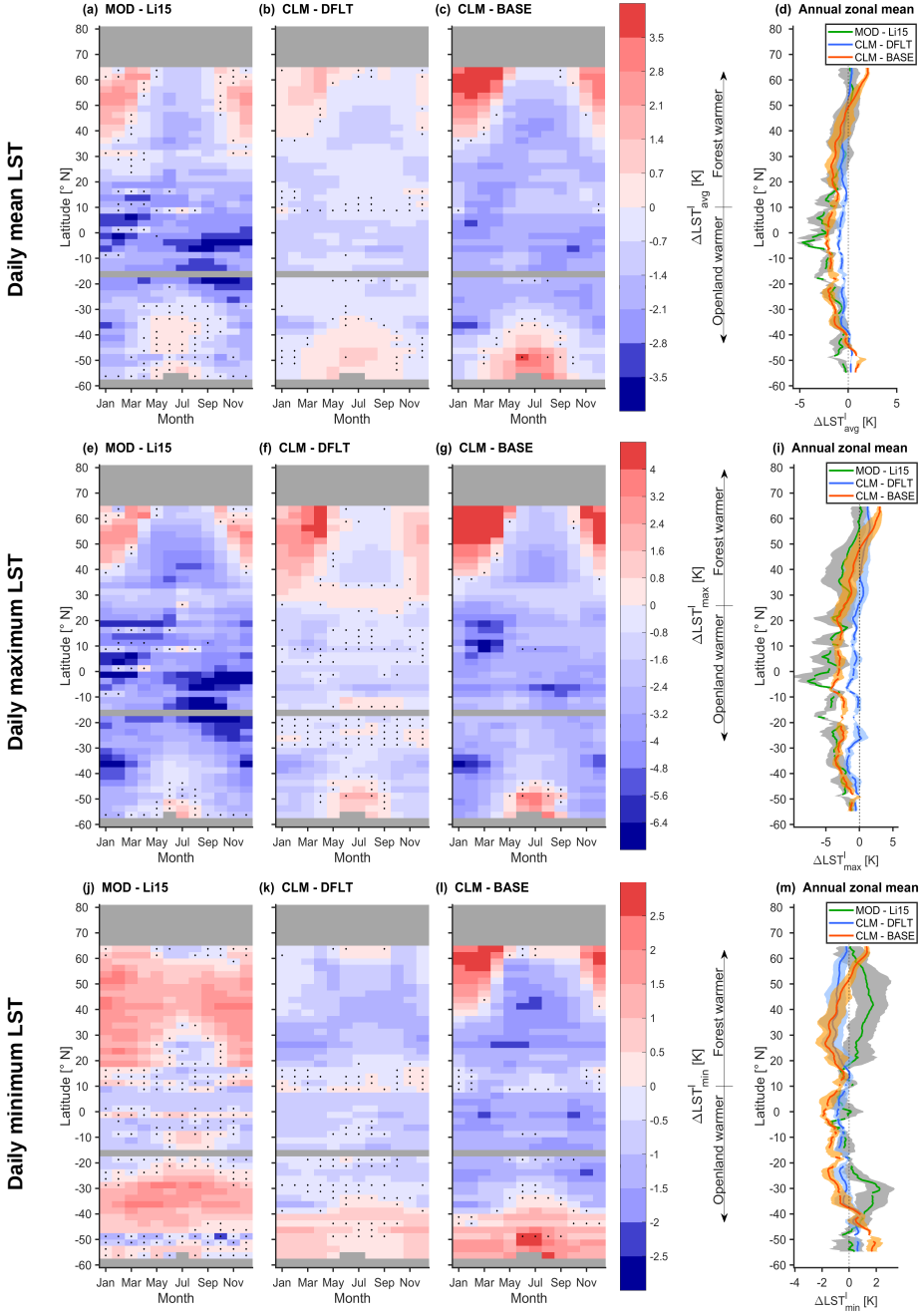


Figure 2.8: *Seasonal and latitudinal variations of (a) daily maximum T2M difference of forest minus open land and (b) ΔLST^{l}_{max} in CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. Only grid cells containing valid data in the MOD - Li15 observations were considered for the analysis.*

Figure 2.9: *(Next page) Seasonal and latitudinal variations of ΔLST^{l}_{avg} in (a) the MOD - Li15 observations, (b) CLM - DFLT, and (c) CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. Panel (d) shows the zonal annual mean of MOD - Li15 (green, interquartile range in grey), CLM - DFLT (blue, interquartile range in blue), and CLM - BASE (red, interquartile range in orange). Note that on this subfigure results have been smoothed with a 4° latitudinally-running mean. Only grid cells containing valid data in the MOD - Li15 observations were considered for the analysis of CLM - DFLT and CLM - BASE. The same for ΔLST^{l}_{max} in panels (e), (f), (g), (h) and for ΔLST^{l}_{min} in panels (i), (j), (k), (l).*


Daily mean LST
Daily maximum LST
Daily minimum LST
(a) MOD - Li15
(b) CLM - DFLT
(c) CLM - BASE
(d) Annual zonal mean
(e) MOD - Li15
(f) CLM - DFLT
(g) CLM - BASE
(i) Annual zonal mean
(j) MOD - Li15
(k) CLM - DFLT
(l) CLM - BASE
(m) Annual zonal mean
Latitude [° N]
Month
Jan Mar May Jul Sep Nov
MOD - Li15
CLM - DFLT
CLM - BASE
Openland warmer
Forest warmer
ΔLST avg [K]
ΔLST max [K]
ΔLST min [K]

2.3.2 Sensitivity experiment to alleviate ET biases in CLM4.5

In the previous section, striking discrepancies between the effect of forests in CLM - BASE and observation-based data were found for ΔET^1. An important driver responsible for these differences was identified to be vegetation transpiration (Fig. 2.7). In addition, it became apparent that the separated soil column configuration impairs the ΔET^1 compared to the shared soil column configuration (Fig. 2.6), despite improving ΔLST^1_{avg} and ΔLST^1_{max} (Fig. 2.9). Hence, in this section we aim to improve the comparability of modeled ΔET^1 to observation-based results by testing a modified parametrisation of vegetation transpiration in a sensitivity experiment called CLM - PLUS. This model configuration comprises (1) a shallower root distribution for open land PFTs, (2) a modified plant water uptake scheme whereby plants only extract water from the 10 % of the roots with best access to soil moisture, (3) altered light limitation of photosynthesis (decreased for C_3 plants and increased for C_4 plants), and (4) altered V_{cmax} values to alleviate ET biases at PFT level compared to the GETA data.

$\Delta\alpha^l$ is only marginally affected by the modifications of CLM - PLUS compared to CLM - BASE (Fig. 2.4 a). This is expected since the modifications are targeted at modifying vegetation transpiration which is not linked directly to albedo. ΔET^1 in CLM - PLUS becomes more positive than in CLM - BASE in all climate zones, thereby better matching the observation-based estimates (Fig. 2.4 e). The improvement is also apparent in the RMSD values which are reduced in CLM - PLUS for all data sets and climate zones, except for GETA in the E-h climate zone (Fig. 2.3 e). The bias in average ET compared to GETA is smaller in CLM - PLUS than in CLM - BASE for all PFTs except for boreal deciduous needleleaf trees and crops (Fig. 2.4 f). Some discrepancies with observation-based ET products nevertheless remain. ΔET^1 in CLM - PLUS is still mostly less positive compared to remote sensing-based observations and GETA, and remains of opposite sign during the warm season in the temperate regions and in a narrow band around the Equator (Figs. 2.6 and 2.4e). This band originates from a negative ΔET^1 around the western part of the equator in Africa and over Indonesia (Fig. 2.5). GLEAM and GETA observations cover these areas which explains the only moderate reduction of RMSD of CLM - PLUS against GLEAM and the increase in RMSD against GETA in the E-h climate zone. On the other hand, the RMSD against MOD - Li15 is reduced considerably in CLM - PLUS, since MOD - Li15 observations are sparse over Africa and Indonesia (Fig. 2.3 e). Also, relative to the in-situ observations of von Randow et al. (2004), biases in CLM - PLUS are reduced, yet not completely eliminated (Table A.2). As a consequence of the improved ΔET^1, we find that CLM - PLUS partly alleviates the positive bias in ΔLST^1_{max} compared to the MOD - Li15 data, especially in the equatorial climate zone which also reduces the RMSD in all but the Arid climate zone (Figs. 2.4 c and 2.3 c). This hints that a realistic representation of ΔET^1 is crucial for resolving the underestimated cooling effect of forests on daily maximum LST. Similarly, RMSD of ΔLST^1_{avg} decreases in the Equatorial and Arid climate zones, whereas it increases in the temperate and snow climate zones (Fig. 2.3 b). At the same time, the RMSD of ΔLST^1_{min} is only marginally increased in all climate zones (Fig. 2.3 d).

2.4 Discussion

The combination of separated soil columns and the further modifications introduced in CLM-PLUS led to substantial improvements in CLM4.5's capability to represent forest/open land contrast. Nevertheless, some biases still persist. In particular, CLM4.5 is still unable to represent the nighttime warming effect of forests in the mid-latitudes exhibited by observational data (Lee et al., 2011; Zhang et al., 2014; Vanden Broucke et al., 2015; Li et al., 2015; Alkama and Cescatti, 2016; Li et al., 2016b). Additionally, there is a remaining positive bias of ΔLST^{1}_{max} compared with MOD - Li15 even though this bias is alleviated to some extent due to the more positive ΔET^{1}. Inadequate representation or omission of several processes in CLM4.5 could be the source of these discrepancies with MOD - Li15. The biases in both ΔLST^{1}_{max} and ΔLST^{1}_{min} could be alleviated by accounting for biomass heat storage, a process which is currently disregarded in CLM4.5. Observed diurnal biomass heat storage fluxes reach an amplitude of 10–20 Wm^{-2} in the mid- and high-latitudes (McCaughey and Saxton, 1988; Lindroth et al., 2010; Kilinc et al., 2012) and 20–70 Wm^{-2} in the tropics (Moore and Fisch, 1986; Meesters and Vugts, 1996; Michiles and Gielow, 2008). Fluxes of this magnitude are sufficient to considerably alter the diurnal temperature cycle in forests and hence potentially resolve the discrepancies in ΔLST^{1}_{max} and ΔLST^{1}_{min} of CLM4.5 with MOD - Li15. While ΔET^{1} in CLM - PLUS is improved against all the considered ET data sets in almost every climate zones, some biases persist especially concerning the seasonality in the mid- and high-latitudes as well as annual mean values around the equator. In CLM - PLUS the focus was on vegetation transpiration, thereby neglecting the contribution from soil and interception evaporation. However, soil evaporation is considerably lower over forests around the equator in CLM - PLUS which might explain the remaining negative ΔET^{1} in this region. We therefore encourage additional sensitivity experiments which also focus on the other components of ET. When testing new model configurations, care should be taken that the implemented modifications do not impair other features of the model, related not only to the water but also the energy and carbon budgets. Reassuringly, we find that global ET averages are only weakly affected in the sensitivity experiment, with an average of 1.43 mm/day in CLM - BASE compared to 1.41 mm/day in CLM - PLUS. These values lie within the range of 1.2 mm/day to 1.5 mm/day estimated from surface water budgets (Wang and Dickinson, 2012). Nevertheless, it would be desirable in future studies to evaluate the biogeochemical effects of the different model configurations investigated here alongside the biogeophysical effects.

For comparison with LST data we used the radiative temperature in CLM4.5 rather than the more common T2M diagnostic which exhibits an observation-contradicting sign in CLM4.5 (compare Figs. 2.9e and 2.8). Such T2M-specific discrepancies with observations could be related to a differing definition of T2M over forests in the model and observations. For example, the differing sign of $\Delta T2M^{1}_{max}$ in climate models using CLM and the observations of Lee et al. (2011) found in Lejeune et al. (2017) might be related to the fact that T2M observations were made 2 to 15 m above the forest canopy whereas T2M of CLM4.5 lies within the forest canopy (Oleson et al., 2013). Therefore, T2M in CLM4.5 should be used with care when comparing to observations.

There are several factors which may affect the comparability of the signal extracted from our CLM4.5 simulations and the considered observational data sets. (1) the different data sources use differing LC information. For example, GLEAM uses the MOD44B product which provides the fraction of each grid cell covered by trees, non-tree vegetation, and non-vegetated land surfaces, whereas MOD - Li15 uses MCD12C1 product which provides the dominant IGBP LC type (Li et al., 2015; Martens et al., 2017). Further, the definition of forest and open land in the Li et al. 2015 dataset can be a source of model-data discrepancy. The methodology applied by Li et al. (2015) relies on the definition of a threshold (80 %) in the coverage of forest (open land) for a pixel to be classified as forest (open land). There are therefore some mixing effects between the forest and open land categories in this dataset, whereas our evaluation method isolates pure signals over forest and open land in CLM4.5. In fact, MODIS albedo retrievals were found to underestimate albedo over grass- and cropland, especially under the presence of snow, and overestimate it over forests due to the heterogeneity of LC within pixels (Cescatti et al., 2012; Wang et al., 2014). Therefore, it is possible that the magnitude of $\Delta\alpha^{\downarrow}$ is underestimated in MOD - Li15 rather than overestimated in CLM4.5. Consistently, in-situ observations of paired forest and open land sites support the higher $\Delta\alpha^{\downarrow}$ found in CLM - BASE (von Randow et al., 2004; Liu et al., 2005). (3) MODIS LST data are retrieved under clear-sky conditions only, whereas we do not mask out cloudy days in the evaluation of the CLM4.5 simulations. (4) the overpass times of the MODIS satellite system are at 1:30 am and 1:30 pm, hence not necessarily coinciding with the daily maximum and minimum LST in CLM4.5. And finally (5), the meteorological conditions within one search window of MOD - Li15 may vary between the different pixels, whereas the different PFT tiles in our CLM4.5 simulations where subject to the exact same atmospheric forcing. However, Li et al. (2015) partly accounted for this effect by applying an elevation adjustment. Moreover, they found little sensitivity of the forest minus open land signal to the size of the chosen window.

In this study, we focused on the contrast between forest and open land. However, we acknowledge that future studies should consider other types of land conversions or land management changes, as an increasing number of studies have demonstrated that other LCCs than de- or reforestation also have remarkable biogeophysical effects (e.g., Davin et al., 2014; Malyshev et al., 2015; Naudts et al., 2016; Thiery et al., 2017; Chen et al., 2018). The two new observation based data sets of Bright et al. (2017) and Duveiller et al. (2018b) assess the biogeophysical consequences of a series of different LCCs globally, thereby enabling the evaluations of the sensitivity to additional types of LC in future studies. An additional advantage of these two studies is that they both provide a signal for a complete conversion from one LC type to another (i.e. they do not rely on a coverage threshold as MOD - Li15). In our evaluation approach we focus on the local climatic impact of forests, thereby neglecting feedback mechanisms between the atmosphere and the land surface. While they appear to be relevant in many climate models (Winckler et al., 2017; Devaraju et al., 2018), their evaluation is prevented by the lack of observations at the moment.

2.5 Conclusions

In this study we evaluate the representation of the local biogeophysical effects of forests in the Community Land Model 4.5 (CLM4.5), using recently published MODIS-based observations of the albedo, evapotranspiration (ET), and land surface temperature (LST) difference between forest and nearby open land. Given the uncertainties in observation-based ET estimates, we further extend our evaluation for this variable by including data from GLEAM v3.1a and GETA 2.0. In our model evaluation we extract a local signal of forests by analysing PFT-level model output, allowing for good comparability with the high-resolution satellite observations. Further, we use a modified version of CLM4.5 which attributes a separated soil column to each plant functional type (PFT), resulting in a more realistic sub-grid contrast between forest and open land.

Overall, the lower albedo over forests in CLM4.5 is in line with the MOD - Li15 observations. However, the albedo contrast between forest and open land is somewhat more pronounced in the model. Ground observations support the stronger albedo contrast in CLM4.5, suggesting that MODIS albedo observations should be used carefully when contrasting different LC types, as satellite observations tend to retrieve a mixed signal of various LC types due to their limited spatial resolution. By suppressing lateral ground heat fluxes, the soil column separation considerably improved the representation of the impact of deforestation on daily mean and maximum LST, resulting in a good agreement with the MOD - Li15 observations. Both exhibit an overall cooling effect of forests on these variables, except for winter at latitudes exceeding 30°. Nevertheless, it appeared that the LST difference of forest minus open land in CLM4.5 tends to have a positive bias compared to observational studies. Also, it emerged that caution is required when comparing 2 m air temperature in CLM4.5 to observational data. This variable is only diagnostic in CLM4.5 and might not be conform with measurements, despite realistic LST values. The nighttime warming effect of forests in the mid-latitudes which emerged in a number of recent observational studies, is not represented by CLM4.5. The biases in the daily maximum and minimum LST signal of forests might be at least partly alleviated by accounting for heat storage in the vegetation biomass. We therefore encourage a modification of CLM which enables the representation of biomass heat storage.

Observation-based ET estimates generally agree on higher ET over forests than open land throughout the year at low latitudes and during summer at mid- and high latitudes. This was however not represented by the CLM4.5 configuration using separated soil columns. In fact, the soil column separation impaired the ET signal of forests in CLM4.5, despite improving the LST signal of forests considerably. Hence, a complete evaluation and verification of this modification of CLM4.5 should be undertaken before including it in future versions of CLM. We succeeded in attenuating the biases in ET and also daily maximum LST in a sensitivity experiment which incorporated modifications on four aspects of the parametrisation of vegetation transpiration: The root distribution, a dynamic plant water uptake instead of the current static one, the light limitation, and the maximum rate of carboxylation.

Historically the most important Land Cover Change (LCC), deforestation is still ongoing in large parts South America, Africa, and Southeast Asia. A realistic representation of the biogeophysical effects of LCC in climate models is needed as a number

of observational studies revealed that they can have a considerable impact on the local climate. An appropriate representation of the effects of LCC is not only a feature land surface models need to have in order to understand the climate of the past and project future climate, but is also a chance to achieve a more realistic simulation of processes at the land surface. To this end, the analysis of model output at PFT level can help revealing model deficiencies that otherwise would have been hidden below the veil of grid-scale aggregation.

Acknowledgments We thank Sonia Seneviratne for her comments on the analysis and are grateful for the technical support of Urs Beyerle during the course of this project. We would like to thank the two anonymous reviewers for their constructive comments which helped to improve the quality of our manuscript considerably. We acknowledge funding from the Swiss National Science Foundation (SNSF) and the Swiss Federal Office for the Environment (FOEN) through the CLIMPULSE and Hydro-CH2018 projects and from the European Union's Horizon 2020 research and innovation programme under grant agreement No 641816 (CRESCENDO).

3

Biomass heat storage dampens diurnal temperature variations in forests

*Envrionmental Research Letters, 14, 084026, https://doi.org/10.1088/1748-9326/ab2b4e**, Ronny Meier[1], Edouard L. Davin[1], Sean C. Swenson[2], David M. Lawrence[2], Jonas Schwaab[1]

Abstract Observational evidence suggests that compared to non-forested areas, forests have a cooling effect on daytime Land Surface Temperature (LST) and a warming effect on nighttime LST in many regions of the world, thus implying that forests dampen the diurnal temperature range. This feature is not captured by current climate models.
Using the Community Land Model 5.0 (CLM5.0), we show that this diurnal behaviour can be captured when accounting for Biomass Heat Storage (BHS). The nighttime release of energy absorbed by the vegetation biomass during the day increases both nighttime LST and ambient air temperature in forested regions by more than 1 K. The daytime cooling is weaker than the nighttime warming effect, because the energy uptake by the biomass is compensated by a reduction in the turbulent heat fluxes during day. This diurnal asymmetry of the temperature response to BHS leads to a warming of daily mean temperatures, which is amplified during boreal summer warm extremes. Compared to MODIS, CLM5.0 overestimates the diurnal LST range over forested areas. The inclusion of BHS reduces this bias due to its dampening effect on diurnal LST variations. Further, BHS attenuates the negative bias in the nighttime LST difference of forest minus grassland and cropland, when compared to MODIS observations. These results indicate that it is essential to consider BHS when examining the influence of forests on diurnal temperature variations. BHS should thus be included in land surface models used to assess the climatic consequences of land use changes such as deforestation or afforestation.

*This publication was slightly changed from its original version to ensure consistency throughout this thesis. Figures and numbers have been updated following the correction of a coding error found after publication (https://doi.org/10.1088/1748-9326/ab4a42).

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]National Center for Atmospheric Research, Boulder, USA

3.1 Introduction

Forests play a critical role in the climate system by regulating land-atmosphere exchanges of greenhouse gases, energy, and water (e.g., Pongratz et al., 2010; de Noblet-Ducoudré et al., 2012). The influence of forests on climate can be divided into two categories of processes: (1) the globally acting biogeochemical processes, representing the release or sequestration of greenhouse gases, and (2) the biogeophysical processes, representing direct alterations of the local energy and/or water budget. While the first category of processes appears to be dominant on a global scale (Pongratz et al., 2010), the latter alter the climate considerably at local to regional scale and are thus essential in determining the influence of forests on local climate (Lee et al., 2011; Li et al., 2015; Li et al., 2016b; Bright et al., 2017; Duveiller et al., 2018b).

In the absence of snow, forests are observed to have a local cooling impact on both Land Surface Temperature (LST) and above-canopy air temperature during the day compared to open land (i.e., grassland and cropland; Lee et al., 2011; Li et al., 2015; Alkama and Cescatti, 2016; Duveiller et al., 2018b). During nighttime, forests tend to be only slightly colder than open land in tropical and subtropical areas and even warmer than open land in the mid-latitudes. Hence, forests tend to dampen the Diurnal Temperature Range (DTR) everywhere but the boreal regions (Lee et al., 2011; Duveiller et al., 2018b). In the mid-latitudes, the sign of the temperature difference of forest minus open land even changes during the diurnal cycle, from a negative daytime difference to a positive one at night (Lee et al., 2011; Zhang et al., 2014). This feature is missing in all the climate models analysed in the LUCID, CMIP5 and LUCAS intercomparisons (Lejeune et al., 2017; Davin et al., 2020). A considerable fraction of the models even exhibits a warming effect of forests on daily maximum 2 m air temperature (T2M) and a cooling effect on daily minimum T2M, opposing observations completely. Such biases could be partly related to the calculation of T2M in models, which does not necessarily correspond to the temperature 2 m above the canopy in forests (Meier et al., 2018; Winckler et al., 2019b). However, the nighttime warming by forests is still not represented in CLM4.5 when considering LST, which is more directly comparable to remote sensing observations (Meier et al., 2018). This systematic model bias suggests that land surface models are either missing or are poorly representing an important process affecting the energy redistribution at the land surface.

There is a general consensus about the processes leading to the daytime cooling effect of forests compared to open land. Namely, forests have higher surface roughness, which results in stronger turbulent heat fluxes, as well as higher evaporative fraction, which is associated with more evaporative cooling (Davin and de Noblet-Ducoudré, 2010; Lee et al., 2011; Vanden Broucke et al., 2015; Schultz et al., 2017). Indeed, reducing the negative bias in the ET difference of forest minus open land in CLM4.5 also attenuated the positive bias in the daytime LST difference between these Land Cover (LC) types (Meier et al., 2018). On the other hand, little is known about the mechanism behind the nighttime warming effect of forests in the mid-latitudes. Lee et al. (2011) hypothesized that under stable conditions warm air from the planetary boundary layer is transported more effectively towards the land surface over forests, due to their higher surface roughness compared to the one of open land. Vanden Broucke et al. (2015)

observed that the nighttime warming of forests is not only related to sensible heat fluxes but also to more incoming longwave radiation. They hypothesized that the latter flux could be explained by (1) a stronger greenhouse gas effect over forests due to a moister boundary layer, (2) higher aerosol loading over forests, or (3) the presence of warmer air due to more turbulent mixing over forests (i.e., the higher incoming longwave radiation is a side effect of the stronger turbulent mixing over forests). An additional explanation for the nighttime warming effect of forests could be that energy accumulated in the plant biomass during the day is released into the canopy space during the night (Schultz et al., 2017; Meier et al., 2018). Due to the high amounts of biomass in forests compared to grasslands or croplands, Biomass Heat Storage (BHS, i.e. heat storage in the biomass itself) is likely much larger over forests and could thus be a driver behind the nighttime warming effect of forests. Indeed, in-situ observational studies indicate that the diurnal amplitude of the Biomass Heat Flux (BHF; i.e., the change of energy stored in the biomass) can be substantial (Aston, 1985; Moore and Fisch, 1986; McCaughey and Saxton, 1988; Meesters and Vugts, 1996; Vogt et al., 1996; Silberstein et al., 2001; Meyers and Hollinger, 2004; Oliphant et al., 2004; Michiles and Gielow, 2008; Garai et al., 2010; Lindroth et al., 2010; Kilinc et al., 2012; Burns et al., 2015). While the BHF is negligible when integrated over longer time scales, it typically exhibits diurnal amplitudes of 15-75 $W\,m^{-2}$ in mature forests (e.g., Oliphant et al., 2004; Haverd et al., 2007; Lindroth et al., 2010; Kilinc et al., 2012; Burns et al., 2015). A flux of this magnitude appears sufficient to considerably alter diurnal temperature variation over forests. Hence, several land surface model have already included the process BHS (e.g., Verseghy et al., 1993; Samuelsson et al., 2011; Boone et al., 2017; Heidkamp et al., 2018).

Still, the relevance of BHS for the local climate has to our knowledge not been assessed at a global scale. In this study, we therefore investigate, how BHS affects diurnal LST and T2M variations over forested regions using the Community Land Model 5.0 (CLM5.0). Swenson et al. (2019) recently introduced a scheme simulating energy storage in leaves and stems in CLM5.0. Here, we extend the BHS scheme of Swenson et al. (2019), by coupling BHS to the biomass carbon stocks simulated by the model with an active carbon and nitrogen cycle. As a consequence, the leaf area index, leaf biomass, stem biomass, and vegetation height are diagnosed directly from the simulated vegetation carbon pools instead of relying on uncertain parameter values, thereby improving the internal consistency within the model. We then present a first global estimate of the diurnal amplitude of the BHF and compare the simulated BHF to in-situ studies to assess the realism of the BHS scheme. In a second step, we assess how BHS affects temperatures at grid cell level and evaluate the modeled DTR with remote sensing observations from the MODerate resolution Imaging Spectroradiometer (MODIS) system. Finally, we investigate how BHS affects the sensitivity of LST to LC at sub-grid scale, which we compare to two MODIS-based data sets (Li et al., 2015; Duveiller et al., 2018b), following the evaluation strategy in Meier et al. (2018).

3.2 Methods and data

3.2.1 Model description

CLM5.0 is the land component of the Community Earth System Model (CESM2; Lawrence et al., 2019). CLM5.0 simulates thermodynamic processes, such as absorption, reflection, and emission of shortwave and infrared radiation, sensible and latent heat fluxes from the vegetation and soil, and heat storage in the soil column. The hydrology of the land surface considers infiltration, runoff, canopy interception, and evapotranspiration, distinguishing between the water, snow, and ice phase. Further, CLM5.0 can represent the exchange of carbon and nitrogen via processes such as photosynthesis, autotrophic and heterotrophic respiration, litterfall, nitrogen deposition, nitrogen fixation, and nitrogen mineralization, denitrification, and fire.

Recently, Swenson et al. (2019) implemented a BHS scheme to a post-release version of CLM5.0 which allows the representation of heat storage in stems and leaves. An additional vegetation temperature, the stem temperature, was introduced in this scheme, besides the already existing bulk canopy temperature (representing the temperature of the leaves). As a consequence, the energy balance is solved both for the leaf and the stem.

The leaf energy balance is closed at each time step, under the assumption that leaf temperature (T_{leaf}) is in balance with the leaf energy fluxes given the high surface area/volume ratio (i.e., the leaf temperature is iteratively adjusted until the energy balance is closed):

$$\overrightarrow{S}_{leaf} + \overrightarrow{L}_{leaf}(T_{leaf}) - H_{leaf}(T_{leaf}) - \lambda E_{leaf}(T_{leaf}) - C_{leaf}\frac{dT_{leaf}}{dt} = 0 \quad (3.1)$$

, where $\overrightarrow{S}_{leaf}$ and $\overrightarrow{L}_{leaf}$ are the net solar and the net longwave radiation absorbed by the leaves, H_{leaf} and λE_{leaf} are the sensible and latent heat fluxes from the leaves, and C_{leaf} is the heat capacity of the leaves.

The energy balance of the stem is computed after calculating T_{leaf} as follows:

$$\overrightarrow{S}_{stem} + \overrightarrow{L}_{stem}(T_{stem}) - H_{stem}(T_{stem}) = C_{stem}\frac{dT_{stem}}{dt} \quad (3.2)$$

, where $\overrightarrow{S}_{stem}$ and $\overrightarrow{L}_{stem}$ are the net solar and the net longwave radiation absorbed by the stem (Eqs. 6 and 7 in Swenson et al., 2019) , H_{stem} is the sensible heat flux from the stem (Eq. 12 in Swenson et al., 2019), C_{stem} is the heat capacity of the stem, and T_{stem} is the stem temperature. Hence, the energy surplus (deficit) of the stem (i.e., the left-hand side of Eq. 3.2) is then taken up (released) by the stem and the stem temperature is adjusted accordingly. The BHF represents the change in energy content of the biomass and is calculated as follows:

$$BHF = C_{leaf}\frac{dT_{leaf}}{dt} + C_{stem}\frac{dT_{stem}}{dt} \quad (3.3)$$

We introduce a few modifications and different parameter choices compared to the original implementation of Swenson et al. (2019). In addition to the leaf area index, the

BHS module requires a number of plant functional type-specific parameters. We assume a dry wood density of 500 kg m^{-3} (Swenson et al., 2019). For the fraction of the biomass that is water (f_w) we use a value of 0.7 for the leaves (Bonan et al., 2018) and a value of 0.5 for the stem, which is representative for the outer section of the stem (Herrington, 1969). For the tree number density, the value of the closest biome of Crowther et al. (2015) is chosen for each Plant Functional Type (PFT). Further, we couple the BHS scheme to the prognostic carbon and nitrogen module in CLM5.0 (referred to as the Bgc-mode in the model community). This allows to calculate the dry leaf mass (M_{leaf}, kg m^{-2}) and dry stem mass (M_{stem}, kg m^{-2}) based on the simulated carbon pools as follows:

$$M_{leaf} = CP_{leaf} * b_{carbon} \tag{3.4}$$

$$M_{stem} = CP_{stem} * b_{carbon} \tag{3.5}$$

, where CP_{leaf} and CP_{stem} are the leaf and stem carbon pools, which are prognostically simulated by CLM5.0, and b_{carbon} is the ratio of dry biomass to carbon with a constant value of 2 kg/kgC (Bonan et al., 2018). M_{leaf} and M_{stem} are subsequently used to compute C_{leaf} and C_{stem}, respectively after Eqs. 9 and 22 in Swenson et al. (2019). This approach enables the calculation of a breast-height diameter (D_{bh}) from M_{stem} instead of assuming a globally constant value by combining Eqs. 10 and 11 of Swenson et al. (2019):

$$D_{bh} = 2 * \sqrt{\frac{M_{stem}}{\pi\, h_{tree}\, N_{tree}\, \rho_{wood}}} \tag{3.6}$$

, where h_{tree} is the tree height, N_{tree} the tree number density, and ρ_{wood} the dry wood density. As in Swenson et al. (2019), the resistance to heat transfer between the interior of the tree and the tree surface, r_{bole}, is linked to the D_{bh} as follows:

$$r_{bole} = r_w * D_{bh} \tag{3.7}$$

, where r_w is resistance to heat transfer per meter of stem diameter. We use r_w as a tuning parameter to achieve realistic BHF values and arrive at the same value of 1000 s m^{-2} as in Swenson et al. (2019). In this study however, r_{bole} may vary spatially and temporally as the D_{bh} is linked to the stem carbon pool. Additionally, the Bgc-module computes the leaf area index, stem area index*, and vegetation height based on the vegetation carbon pools instead of using values estimated from remote sensing observations (Lawrence et al., 2018). The source code used in the simulations is available at https://github.com/RonnyMeier/ctsm/tree/ERL-106754_corrected.

3.2.2 Simulation setup

The simulations presented in this study are analysed over the period of 2002 to 2010 (since the observational data described in the section below start from 2002 and the atmospheric forcing is available only until 2010). They were run in offline mode, forced by the

*This variable is called stem area index in CLM. However, this variable was initially called "the stem and dead leaf area index" when introduced in CLM (Zeng et al., 2002) and exhibits a distinct seasonal cycle with a peak for deciduous trees in autumn in the higher latitudes. As such, the name "stem area index" is misleading.

Global Soil Wetness Project (GSWP3) data set, at 0.5° spatial resolution (Kim, 2017). The percentage of the different PFTs within each grid cell is derived from MODerate resolution Imaging Spectroradiometer (MODIS) observations, as described in Lawrence and Chase (2007), and is kept constant throughout the simulation at the coverage of the year 2000 (Fig. B.1). We use separated soil columns in our simulations to suppress unrealistically large lateral ground heat fluxes between different PFTs (Schultz et al., 2016; Meier et al., 2018). To isolate the effect of BHS, we run a control simulation with no BHS (CLM - CTL). The second simulation, CLM - BHS, is run in the exact same configuration but with an active BHS scheme. As in Swenson et al. (2019), we raise the upper cap of the Monin-Obukhov stability parameter in the surface layer from 0.5 to 100 to assess the full impact of BHS in both simulations (thereby making the atmospheric stability virtually unconstrained). Both the CLM - CTL and CLM - BHS simulations start from the same initial state and are run for the years 1997 to 2010 (Note that this experimental design mutes feedbacks arising from BHS on the simulated carbon stocks themselves). There is therefore an additional 5 years of spin up before the analysis period of 2002 to 2010. The initial state in 1997 is retrieved by running the model for 146 years at 0.5° spatial resolution, starting from an already spun-up, pre-industrial state, which was interpolated from the original resolution of 0.9° × 1.25° to 0.5° × 0.5° resolution. For this additional spin up, we first cycle five times through the atmospheric forcing data of the years 1901 to 1910 for the period of 1851 to 1900. After, we force the model with the reanalysis data of the years 1901 to 1996. During this spin up, BHS remains inactive. Fig. B.2 illustrates the resulting aboveground biomass of different LC types.

3.2.3 Observational data

We evaluate the modeled LST with MODIS observations both at grid cell level and in terms of the sub-grid difference between forest minus open land. The MODIS instruments aboard the satellites Aqua and Terra provide LST measurements at approximately 01:30 and 13:30 solar time at 0.05° × 0.05° resolution. For the evaluation at grid cell level we use the data of the monthly MYD11C3 data product from July 2002 to December 2010. After masking out observations with a reported LST error estimate larger than 1 K and/or an emissivity error estimate larger than 0.01 (as in Li et al., 2015) and discarding pixels with a land fraction lower than 80 %, the original data is averaged to the model resolution of 0.5° × 0.5°. The resulting monthly values are then used to derive a multi-year monthly average. From this multi-year monthly average we then calculate the multi-year total average, excluding pixels with no valid data for at least one month.

The model performance in terms of its representation of the local LST difference of forest minus open land is evaluated using the two MODIS-based data sets of Li et al. (2015) and Duveiller et al. (2018b), subsequently called MOD - Li15 and MOD - Du18. These data sets compare LST over different LCs within moving windows of 9 by 5 pixels or 5 by 5 pixels to infer a LST difference of forest minus open land at roughly 01:30 and 13:30 solar time. Given the relatively high resolution of the two data sets, it is a reasonable assumption that the different pixels within the moving window experience similar climatic conditions. Thus, the data sets are assumed to capture the local impact of a conversion from forest to open land.

MOD - Li15 and MOD - Du18 differ in methodology, time frame, and the MODIS products utilized. MOD - Li15 uses the MYD11C2 product extracting data from the period 2002-2012. On the other hand, MOD - Du18 employs MYD11C3 data from 2008 to 2012. The MYD11C2 and MYD11C3 products are both aggregated from daily LST observations (MYD11C1) to 8 day average values in the case of MYD11C2 and monthly average values in the case of MYD11C3 (Wan and Hulley, 2015a; Wan and Hulley, 2015b). Although the two data sets exhibit a comparably high spatial resolution, the observed pixel usually still comprises different LC types. Thus, it is challenging to isolate a LST over forest or open land only. MOD - Li15 defined a threshold of 80 % in the coverage of a certain pixel by forest or open land for the pixel to be classified as forest or open land, respectively, and calculated the difference between the forest minus the open land pixels within search windows of 9 by 5 pixels. The LST of the different pixels was corrected for elevation differences using a lapse rate which was inferred from the MODIS observations. Further, the comparison samples were masked for an elevation difference of less then 500 m. The more recent MOD - Du18 data set on the other hand used a multiple linear regression model within windows of 5 by 5 pixels to establish a relationship between the fraction of different vegetation types and LST. Additionally, Duveiller et al. (2018b) applied stricter criteria to mask out pixels with too high elevation variability but did not apply an elevation adjustment to the LST data.

3.2.4 Model evaluation

The data sets of Li et al. (2015) and Duveiller et al. (2018b) depict the local LST difference of forest minus open land under similar atmospheric forcing. We extract a comparable signal from our simulations by calculating the sub-grid difference of forest minus open land in the variables of interest following the approach of Malyshev et al. (2015). In this approach the average in a certain variable over the forest tiles within a grid cell is subtracted from the average over the open land tiles within the same grid cell (i.e. under the same atmospheric forcing). Note that this difference is comparable to the signal of an open-land-to-forest transition (i.e., afforestation or reforestation) rather than a deforestation signal. A more detailed description of this methodology can be found in Meier et al. (2018).

The LST is calculated as the weighted average between the leaf temperature (T_{leaf}) and the ground temperature (T_{grnd}):

$$LST = e_v * T_{leaf} + (1 - e_v) * T_{grnd} \tag{3.8}$$

The vegetation emissivity, e_v, is calculated as in Eq. 4.20 in Lawrence et al. (2018):

$$e_v = 1 - e^{-(L+S)/\bar{\mu}} \tag{3.9}$$

, where L and S are the leaf area index and stem area index, respectively, and $\bar{\mu}$ is the average inverse optical depth for longwave radiation (with a value of 1). T_{grnd} is inferred from the snow temperature (T_{snow}), the temperature of the top soil layer (T_{soil}), and the surface water temperature (T_{H2O}) as follows:

$$T_{grnd} = \left(f_{snow} * (T_{snow})^4 + (1 - f_{snow} - f_{H2O}) * (T_{soil})^4 + f_{H2O} * (T_{H2O})^4 \right)^{1/4} \tag{3.10}$$

, where f_{snow} and f_{H2O} are the fraction of the ground covered by snow and liquid water, respectively. Since the MODIS observations are available at around 01:30 and 13:30 solar time, LST values in the model are output at the corresponding time steps. As in Duveiller et al. (2018b), we assume that the LST difference at 01:30 and 13:30 reasonably represent the daily minimum and maximum LST difference (we find that this assumption is reasonable in the context of CLM5.0 output). Hence, ΔLST^{1}_{min} and ΔLST^{1}_{max} refer to the LST difference of forest minus open land at 01:30 and 13:30, respectively. In addition, we analyse the effect of forests on the DTR, which can be calculated as the difference of ΔLST^{1}_{max} minus ΔLST^{1}_{min} (Duveiller et al., 2018b). Accordingly, this variable is called ΔDTR^{1}. While the advantage of LST is that there exist observations with a global coverage, T2M is used more frequently as a temperature metric and is thus more readily interpreted. We therefore also analyse the impact of BHS on T2M at grid cell level (Eq. 5.58 in Lawrence et al., 2018). T2M is a diagnostic variable in CLM5.0, representing the air temperature 2 m above the apparent sink of sensible heat (i.e., 2 m above the sum of the roughness length for sensible heat and the displacement height). Note that this height normally lies below the canopy for forests but above the canopy for shorter vegetation, such as grassland or cropland.

3.3 Results

3.3.1 Magnitude of simulated BHFs

The diurnal range of the BHF (daily maximum of BHF minus daily minimum of BHF) simulated in CLM - BHS generally lies in the range 30-75 $W\,m^{-2}$ and even exceeds 75 $W\,m^{-2}$ in few densely forested areas in the tropics and the northern mid-latitudes (Fig. 3.1). The simulated BHFs are small in areas with low forest coverage (cf. Figs. 3.1 a and B.1 a), since the main contribution to the BHF comes from the stem of trees, which have a much larger storage capacity than the leaves. The simulated diurnal range of the BHF is of similar magnitude as found in in-situ studies (Fig. 3.1 b; McCaughey and Saxton, 1988; Moore and Fisch, 1986; Meesters and Vugts, 1996; Silberstein et al., 2001; Oliphant et al., 2004; Haverd et al., 2007; Michiles and Gielow, 2008; Garai et al., 2010; Lindroth et al., 2010; Kilinc et al., 2012). However, the spatial variability in the observations is considerable and much higher than in the model, which can be related to several factors. First, the observational studies utilize differing methodologies to estimate BHS. Second, those studies are conducted for a limited time period and are thus impacted strongly by the meteorological conditions during the measurement campaign. Finally, the geometry of the vegetation at the site has an impact on BHS. An example for the latter factor is the study of Kilinc et al. (2012), which was conducted at a site with relatively few but enormous trees (average mass of 29 Mg per tree), thereby resulting in a relatively small BHF when compared to the large amount of aboveground biomass (AGB). The opposite is the case in Burns et al. (2015), where the exceptionally high tree number density of 0.4 m^{-2} is contributing to a comparably high BHF range. As can be seen in Fig. 3.1 (b), CLM - BHS can not capture such exceptionally large BHFs, which is related to two reasons: First, the land surface characterization in the simulation is not

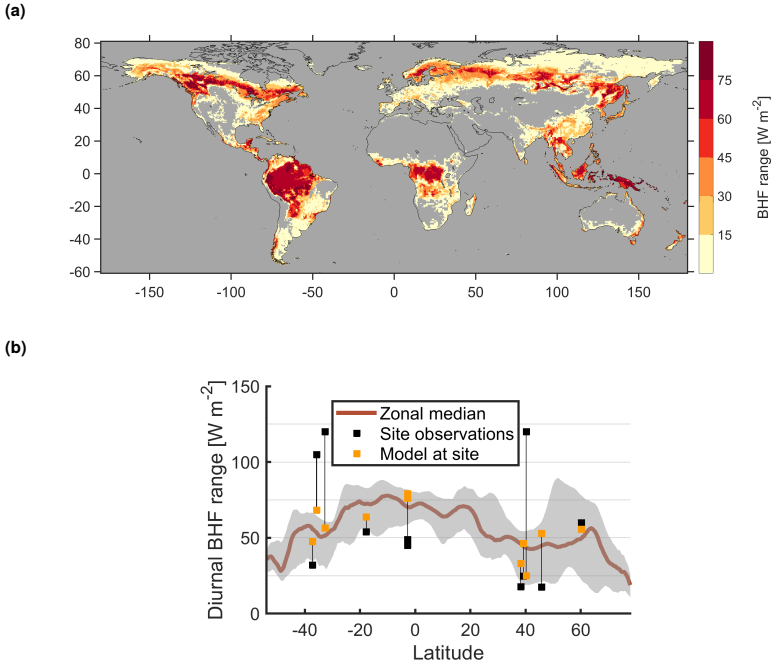


Figure 3.1: *Panel (a) dipslays the diurnal range of the biomass heat storage fluxes averaged over all plant functional types (PFTs, including shubs, crops, and grassland). Panel (b) shows the zonal median (brown line) and the range between the 10 % and 90 % percentile (grey shaded area) of the diurnal range averaged over all forest PFTs. Black squares depict different observational studies and the orange squares corresponding modeled values connected by a black string (for more information see Table B.1).*

necessarily representative for a specific site. For example, the tree number density for temperate evergreen needleleaf forest reported in Crowther et al. (2015) is much lower than the observed value at the site of Burns et al. (2015) (0.03 m^{-2} versus 0.4 m^{-2}). Second, the atmospheric conditions observed at a site during a measurement can differ considerably from the global GSWP3 data set. In fact, the BHS scheme implemented in CLM5.0 can capture the BHFs and diurnal temperature variations observed in Burns et al. (2015), when the model is forced with observed vegetation parameters and atmospheric conditions (Swenson et al., 2019).

3.3.2 Impact of BHS at grid cell level

The comparison of the CLM - BHS and the CLM - CTL simulations during boreal summer (JJA) shows that BHS has a considerable impact on T2M in areas with a large fraction covered by forests (cf. Figs. 3.3 and B.1 a). The diurnal T2M impact induced by BHS is

asymmetric: While the daily minimum T2M often increases by more than 1 K, the daily maximum cooling effect of BHS rarely exceeds 0.1 K (Fig. 3.3 a and b). As a consequence, daily average T2M tends to be higher in CLM - BHS. This warming frequently exceeds 0.4 K both for daily mean T2M and daily mean LST (see Fig. B.3 for LST). BHS also has important implications regarding extreme conditions. In fact, the warming effect of BHS during the 5 % hottest days of JJA is stronger in most regions than the effect on the mean, frequently exceeding 1 K over forested areas (Fig. 3.3 d and Fig. B.4 c). In addition, the implications of BHS vary with season. For example, the daytime cooling effect of BHS is more pronounced during winter in the northern latitudes, resulting in a weaker increase of daily mean T2M and daily mean LST due to BHS during this season (Figs. B.5 and B.6). The temperature impact of BHS displayed in Figs. 3.3, B.3, B.5, and B.6 originates mostly from the forest PFTs. When considering only the forest PFTs, BHS often increases the daily minimum T2M by more than 2 K (Fig. 3.4 b). On the other hand, the impact of BHS over the open land PFTs is marginal, due to the negligible storage capacity of leaves (Fig. 3.4 d).

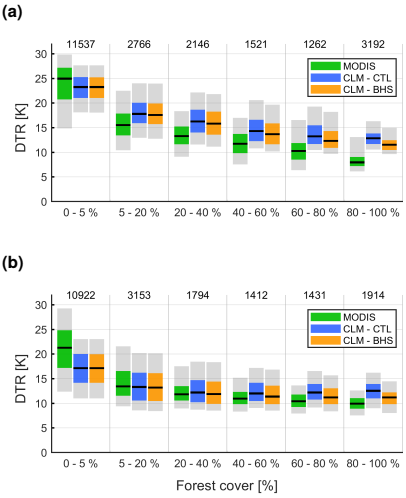


Figure 3.2: *Grid-cell-level annual mean diurnal temperature range in MODIS (green), CLM - CTL (blue), and CLM - BHS (orange) binned by different percentages of forest coverage. Displayed are the median (black line), interquartile range (coloured area), and the range from the 5 % to the 95 % quantile (shaded area) between (a) 30° S and 30° N and (b) 30° N/S and 60° N/S. The numbers above the boxplots indicate the number of grid cells belonging to the respective category.*

The increase in nighttime temperatures and decrease in daytime temperatures resulting from BHS, dampen diurnal temperature variations in forested regions leading to a better agreement with the MODIS data in these regions, as illustrated in Fig. 3.5. The MYD11C3 data show a clear signature of decreasing DTR (LST_{max} - LST_{min}) with increasing forest fraction (Fig. 3.2). The highest DTR-values are observed in desert regions, where the DTR frequently exceeds 25 K, whereas the DTR in grid cells with a considerable amount of forest coverage typically lies in the range of 8 to 15 K (Fig. 3.5 a). In comparison to MODIS, CLM - CTL tends to exhibit negative bias in DTR over regions with low forest coverage, and a positive bias over regions with higher forest coverage both in the tropics and subtropics as well as in the mid-latitudes (Figs. 3.5 b and 3.2). This finding is supported by a pattern correlation coefficient of 0.48 between the DTR bias of CLM - CTL (Fig. 3.5 b) and the forest fraction in the LC data (Fig. B.1 a). BHS reduces part of the positive DTR bias over forested regions, by dampening the DTR (Fig. 3.5 c, d, and e). The relative reduction of the DTR bias in forested regions is most distinct in the northern mid-latitudes, where CLM - CTL typically overestimates the DTR by 1 to 5 K. The relative improvement is more moderate in tropical regions, as CLM - CTL frequently overestimates the DTR by more than 5 K, which is not alleviated completely

by BHS. Therefore, the overall tendency to overestimate the DTR in forested regions and underestimate the DTR in sparsely vegetated regions partly persists in CLM - BHS, despite these improvements (Fig. 3.2).

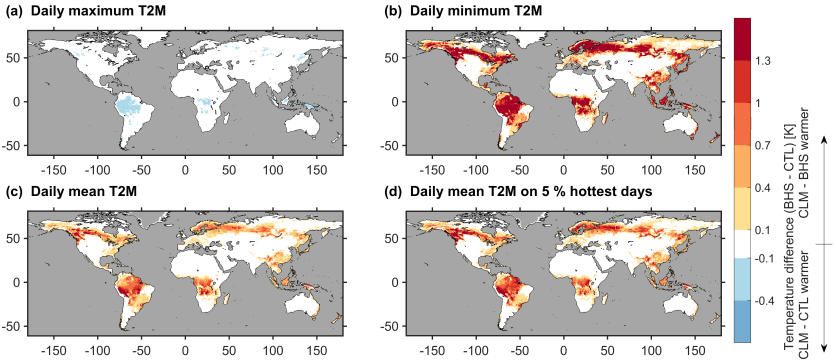


Figure 3.3: *Difference of CLM - BHS minus CLM - CTL during boreal summer (JJA) in (a) daily maximum, (b) daily minimum, and (c) daily average 2 m air temperature. Panel (d) displays the daily average 2 m air temperature difference of CLM - BHS minus CLM - CTL during JJA averaged over the days when the atmospheric temperature exceeded its 95 % percentile.*

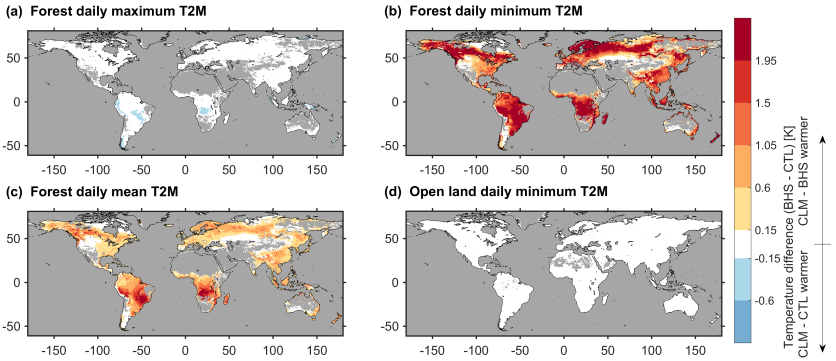


Figure 3.4: *Difference during boreal summer (JJA) of CLM - BHS minus CLM - CTL in (a) daily maximum, (b) daily minimum, and (c) daily average 2 m air temperature over the forest plant functional types (PFTs). Panel (d) displays the daily minimum 2 m air temperature difference of CLM - BHS minus CLM - CTL during JJA averaged over the grassland and cropland PFTs.*

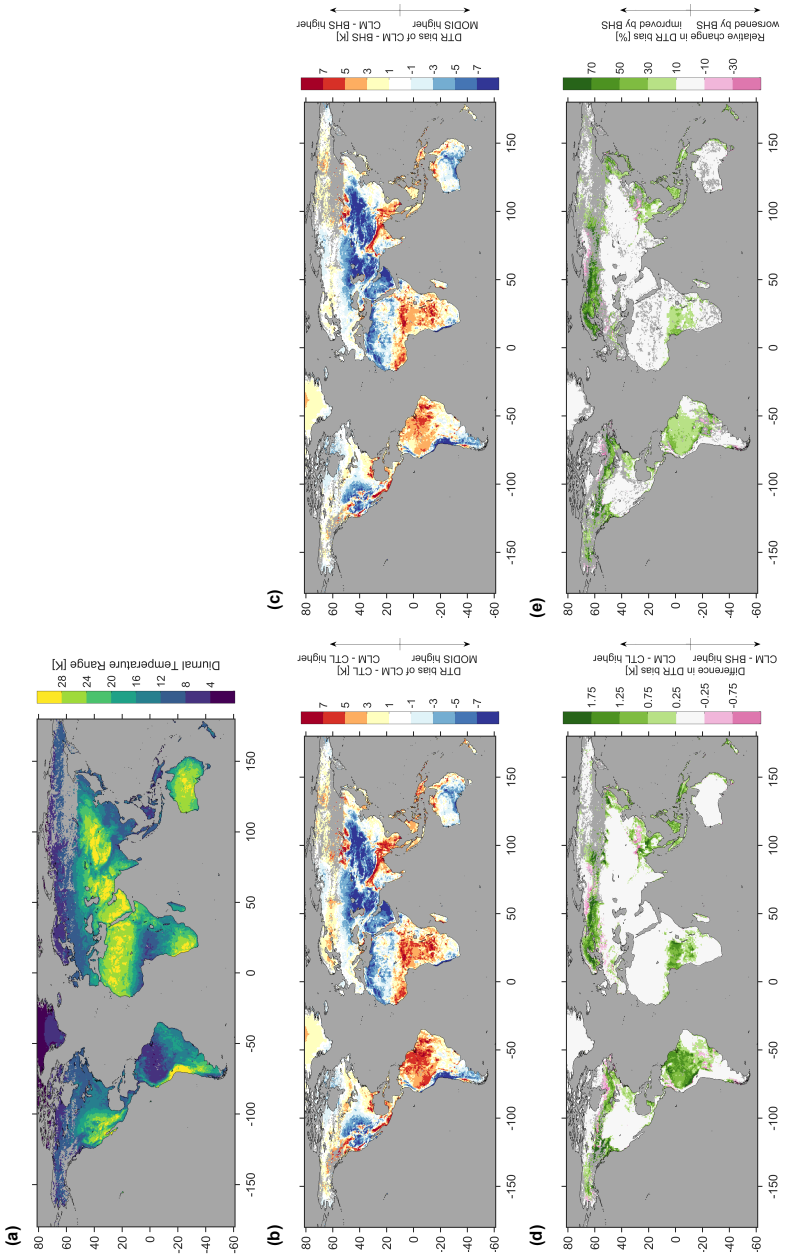


Figure 3.5: ***Comparison of diurnal temperature range (DTR) in MODIS and CLM5.0:*** *Panel (a) displays the mean DTR in the MODIS observations. Panels (b) and (c) show the bias in DTR of CLM - CTL and CLM - BHS compared to MODIS. Panel (d) shows the improvement in the DTR bias (i.e., absolute of panel (b) minus absolute of panel (c)) and panel (e) the improvement relative to the DTR bias of CLM - CTL due to BHS (i.e., panel (d) devided by absolute of panel (b)). For relative improvement regions with a bias smaller than 1 K have been masked out.*

3.3.3 Impact of BHS on LST sensitivity to LC

In this section we evaluate how BHS affects the LST sensitivity to LC, by contrasting the local difference of forest minus open land in LST_{max} and LST_{min} of the CLM5.0 simulations with the two MODIS-based datasets of Li et al. (2015) (MOD - Li15) and Duveiller et al. (2018b) (MOD - Du18). ΔLST^{l}_{min} is on average in the order of 1 K higher in CLM - BHS as compared to CLM - CTL, resulting in a considerable reduction of the global root-mean-squared deviation (RMSD) in comparison to both observational data sets (Fig. 3.6). In contrast to the clearly negative ΔLST^{l}_{min} in CLM - CTL, ΔLST^{l}_{min} of CLM - BHS is close to zero in tropical regions as observed in the MODIS-based products. However, there is still some disagreement on the nighttime signal of forests between CLM - BHS and the observations in the extra-tropics. The nighttime warming by forests between 20° N and 55° N is sometimes too weak in the BHS simulation, especially in comparison to MOD - Li15 (Fig. 3.6 d and h). However, the nighttime warming by forests at these latitudes is less pronounced in MOD - Du18 than in MOD - Li15, highlighting a certain degree of uncertainty in the observations. Similarly, CLM does not capture the distinctly positive ΔLST^{l}_{min} between 25° S and 40° S which is observed in MOD - Li15, whereas the BHS simulation agrees with MOD - Du18 overall at these latitudes. At high-latitudes on the other hand, BHS impairs the already positively-biased ΔLST^{l}_{min}. During the day, BHS consistently decreases ΔLST^{l}_{max} by about 0.1 K, which causes only a small change in RMSD (Fig. B.8). In offline simulations, BHS appears therefore mostly relevant for the LC sensitivity of the model during night, whereas its effect is less distinct during the day.

3.4 Discussion and conclusions

3.4.1 Diurnal asymmetry of BHS impact on temperature

We found that BHS affects daytime temperature much less than nighttime temperature, although roughly the same amount of energy absorbed by the vegetation during the day is released during night (i.e. the daily average BHF is close to zero).We hypothesize that this diurnal asymmetry of the temperature impact is related to a different structure of the surface layer during daytime compared to the nighttime. During the day, the surface layer tends to be unstable and the turbulent heat fluxes are directed from the land surface towards the atmosphere. An initial reduction of the surface temperature due to the energy uptake by the vegetation decreases the surface layer instability and thus also the turbulent heat fluxes. The lower turbulent

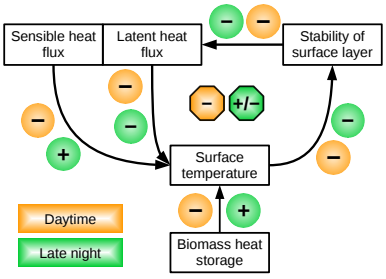


Figure 3.7: *Sketch of the daytime (orange) and and late night (green) feedback from turbulent heat fluxes induced by biomass heat storage. A positive (negative) sign indicates that an increase in the first variable would increase (decrease) the second variable. The octagons with the black edge in the middle of the loop indicate the sign of the entire feedback loop.*

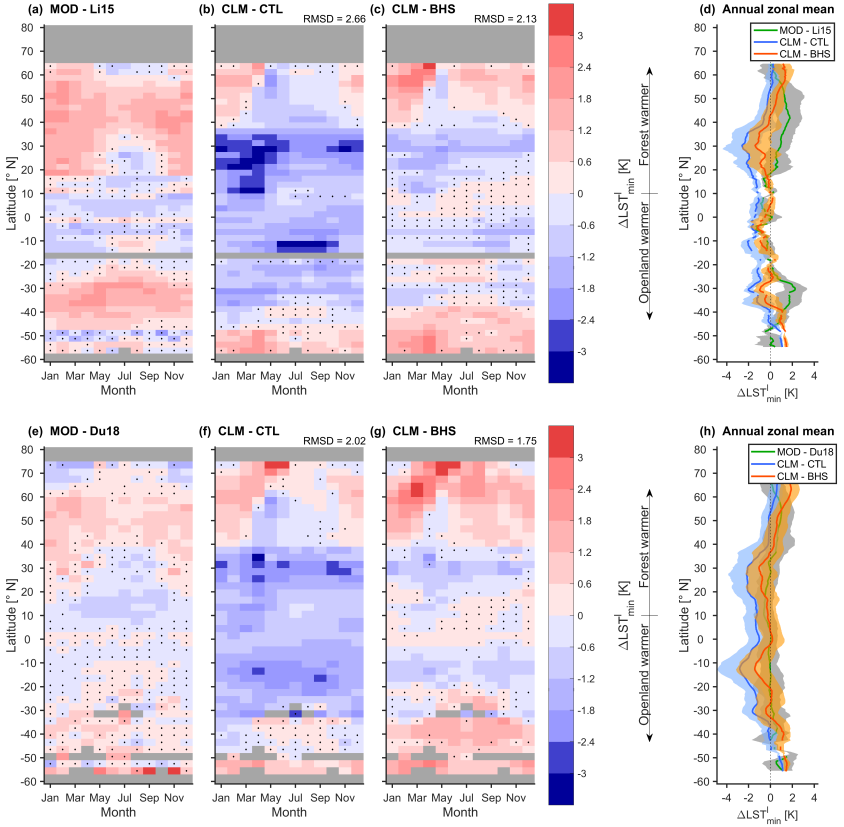


Figure 3.6: ***Nighttime land surface temperature difference between forest minus open land (ΔLST$^{l}_{min}$):*** *Seasonal and latitudinal variations of ΔLST$^{l}_{min}$ in (a) the MOD - Li15 observations, (b) CLM - CTL, and (c) CLM - BHS. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. The numbers next to the titles are the area-weighted spatiotemporal root-mean-squared deviation of the respective simulation against the MOD - Li15 data set (Meier et al., 2018). Panel (d) shows the zonal annual mean of MOD - Li15 (green, range between the 10th and 90th percentiles in grey), CLM - CTL (blue, range between the 10th and 90th percentiles in blue), and CLM - BHS (red, range between the 10th and 90th percentiles in orange). Note that on this subfigure results have been smoothed latitudinally with a simple moving average over 4°. The same with MOD - Du18 in panels (e), (f), (g), and (h). For values over different latitudinal bands see Fig. B.7.*

heat fluxes lead to a higher surface temperature (because less energy is transported away from the surface) and form therefore a negative feedback to the initial temperature decrease (Fig. 3.7). In other words, the energy uptake by BHS during day is mainly compensated by a reduction of the turbulent heat fluxes. During the late night on the other hand, the surface layer is often stable. Therefore, the energy released from the biomass is compensated less efficiently by increased turbulent heat fluxes. Further, the sensible heat flux can be directed towards the land surface during the late night. In this case, the increase in surface temperature due to BHS reduces the surface layer stability and the suppression of turbulent mixing, resulting in more intrusion of relatively warm boundary layer towards the surface. Consequently, the feedback from the surface layer to the initial surface temperature increase can even be an additional warming of the surface (Fig. 3.7).

This behaviour can be seen in point scale simulations in the tropics at the site of Michiles and Gielow (2008). As long as there is solar irradiance, the energy taken up/released by the vegetation is compensated for to a large extent by a reduction/increase in turbulent heat fluxes (Fig. 3.8 a, compensation calculated from the difference of the respective flux between a CLM - CTL-like simulation and that of a CLM - BHS-like simulation). Fig. 3.8 b displays the fraction of the BHF that is compensated by decreased/increased emission of longwave radiation by the land surface (f_{rad}) against the gradient between the atmospheric temperature and the vegetation temperature (as an indicator for near-surface stability). f_{rad} remains low and surprisingly constant throughout the day, indicating an only weak reduction of the LST when BHS is positive. Once the sun sets, the surface layer grows more stable, thereby inhibiting the compensation by the turbulent heat fluxes. Therefore, f_{rad} gradually increases during early night up to roughly 0.6 and remains relatively constant during the second half of the night (accompanied by a strong nighttime increase in LST). As hypothesized before, sensible heat even provides a small positive

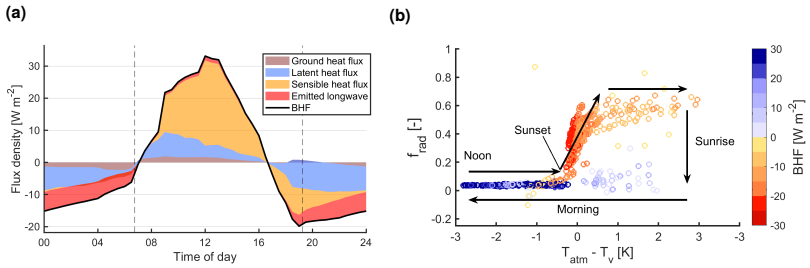


Figure 3.8: *Panel (a), the average diurnal cycle during boreal winter (DJF) at 2.75° S/60.25° W (location of Michiles and Gielow, 2008) of the biomass heat flux (BHF) in black. The shaded areas indicate the amount of the BHF compensated for by a change in the ground heat flux (brown), latent heat flux (blue), sensible heat flux (orange) and the emitted longwave radiation (red) compared to a simulation without biomass heat storage (BHS; compensation calculated from the difference of a simulation without BHS minus a simulation with BHS in the respective flux). The dashed lines indicate the sunrise and sunset. Panel (b) displays the fraction of the BHF that is compensated by emitted longwave radiation (red area divided by black curve in panel a) against the gradient between the atmospheric temperature and the vegetation temperature (as an indicator for near-surface stability). The colour of the markers displays the BHF. Half-hourly values of the monthly-averaged diurnal cycles were used for this plot.*

feedback during the late night, due to the lower surface stability resulting from BHS (The dark red area in Fig. 3.8 a from 2:00 to 7:00 am indicates that sensible heat has a warming effect compensating part of the emitted longwave radiation effect). A similar diurnal cycle of f_{rad} occurs in the mid-latitudes at the site of Lindroth et al. (2010) during the summer months (Fig. B.9).

The strong dependency of f_{rad} on the surface layer stability can also explain why the daytime cooling effect of BHS in the mid-latitudes is more pronounced during winter. During this season the daytime compensation of BHF by reduced turbulent heat fluxes is less efficient as stable conditions occur more frequently than during summer (Chan and Wood, 2013). Hence, a larger proportion of the BHF is compensated by reduced emission of longwave radiation. At the site of Lindroth et al. (2010), f_{rad} often exceeds 0.1 during the mid-latitudinal winter even when the BHF is positive (i.e., the biomass takes up energy), which is normally not the case during summer (cfs. Figs. B.9 and B.10). Hence, the characteristic diurnal cycle of f_{rad} which emerges in the tropics and during summer in the mid-latitudes, does not develop as clearly during DJF.

According to our modeling results, the diurnally asymmetric temperature difference between forest and nearby open land, which was found in observational studies (Lee et al., 2011; Vanden Broucke et al., 2015; Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b), can thus be explained by differing relevant processes during day and night. In various studies it was found that the differences in the evaporative fraction and albedo are more relevant for the temperature difference of forest minus open land during day (Liu et al., 2005; Vanden Broucke et al., 2015; Meier et al., 2018), whereas it appears that BHS mainly affects nighttime temperatures (Swenson et al., 2019 and this study). However, another part from this asymmetry is likely related to the higher turbulent heat fluxes over forests due to their high surface roughness (Schultz et al., 2017). The sensible heat flux can thus have a cooling effect over forests during day and a warming effect during night compared to open land, if this flux changes the vertical direction during the course of the day (which is normally the case; Liu et al., 2005; Vanden Broucke et al., 2015).

3.4.2 Limitations and knowledge gaps

While this study demonstrates the importance of BHS on a global scale, it could not address several aspects. The evaluation of the DTR in the CLM5.0 simulations with MODIS at grid-cell-level is to some degree an unfair comparison, as part of the biases could be the result of discrepancies between the atmospheric forcing used in our simulations and the actual conditions during the MODIS observations. As a consequence, even a perfect land surface model could not achieve a perfect agreement with the observations. Unfortunately, it is not straight-forward to assess the disagreement between MODIS and GSWP3 as LST is not directly comparable to atmospheric temperature at the lowest level. Yet, we find that the DTR in the MODIS LST data often considerably exceeds the DTR in the GSWP3 forcing, at locations where CLM5.0 underestimates the DTR in LST (cfs. Fig. 2.2 and Fig. B.11). Also, the MODIS observations can be retrieved only under cloud-free conditions. Unfortunately, the GSWP3 forcing does not contain any cloud cover information, hence making it difficult to mask out cloudy days in the model output. Further, the disagreement in DTR between CLM5.0 and MODIS could also be

related to bare soil or open land PFTs, where BHS is unimportant due to the small (or non-existent) biomass of these LC types. In fact, our analysis revealed that the DTR is often underestimated by CLM5.0 in sub-tropical and mid-latitudinal regions, which are mostly covered by bare soil or open land (cf. Figs. 2.2 b and B.1). Hence, the missing process of BHS is likely not the only source for the difference between our CLM5.0 simulations and the MODIS DTR.

At least two potential feedbacks to BHS cannot be addressed with our simulation set up. First, land-only simulations inherently mute the feedbacks from the atmosphere. We found that approximately 90 % of the energy-uptake by the biomass is compensated by a reduction of the turbulent heat fluxes during the day. BHS therefore reduces the daytime energy input from the land surface into the boundary layer. This reduced energy input could result in dynamical feedbacks in the boundary layer. It seems therefore necessary to analyse simulations coupled to the atmosphere to capture the full effect of BHS. Another feedback of BHS that we do not assess in this study are the changes in vegetation structure induced by BHS. The alteration of the local climate due to BHS affects primary productivity and respiration (not shown). This can not only alter the vegetation phenology (e.g., vegetation height, leaf area index, biomass) but also the carbon budget of the land surface. We therefore encourage further research in this direction.

Finally, there remain several aspects of energy storage at the land surface that could be further improved in the CLM modeling framework. First, the model we use restricts the energy storage at the land surface to the ground and biomass energy storage, neglecting the sensible and latent heat stored in the canopy air space, which can be of comparable magnitude in observations as BHS (e.g., Michiles and Gielow, 2008; Lindroth et al., 2010; Kilinc et al., 2012). Bonan et al. (2018) find that introducing a multi-layer canopy model, which represents the roughness sublayer and thus the canopy air space, did reduce but not alleviate completely the positive bias in the DTR in forests. However, this model version did not account for BHS in the stem. Hence, the combination of a multi-layer canopy model and heat storage in the stem could potentially alleviate the overestimation of the DTR in forests completely. Second, we assume a uniform stem temperature in our model, which is of course not the case in reality. The existing BHS parametrisation could thus be refined with the "force-restore" method, which represents two stem layers (Haverd et al., 2007), or even by including multiple stem layers, as proposed in the "analog model" by Herrington (1969). Further, a number of parameters such as the tree number density are assumed to be globally constant for the different PFTs. In reality however, such parameters can vary considerably within a PFT. Hence, it could be beneficial to include spatially explicit data if available from observations.

3.4.3 Implications for LCC impacts on climate

Recently, the local and regional climate impact of LC changes such as deforestation has been a heavily discussed topic in the literature. These climatic impacts are often assessed by contrasting climate model simulations that differ in LC. However, all of the CMIP5, LUCID, and LUCAS regional climate models miss or underestimate the observed nighttime cooling effect and daytime warming effect associated with deforestation (Lejeune et al., 2017; Davin et al., 2020). This indicates, that current climate models miss an important

part of the climate signal of forest-related land-use changes. Up to now, BHS has often been neglected in global climate models, although observed BHFs are of non-negligible magnitude across the diurnal cycle. Given our results of including BHS in CLM5.0, we reach the following three main conclusions regarding the role of BHS in the climate system: (1) BHS dampens the DTR in forested regions, thereby improving the model agreement with MODIS observations. (2) BHS affects LST stronger under stable surface layer conditions as compared to unstable conditions. As a consequence, the BHS-induced temperature increase during the night tends to exceed the temperature reduction during the day. (3) The effect of BHS is especially pronounced during warm extremes. These conclusions indicate that the representation of BHFs is important, especially in the case of diurnal temperature variations and/or temperature extremes in forests. Climate models used to assess the biogeophysical impact of forest-related land-use changes should thus consider including BHS in their land surface component.

Acknowledgements We would like to express our gratitude to the two reviewers, who helped to improve the quality of this publication considerably. We acknowledge funding from the Swiss National Science Foundation (SNSF) and the Swiss Federal Office for the Environment (FOEN) through the CLIMPULSE project and we thank the Swiss National Supercomputing Centre (CSCS) for providing computing resources.

Box 1: Preliminary results on atmospheric feedbacks from BHS

Introduction As mentioned in the article, the compensation of BHS by the turbulent heat fluxes implies that BHS might trigger relevant feedbacks in the atmosphere. Here, I briefly discuss two specific effects of BHS that appear in coupled regional climate simulations over a maximally forested Europe. This topic is currently pursued more systematically in a Master's thesis by Chenwei Xiao.

Model simulations The simulations presented here follow the FOREST protocol of the LUCAS phase one (Davin et al., 2020). They were performed with the atmospheric component of COSMO-CLM coupled to CLM5.0. Coupling previous versions of CLM to COSMO-CLM has proven to improve the model performance over Europe considerably (Davin et al., 2011; Davin and Seneviratne, 2012; Davin et al., 2016). The simulations cover a Europe, where all the land surface except for bare soil has been converted to the dominant forest type. They are conducted at 0.44° resolution and are analysed from 1986-2015 following a seven year spinup from 1979-1985. The effect of BHS is assessed by comparing two simulations: CC - DFLT, COSMO-CLM coupled to the default version of CLM5.0 (with an upper cap for the Monin-Obukhov stability parameter at 0.5), and CC - BHS, where BHS is active and the upper cap of the Monin-Obukhov stability parameter is raised to 100.

Summertime increase in convective precipitation It appears that BHS increases the amount of convective precipitation during the summer months in Europe (Fig. 3.9). Most likely linked to it is an increase in the latent heat flux and the convectively available potential energy (not shown). Note however that convection is parametrised in the presented simulations due to their relatively coarse resolution, which can impede the representation of clouds and precipitation (Hohenegger et al., 2009; Prein et al., 2015; Hentgen et al., 2019). Therefore, it is possible that this signal is an artefact of the parametrisation of convection (i.e., that precipitation events are allocated to convection more frequently without an overall change in precipitation).

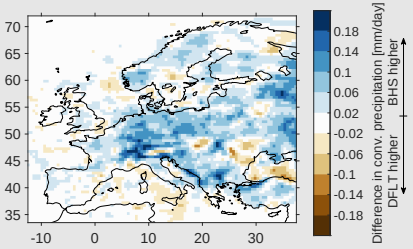


Figure 3.9: *Difference in JJA convective precipitation of CC - BHS minus CC - DFLT.*

Nighttime cooling in winter During the winter months, CC - BHS is colder at night than CC - DFLT in most regions of Europe, which is not necessarily expected from the nature of the process (Fig. 3.10 a). An energy balance decomposition after Luyssaert et al. (2014) reveals that the main contributor to this is the difference in the sensible heat flux between CC - BHS and CC - DFLT (Fig. 3.10 c). This term overcompensates the nighttime warming induced by BHS (Fig. 3.10 b) in many regions, leading to a colder temperatures in CC - BHS overall. This effect likely originates partly from the removal of the stability cap in CC - BHS, which has an amplified effect during night and during winter, when the surface layer is frequently stable.

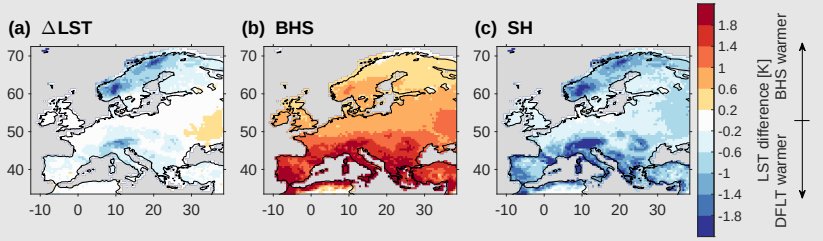


Figure 3.10: *LST difference at 01:30 solar time in DJF of CC - BHS minus CC - DFLT (a). Contribution to this difference from BHS (b) and the difference in the sensible heat flux (c) after the energy balance decomposition of Luyssaert et al. (2014). Note that only the two most relevant terms of the energy balance decomposition are shown here.*

4

Empirical estimate of forestation-induced precipitation changes in Europe

This manuscript is accepted in Nature Geoscience[*], Ronny Meier[1], Jonas Schwaab[1], Sonia I. Seneviratne[1], Michael Sprenger[1], Elizabeth Lewis[2], Edouard L. Davin[1]

Abstract Land cover changes can affect the climate by altering the water and energy balance of the land surface. Numerous modeling studies have indicated that alterations at the land surface can result in considerable changes in precipitation. Yet, land cover-induced precipitation changes remain largely unconstrained by observations. Here we use an observation-based continental-scale statistical model to show that forestation of rainfed agricultural land in Europe triggers substantial changes in precipitation. Locally, we find an increase in precipitation following forestation, in particular in winter, which is supported by a paired rain gauge analysis. In addition, forests are estimated to increase downwind precipitation in most regions during summer. In contrast, the downwind effect in winter is positive in coastal areas but near-neutral and negative in Continental and Northern Europe, respectively. The combined local and non-local effects of a realistic reforestation scenario, constrained by sustainability safeguards, are estimated to increase summertime precipitation by 7.6±6.7 % on average over Europe (0.13±0.11 mm/day), potentially offsetting a substantial part of the projected precipitation decrease from climate change. We therefore conclude that land cover-induced alterations of precipitation should be considered when developing land management strategies for climate change adaptation and mitigation.

[*]This manuscript was changed from its submitted version to ensure consistency throughout this thesis.

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

[2]School of Engineering, Newcastle University, Newcastle, United Kingdom

4.1 Introduction

Observations show that Land Cover Changes (LCCs) alter local temperatures considerably through biogeophysical processes (Lee et al., 2011; Li et al., 2015; Duveiller et al., 2018b; IPCC 2019c). Such observational constraints have enabled the evaluation of the Land Cover (LC)-temperature coupling in climate models (Lejeune et al., 2017; Winckler et al., 2017; Duveiller et al., 2018a; Meier et al., 2018; Meier et al., 2019). Alongside temperature, water availability is amongst the most important climate drivers for life on Earth. The high precipitation amounts over tropical rainforests are partly sustained by the presence of the forests themselves (Spracklen et al., 2012; Lejeune et al., 2015). On the other hand, precipitation can increase locally after clearing a forest patch in the Amazon (Khanna et al., 2017). Outside of the tropics, modeling studies suggest that LCCs can trigger substantial changes in precipitation (Yosef et al., 2018; Belušić et al., 2019). However, such studies lack solid observational constraints at regional to global scales, as little work has been conducted on estimating the LC-precipitation coupling in the mid-latitudes based on observations (Perugini et al., 2017). This is likely owed to the difficulty of isolating such a signal. The distribution of a certain LC type itself is often influenced by the precipitation climatology, by other factors (e.g., elevation) that in turn affect precipitation, or systematic preferences for certain terrain types from human land use (Sandel and Svenning, 2013). A further complication is that LC-precipitation couplings do not necessarily occur at the location of the LC, but might also have an influence on precipitation further downwind. Thus, they are difficult to observe especially in regions of heterogeneous LC, as in Europe.

In this study, we estimate the impact of converting rainfed agricultural land (AL_r; including pasture) to forest on precipitation over Europe. We focus on this LC conversion since it has been the dominant LCC over the last decades in many European regions (Fig. C.8; Fuchs et al., 2013; Hansen et al., 2013; Fuchs et al., 2014; McGrath et al., 2015). An estimated 20 % of the European land area has undergone de-, re-, or afforestation between 1900 and 1990 mainly in exchange for cropland and grassland (Fuchs et al., 2014). In addition, reforestation has been proposed as a tool to mitigate greenhouse gas emissions (Griscom et al., 2017a), while potentially providing additional co-benefits for biodiversity (Navarro and Pereira, 2015), soil protection (Navarro and Pereira, 2015), and the local climate (Lee et al., 2011; Li et al., 2015; Duveiller et al., 2018b). For simplicity, we assume here that forestation corresponds only to the conversion of AL_r to forest. We acknowledge however that other forms of human land use, such as settlements, might also be suitable for forestation. Also, we use forestation as a general term, not distinguishing between afforestation and reforestation. Due to the lack of constraints from previous studies, we employ two independent methodologies to corroborate our results (Fig. 4.1): (1) We search for suitable closely-located rain gauge site pairs in the Global Sub-Daily Rainfall Dataset (GSDR; Lewis et al., 2019) and Global Historical Climatology Network (GHCN; Menne et al., 2012b; Menne et al., 2012a), which differ in the AL_r and forest fractions. Such an approach has previously been applied to assess the effect of LC on temperature (Lee et al., 2011; Zhang et al., 2014) and the surface energy budget (Liu et al., 2005; Juang et al., 2007; Vanden Broucke et al., 2015). (2) We statistically disaggregate the monthly precipitation climatology over 1986 to 2015 from MSWEP v2.2 (Multi-Source Weighted-

Ensemble Precipitation version 2.2; Beck et al., 2019b) with a Generalized Additive Model (GAM) to infer the contribution of LC on the spatial distribution of precipitation. Statistical models were successfully used as a tool to assess both local (Schwaab et al., 2020) and remote (Cohn et al., 2019) temperature changes induced by LCC. For the GAM presented here, we consider topographic effects, which strongly modulate the spatial distribution of precipitation (Houze Jr., 2012), by including a number of topographic metrics as predictors. The ERA5-Land 2 m temperature climatology (C3S, 2019) serves as an indicator for climatic conditions, as the amount of precipitation varies considerably across climate zones. Further, the atmospheric circulation is an important driver of the spatial distribution of precipitation (Daly et al., 2008) and is therefore represented in the GAM by various metrics based on ERA5 air parcel trajectories (Sprenger and Wernli, 2015). Finally, we employ the CORINE LC data (Kosztra et al., 2019) as predictors in the GAM to examine the relation between LC and precipitation. Alongside the local LC, we consider the upwind LC fractions, calculated from the backwards ERA5 air parcel trajectories. Therefore, we can not only isolate the local effect of forestation but also its downwind effect.

4.2 Methods

Fig. 4.1 provides an overview of the overall procedure and how the various datasets (Table 4.1) and analysis steps are connected.

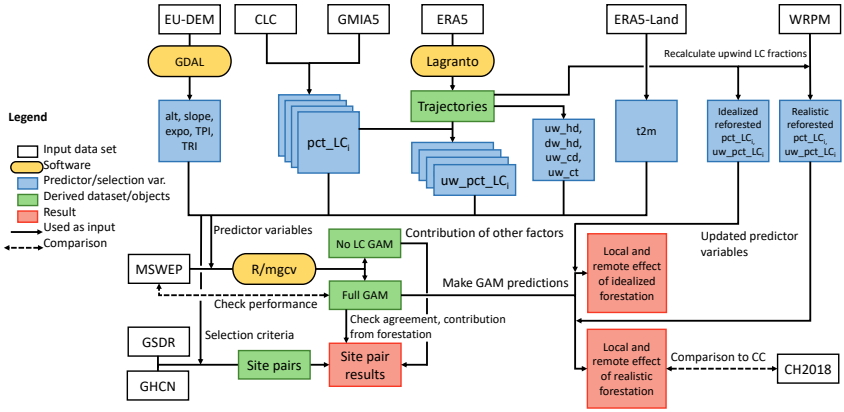


Figure 4.1: ***Overview of the steps, softwares, and input data sets described in the Methods and how they are interconnected.*** *White boxes, input data sets and yellow boxes softwares used in the study (Table 4.1). The blue boxes show variables that were used as predictor variables in the GAMs and for the selection criteria in the site pair analysis, the green boxes derived data sets and objects, and red boxes the final results.*

4.2.1 Gridded data

We employ a number of gridded data sets, which are used for two purposes: (1) They serve as criteria to find suitable site pairs in the GSDR and GHCN data. (2) They are used as response variables and predictors for the GAM. All the data are used at a regular 0.1° grid, which is the native resolution of the MSWEP data set. The study domain of the GAM is confined to the coverage of the CORINE LC data. Turkey was removed from the analysis, since it is surrounded by areas not covered by CORINE, which hampers the calculation of the upwind LC fractions.

Table 4.1: *Overview of the different data sets and softwares used in this study, the variables derived from them, their reference, and their availability.*

Date set	Variables	Reference	Data availability
GSDR	In-situ precipitation	Lewis et al. (2019)	Upon request from Elizabeth Lewis
GHCN-Daily v3.20	In-situ precipitation	Menne et al. (2012b) Menne et al. (2012a)	ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/
MSWEP v2.2	Gridded precipitation	Beck et al. (2019b)	https://platform.princetonclimate.com/PCA_Platform/
EU-DEM v1.1	alt, slope, expo, TPI, TRI	EEA (2014)	https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1
CLC	pct_LC_i, $prox\_LC_i$, $uw\_pct\_LC_i$	Kosztra et al. (2019)	https://land.copernicus.eu/pan-european/corine-land-cover/clc-2000
GMIA5	pct_{agr_irr}, $prox_{agr\_irr}$, $uw\_pct_{agr\_irr}$	Siebert et al. (2013)	http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas
ERA5	uw_hd, dw_hd, uw_cd, $uw\_pct\_LC_i$, uw_ct	C3S (2017)	https://cds.climate.copernicus.eu
ERA5-Land	T2M	C3S (2019)	https://cds.climate.copernicus.eu
WRPM	Reforestation potential	Griscom et al. (2017a), Griscom et al. (2017b)	https://zenodo.org/record/883444
CH2018	Climate change signal	CH2018 (2018)	https://doi.org/10.18751/climate/scenarios/ch2018/1.0
Software	**Variables**	**Reference**	**Data availability**
GDAL	alt, slope, expo, TPI, TRI	-	https://cran.r-project.org/web/packages/rgdal/index.html
LAGRANTO	uw_hd, dw_hd, uw_cd, $uw\_pct\_LC_i$, uw_ct	Wernli and Davies (1997), Sprenger and Wernli (2015)	www.lagranto.ethz.ch
mgcv	Estimated precipitation changes	Wood (2011), Wood (2017)	https://cran.r-project.org/web/packages/mgcv/index.html

Precipitation fields

We use the monthly precipitation climatologies from MSWEP v2.2 over the period 1986-2015 as the response variable to train the GAMs. This data set is the latest version of the Multi-Source Weighted-Ensemble Precipitation data set (Beck et al., 2019b). For simplicity we refer to it as MSWEP. It is available globally at 0.1° spatial and 3-hourly temporal resolution. As indicated by its name, multiple rain gauge data sets, remote sensing products, and reanalysis products were used to create MSWEP (Beck et al., 2017a). This data set systematically outperforms other gridded precipitation data sets when evaluated with rain gauge data and streamflow measurements (Beck et al., 2017b) as well as radar measurements (Beck et al., 2019a). We calculate monthly precipitation climatologies from MSWEP over the period 1986-2015 (Fig. 4.2).

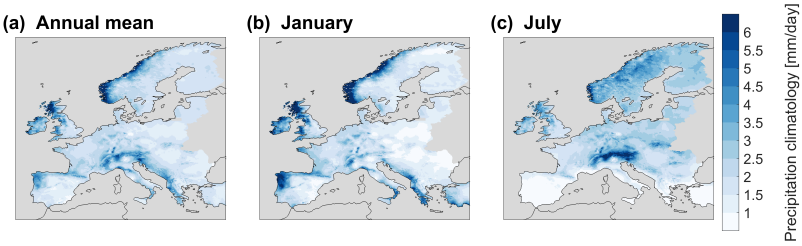


Figure 4.2: *Climatology in daily precipitation of MSWEP over 1986-2015 averaged over the entire year (a), January (b), and July (c).*

Topography

Topography plays a crucial role in the spatial distribution of precipitation (Daly et al., 2008; Houze Jr., 2012; Lu, 2019). We use the Digital Elevation Model over Europe version 1.1 (EU-DEM v1.1; EEA, 2014) to derive a number of relevant metrics. First, we compute the mean elevation at 0.1° resolution from the original 30 m resolution. Besides, we calculate the slope, exposition of the slope, the Topographic Position Index (TPI), and the Terrain Ruggedness Index (TRI) using the gdaldem tool of the Geospatial Data Abstraction software Library 2019 (GDAL, Fig. 4.3). The TPI describes the elevation of a point in comparison to the surrounding areas. It is defined as the elevation difference between the focal grid cell and the mean of the eight surrounding grid cells. The TRI is a measure of terrain variability and calculated as the mean of the absolute elevation differences between the focal grid cell and the eight adjacent grid cells. The exposition of the slope and the TPI are calculated from the 0.1° elevation data, whereas the slope and TRI are computed at the original resolution of EU-DEM and than averaged over the 0.1° grid.

Table 4.2: *Aggregation of the CLC classes (https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html) to the LC types used in this study. Left column the names of the classes used in this study, middle column CLC classes aggregated to respective class, and right column fraction of study domain covered by respective class.*

Name	CLC classes	Fraction [%]
Artificial surfaces	1. Artificial Surfaces (classes 1-11)	3.7
Rainfed agricultural land	2. Agricultural areas (classes 12-22)[a]	39.2
Irrigated agricultural land	2. Agricultural areas (classes 12-22)[b]	1.8
Forest	3.1 Forest (classes 23-25)	32.7
Natural low vegetation	3.2.1 Natural grassland and 3.2.2 Moors and heathland (classes 26-27)	5.3
Shrubland	3.2.3 Sclerophyllous vegetation and 3.2.4 Transitional woodland/shrub (classes 28-29)	6.9
Natural bare land	3.3 Open spaces with little or no vegetation (classes 30-34)	4.0
Wetland	4. Wetlands (35-39)	2.7
Open water	5. Water bodies (classes 40-44)	3.9

[a] "2. Agricultural areas" in CLC minus irrigated agricultural land.
[b] Minimum of irrigated agricultural land fraction according to Global Map of Irrigated Areas version 5 and "2. Agricultural areas" in CLC.

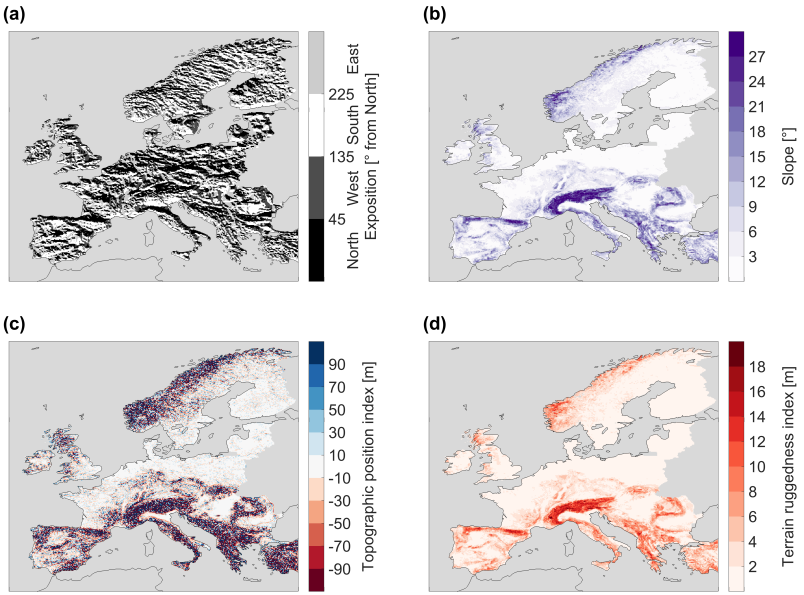


Figure 4.3: *Exposition (a), slope (b), topographic position index (c), and terrain ruggedness index (d) based on EU-DEM.*

Land cover information

For the LC information, we regrid the CORINE Land Cover (CLC) data from the year 2000 at 100 m resolution to 0.1° resolution (Kosztra et al., 2019). The original 44 LC classes are aggregated to the classes artificial surfaces, agricultural land, forest, shrubland, natural low vegetation, bare soil, wetland, and open water according to Table 4.2. Further, we differentiate irrigated agricultural land using the Global Map of Irrigation Areas version 5 (GMIA5; Siebert et al., 2013), because irrigation itself can trigger substantial modifications of the precipitation distribution (DeAngelis et al., 2010; Thiery et al., 2017). The fraction of irrigated agricultural land in a grid cell from GMIA5 is used, unless it exceeds the fraction of agricultural land in CLC. In this case, it is set to the fraction of agricultural land in CLC, to assure the fraction of irrigated agricultural land does not exceed the total fraction of agricultural land in CLC (Fig. 4.4). The remainder of the agricultural land in CLC is then attributed to AL_r.

Trajectory-based fields

The spatial distribution of precipitation is not only affected by the properties of the land surface at a given location, but also the atmospheric motion prior to reaching this location. In particular the distance to the coast along the path of atmospheric motion is an important driver of how much moisture is available for precipitation (Daly et al., 2008), as the oceans are the most important moisture source. To account for the relationship between the atmospheric circulation and precipitation, we use the LAGRangian ANalysis TOol (LAGRANTO) based on hourly ERA5 reanalysis data (Wernli and Davies, 1997; Sprenger and Wernli, 2015). Every six hours between January 1986 and December 2015, we calculate two kinematic trajectories five days backward in time for all land grid cells where the precipitation exceeded 1 mm/hr. We choose the two pressure levels that have produced most precipitation in the ERA5 data as starting heights, where the precipitation production at a specific level was calculated from the volumetric rain and snow water content of this level minus the volumetric rain and snow water content of the level above (Fig. C.16). The calculations are done in time steps of 10 min using hourly ERA5 wind field data at 0.5° resolution. LAGRANTO enables to track the position of the air parcel, the specific humidity, the land-sea mask, the elevation of the surface, and precipitation at the surface of ERA5, as well as the LC fractions from CLC. In the post-processing, we follow the trajectories backwards in time until the coast to determine the distance of the final precipitation event to the coast along the trajectory (uw_cd), the time it took to reach the coast (uw_ct), and the elevation difference between the highest topographic point the airmass surmounted on its path from the coast and the end point (uw_hd; Fig. 4.5 d and e). In addition, we compute the average fraction of each LC class the airmass overpassed during the previous five days, which are called upwind LC fractions here (uw_pct_LC$_i$; Fig. 4.5 a-c). In the calculation of the upwind LC fractions we stop tracking the trajectory if the specific humidity of the parcel falls below 0.05 g/kg or if a land point is reached according to the ERA5 land-sea mask that is outside of the CLC domain. Collecting all

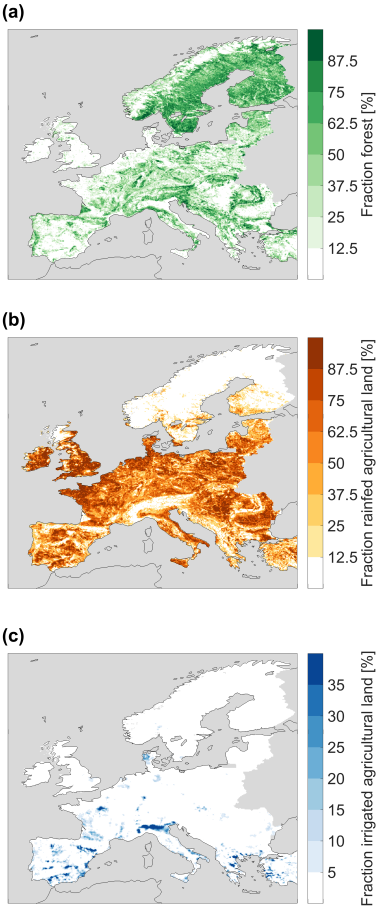


Figure 4.4: *Fraction of land covered by forest (a), AL$_r$ (b), and irrigated agricultural land (c).*

trajectories for a given location and month, we then compute the median value for each of those metrics to retrieve the predictor fields for the GAMs. Using a similar approach, we calculate the median of the elevation difference between highest topographic point the airmass surmounted during the day following the precipitation event and the point where the precipitation occurred (dw_hd; Fig. 4.5 f).

2 m Temperature

Finally, we use the ERA5-Land monthly 2 m temperature climatology from 1986-2015 as a predictor in the GAM, which is available at the required 0.1° resolution (C3S, 2019).

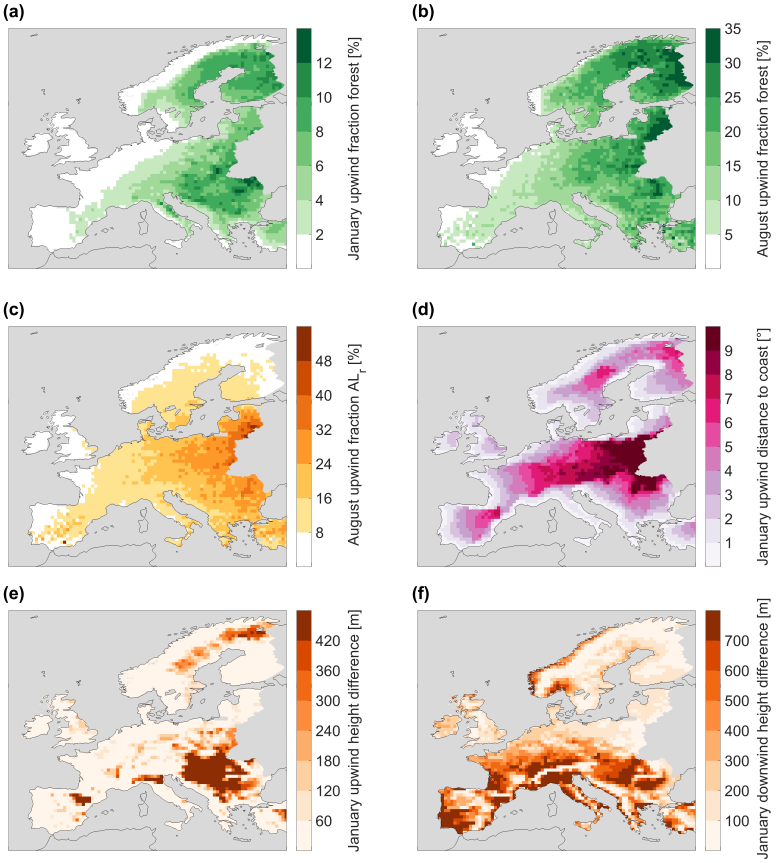


Figure 4.5: ***Examples of trajectory-based fields.*** *Upwind forest fraction in January (a) and August (b). Panel (c) shows the upwind fraction of AL_r in August. Panel (d) shows the upwind distance to coast, panel (e) the upwind height difference, and panel (f) the downwind height difference in January.*

4.2.2 Site pair analysis

For the paired rain gauge station analysis we employ the GSDR data set (Lewis et al., 2019) of the following countries: Belgium, Spain (Catalonia), Finland, Germany, Ireland, Italy, Norway, Portugal, Switzerland, and the UK. In addition, we include the data from the Integrated Surface Database (Smith et al., 2011), which is not country specific. GSDR has been extensively quality controlled for common mechanical, recording, and storage errors in rainfall data as well as being checked against neighbouring gauges (Blenkinsop et al., 2017). The station data are aggregated from the original hourly to a three-hourly resolution to match with the temporal resolution of MSWEP. We exclude stations with more than 50 % missing data arriving at a total of 3481 stations in the study domain. Besides GSDR, we also employ the GHCN-Daily v3.20 (Menne et al., 2012a; Menne et al., 2012b) rain gauge collection, which is available at daily resolution and has 2376 stations with less than 50 % missing days available in the study domain (Fig. C.10). Given the differing temporal resolution and quality checks of the two data sets, we only look for site pairs within one data set. Further, we neglect any measured data prior to 1986. For all of these stations we extract the values of the gridded data sets at the respective location to check for the following criteria to find station pairs: (1) The first site has at least 20 % more forest coverage and at least 20 % less coverage by AL_r than the second site according to the CLC data. Note that the site comparison does not correspond to a pure local comparison, but rather to a comparison of having 20 % more/less of AL_r/forest within a grid cell of roughly 11 km by 7 km. The isolated signal does therefore not only include the effect of foresting one specific location but also the effect of foresting areas in close proximity. (2) The sum of the AL_r and forest fractions is at least 50 % at each site. (3) The great-circle distance between the two sites is less than 0.75° corresponding to 83.5 km. (4) They have at least three overlapping years of measurements. (5) They differ less than 25 m in altitude, 5° in slope, 3 m in TRI, 20 % in the open water fraction, 20 % in the upwind open water fraction, and 0.25° in the annual mean uw_cd. For each station in a pair, we then calculate the monthly rainfall climatology considering only times during which both stations have collected data. As an additional quality check, we discard station pairs that had a smaller sample size than 400 and 50 for any month in GSDR and GHCN, respectively (corresponding to 50 measurement days). An analysis of the sensitivity of the site pair analysis on the selection criteria is presented in Supplement C.1.

4.2.3 GAM construction

GAMs express the expected value, $\mathbb{E}$, of a response variable, Y_i, as the sum of a number of smooth functions (Wood, 2017):

$$g(\mathbb{E}(Y_i)) = f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + ... \tag{4.1}$$

, where g is a link function, f_j are the smooth functions, and x_k are different predictor variables. GAMs possess a number of features that are advantageous for our analysis: (1) The smooth functions allow to empirically capture non-linear relations

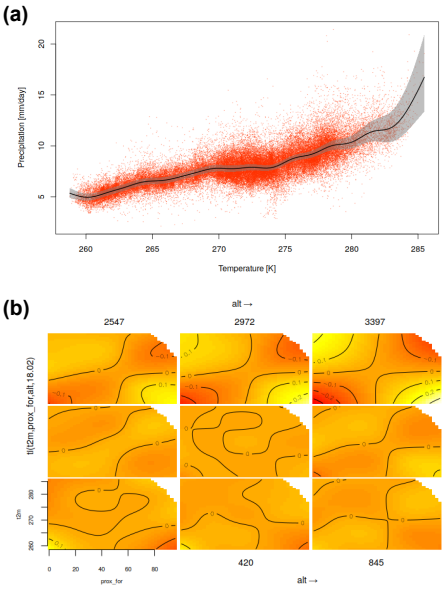


Figure 4.6: ***Example of smooth functions.*** *Panel (a), thin plate regression spline ('s' smooth) of 2 m temperature in January (black) with standard error (grey) and partial residuals (orange). Panel (b), tensor product smooth ('ti' smooth) for the proximity forest, 2 m temperature, and altitude term in January. Each panel shows a two dimensional smooth for proximity forest (X-axis) and 2 m temperature (Y-axis) at a specific altitude. Contours are the precipitation smooth. Note that a smooth of variables that are present in several other smooths cannot be interpreted in isolation.*

between the predictor and the response variables without a prior assumption regarding the shape of this relation. (2) The smooth functions can depend on several predictor variables. (3) The additive structure of GAMs makes them readily interpretable. (4) The smooth functions are penalized for their 'wiggliness', which prevents overfitting of the model. In this case, the response variable is the monthly precipitation climatology of MSWEP, which is assumed to follow a gamma distribution, g is the natural logarithm, and the f_j terms are summarized in Table 4.3. The logarithmic link function was chosen to avoid harming the assumption of constant variance of residuals. Precipitation changes from forestation are estimated in a spatially-explicit way with our GAM, by including interaction terms between the LC fractions and other spatially-varying variables ('prox_LC$_i$, T2M, alt' and 'uw_pct_LC$_i$, T2M, uw_ct' smooths) and by directly including interactions terms of the LC fractions with the spatial dimensions ('prox_LC$_i$, lat, lon' and 'uw_pct_LC$_i$, lat, lon' smooths).

We employ the extensive R package Mixed GAM Computation Vehicle with automatic smoothness estimation ('mgcv'; Wood, 2011; Wood, 2017) to construct a GAM for each month of the MSWEP precipitation climatology. To shorten computation times the model is fitted with the 'bam' function, which is numerically optimized for large datasets, using the fast REstricted Maximum Likelihood method (fREML) to estimate the smoothing parameters (if not specified) and discretisation of covariate values, which increases the computational efficiency of the fREML method (Wood et al., 2017; Li and Wood, 2020). We use thin plate regression splines ('s' smooths) for the one-dimensional terms, which are invariant in space and therefore called global terms here, and tensor product smooths ('ti' smooths) for the multi-dimensional terms, which are suitable if other terms using the respective predictors are used (Fig. 4.6). To avoid unrealistically noisy results, the smoothing parameters were chosen manually for the multi-dimensional terms. For the

Table 4.3: *Smooths used to construct the GAM. First column the predictor variables used in the respective smooth. Second column names the type of smooth used. Third column shows the maximum number of nodes of the smooth (higher k more flexibility) and column sp the smoothing parameter applied (higher sp results in a smoother function). In blue smooths that were removed when fitting the 'no LC' GAM. Last column shows the maximum p-value across the GAMs for the individual months of the respective smooth and in brackets number of months for which this smooth had a p-value larger than 0.01. For the LC-related smooths values are given in order of forest, AL_r, shrubland, natural low vegetation, artificial surfaces, irrigated agricultural land, natural bare land, open water, and wetland.*

Variable(s)	Type	k	sp	Description	Max p-value
alt	s	15	estimated	global altitude term	3.1e-29 (0)
alt, lat, lon	ti	5	5	spatially varying altitude term	7.1e-19 (0)
TPI	s	10	estimated	global TPI term	7.4e-42 (0)
expo	s	10	estimated	global exposition term	1.3e-7 (0)
expo, lat, lon	ti	5	5	spatially varying exposition term	8.6e-22 (0)
slope	s	15	estimated	global slope term	1.9e-2 (1)
slope, lat, lon	ti	5	5	spatially varying slope term	7.2e-34 (0)
TRI	s	10	estimated	global TRI term	8.6e-2 (2)
T2M	s	20	estimated	global 2 m temperature term	1.0e-71 (0)
uw_hd	s	10	estimated	global upwind height difference term	1.2e-41 (0)
dw_hd	s	10	estimated	global downwind height difference term	8.8e-76 (0)
uw_cd	s	10	estimated	global upwind coast distance term	4.2e-45 (0)
prox_LC$_i$	s	10	estimated	global term for local effect of each LC class	1.6e-4 (0), 5.1e-5 (0), 1.45e-13 (0), 8.2e-3 (0), 1.5e-1 (3), 1.2e-14 (0), 1.4e-4 (0), 2.2e-2 (1), 2.1e-8 (0)
prox_LC$_i$, lat, lon	ti	3	5	spatially varying local effect of each LC class	2.6e-2 (1), 2.3e-6 (0), 1.3e-2 (1), 1.7e-1 (1), 1.1e-2 (1), 5.4e-1 (3), 2.7e-2 (1), 2.1e-4 (0), 9.8e-1(4)
prox_LC$_i$, T2M, alt	ti	5	5	term varying with mean climate for local effect of each LC class	7.3e-27 (0), 1.3e-20 (0), 2.1e-9 (0), 5.0e-6 (0), 1.9e-2 (2), 1.7e-15 (0), 5.3e-12 (0), 3.2e-2 (1), 5.4e-8 (0)
uw_pct_LC$_i$	s	10	estimated	global upwind term for each LC class	1.9e-15 (0), 5.5e-18 (0), 4.4e-11 (0), 1.5e-8 (0), 1.1e-24 (0), 2.9e-58 (0), 5.4e-19 (0), 2.0e-21 (0), 2.0e-2 (1)
uw_pct_LC$_i$, lat, lon	ti	3	5	spatially varying upwind term for each LC class	6.0e-1 (2), 3.2e-2 (3), 1.8e-5 (0), 1.5e-9 (0), 2.2e-2(1), 3.6e-5 (0), 0.34e-2 (1), 2.2e-2 (2), 2.4e-1 (2)
uw_pct_LC$_i$, T2M, uw_ct	ti	5	5	upwind term varying with mean climate for each LC class	4.9e-12 (0), 6.7e-7 (0), 1.9e-44 (0), 4.5e-33 (0), 5.0e-3 (0), 2.0e-60 (0), 1.1e-3 (0), 5.3e-3 (0), 2.0e-11 (0)
lat, lon	ti	15	5	spatial term to capure large scale variation of precipitation	0 (0)

local effect of LC, we also account for proximity effects by computing proximity LC fractions giving a third weight to the central grid cell and a sixth weight to each of the four grid cells that share an edge with the central grid cell (prox_LC$_i$).

4.2.4 Performance assessment of GAM

Across all months, the GAM achieves an adjusted R^2 of 0.93, a RMSE of 0.30 mm/day, and an index of agreement (Duveiller et al., 2016) of 0.96. Overall, the residuals still show some spatial correlation up to a distance of approximately 1.5°, which we accept in order to not have to over-fit the precipitation field with the purely spatial term (Figs. C.15 and C.17). The GAM slightly deteriorates the agreement with the precipitation climatologies at the GSDR rain gauges compared to the original MSWEP data (Fig. C.14). However, considerable differences are already present between GSDR and MSWEP, likely owed to the different spatial and temporal scales of these data sets. These differences are not

systematic however. The significance of the individual smooths was tested in a type III ANOVA, using the 'mgcv' function 'anova'. Overall, all the smooths remain highly significant throughout the year, with a few exceptions mostly for interactions terms of some variables with latitude and longitude (Table 4.3). Note however, that the significance of the terms is likely to be slightly overestimated due to the spatial auto-correlation of the residuals (Dormann et al., 2007).

To compare the results of the GAM with the site pairs, we make a GAM prediction extracting the forcing data from each individual site in a site pair. Due to the differing spatial scales of the rain gauge data (point scale) and the MSWEP data (0.1° grid), we use the local CLC fractions at the sites (within 0.1° grid cell) as the proximity when conducting the GAM prediction. Besides this, we make a second GAM prediction, using only the local CLC AL_r and forest LC information from the individual sites, but setting all other predictor variables to the values at the location of the site with more forest coverage (Site 1). The AL_r and forest LC fractions at the second site are computed as follows to ensure the LC fractions still add up to 100 %:

$$\begin{aligned} \overline{\Delta Fr} &= mean(pct_{for1} - pct_{for2}, pct_{agr2} - pct_{agr1}), \\ \Delta Fr &= min(pct_{for1}, 100 - pct_{agr1}, \overline{\Delta Fr}), \\ pct^*_{agr1} &= pct_{agr1} + \Delta Fr,\ pct^*_{for1} = pct_{for1} - \Delta Fr \end{aligned} \tag{4.2}$$

, where pct_{for1} and pct_{agr1} are the fraction forest and fraction AL_r at the site with more forest, pct_{for2} and pct_{agr2} the respective fractions at the site with less forest, ΔFr the imposed LCC, and pct^*_{agr1} and pct^*_{for1} the new fractions of AL_r and forest to compute the contribution from the differences in those two LCs. A GAM prediction at Site 1 with these updated LC fractions is then subtracted from the original prediction at Site 1 to isolate the actual contribution from the differences in the AL_r and forest fractions, while excluding contributions from other factors, such as elevation. To further corroborate our results, we fit a GAM without the local LC fractions as predictors (blue terms in Table 4.3). Note that the linear fits in Figs. 4.8, C.13, and C.14 are produced with the Matlab function 'fitlm' using the option 'RobustOpts' to be less affected by outliers when fitting. A discussion on the sensitivity of the GAM results on its structure is provided in Supplement C.2.

4.2.5 Estimating the theoretical effect of forestation on precipitation

When estimating the forestation-induced precipitation change, we confine ourselves to foresting 20 % of the land surface, to (1) ensure the imposed LCC is of realistic magnitude and (2) prevent moving outside of the boundaries of the predictor space for the AL_r and forest LC fractions. Further, the GAM estimate of the local precipitation change from foresting AL_r on 20 % of the land surface ($\Delta P^1_{20\,\%}$) is evaluated with the site pair analysis. We determine $\Delta P^1_{20\,\%}$ using one LC map with 10 % more forest and 10 % less AL_r and a second LC map with 10 % less forest and 10 % more AL_r relative to CLC. When creating these maps, we prevent LC fractions from falling below 0 % or above 100 %. In case it is impossible to convert 10 % of the land surface in one direction, we allow for more than

10 % conversion in the other direction up to a total conversion of 20 %:

$$\begin{aligned} \Delta LC_{tot} &= min(20\,\%, prox_{agr} + prox_{for}), \\ \Delta LC_1 &= min(prox_{agr}, max(10\,\%, \Delta LC_{tot} - prox_{for})), \\ and\ \Delta LC_2 &= min(prox_{for}, max(10\,\%, \Delta LC_{tot} - prox_{agr})) \end{aligned} \tag{4.3}$$

, where $prox_{agr}$ and $prox_{for}$ are the proximities AL_r and forest in the CLC, ΔLC_{tot} is the total LCC between the map with more forest cover and the map with more AL_r, and ΔLC_1 and ΔLC_2 are the imposed changes in the AL_r and forest fraction for the forested map and the deforested map, respectively. ΔLC_1 is than added to the forest fraction and subtracted from the fraction of AL_r to create the forested LC map (LC_1), while the opposite is done with ΔLC_2 to create the deforested map (LC_2). To avoid the estimation of the forestation effect in regions with no AL_r available for forestation in Scandinavia, we prevent forestation at grid cells where the average fraction of AL_r within a moving window of 3.1° is below 8 %. With this procedure, it is possible to achieve 20 % forestation almost everywhere in Europe (Fig. 4.7 a). Then, we predict precipitation with each of the two new LC maps ($P_{pred}(LC_1)$ and $P_{pred}(LC_2)$) and calculate the difference of the two predictions to estimate the local change in precipitation due to foresting 20 % of the land surface:

$$\Delta P^{l}_{20\,\%} = P_{pred}(LC_1) - P_{pred}(LC_2) \tag{4.4}$$

Using the same procedure, we calculate the upwind LC fractions if the land was transformed according to the mentioned rules when processing the wind trajectories (Fig. 4.7 b and c; $uw\_pct\_LC_i(LC_1)$ and $uw\_pct\_LC_i(LC_2)$). Again, the two GAM predictions for each of the resulting upwind LC fractions are subtracted to determine the remote change in precipitation due to foresting AL_r on 20 % of the land surface, $\Delta P^{rem}_{20\,\%}$:

$$\Delta P^{rem}_{20\,\%} = P_{pred}(uw\_pct\_LC_i(LC_1)) - P_{pred}(uw\_pct\_LC_i(LC_2)) \tag{4.5}$$

4.2.6 Realistic forestation scenario

Finally, we estimate the precipitation change of reforesting Europe according to the World Reforestation Potential Map (WRPM; Griscom et al., 2017b), which estimates the reforestation potential relative to the LC state from 2000-2009. Starting from the original binary map indicating whether a location is suitable for reforestation at 0.0083° resolution, we calculate the area fraction which is suitable for reforestation at 0.1° resolution. This fraction we then add to the forest and subtract from the AL_r fractions of CLC, again making sure the AL_r cover fraction does not fall below 0 % and the forest fraction does not exceed 100 %. The estimate of the precipitation change due to this reforestation scenario then amounts to the difference of a GAM prediction with the forested local and upwind LC fractions minus a GAM prediction with local and upwind LC fractions according to CLC.

4.2.7 Comparison to climate change signal

The EURO-CORDEX EUR-11 regional climate simulations, used for comparison to the climate change signal, stem from the CH2018 collection (CH2018, 2018). EURO-CORDEX simulations at 0.44° resolution are excluded from the analysis, since they generally represent present day precipitation fields worse than the ensemble at 0.11° resolution (Prein et al., 2016). Compared to the original EURO-CORDEX ensemble (Jacob et al., 2014), some simulations were removed in CH2018 following quality checks. We use data from nine different RCM/GCM combinations with a total of ten available ensemble members at 0.11° resolution for the RCP4.5 and RCP8.5 scenarios to calculate multi-model mean precipitation climatologies for the periods 1986-2015 and 2071-2100 (Table 4.4, note that MPI-CSC-REMO2009/MPI-M-MPI-ESM-LR provides two ensemble members). The original data on a rotated grid were interpolated to a regular 0.1° grid using distance-weighted remapping of CDO. Then, we compute the area-weighted change in precipitation between 2071-2100 and 1986-2015 over five aggregated biogeographical regions compared to the climatology of 1986-2015. For the realistic forestation scenario, we compute the precipitation changes from this scenario compared to the MSWEP precipitation climatology of 1986-2015.

Table 4.4: *Ensemble of EUR-11 regional climate model (RCM) simulations forced by RCP4.5 and RCP8.5 in CH2018. Middle column name of driving general circulation model (GCM) and right column number of ensemble members.*

RCM Name	Driving GCM	N
CLMcom-CCLM4-8-17	ICHEC-EC-EARTH	1
CLMcom-CCLM4-8-17	MOHC-HadGEM2-ES	1
CLMcom-CCLM4-8-17	MPI-M-MPI-ESM-LR	1
DMI-HIRHAM5	ICHEC-EC-EARTH	1
MPI-CSC-REMO2009	MPI-M-MPI-ESM-LR	2
SMHI-RCA4	ICHEC-EC-EARTH	1
SMHI-RCA4	IPSL-IPSL-CM5A-MR	1
SMHI-RCA4	MOHC-HadGEM2-ES	1
SMHI-RCA4	MPI-M-MPI-ESM-LR	1

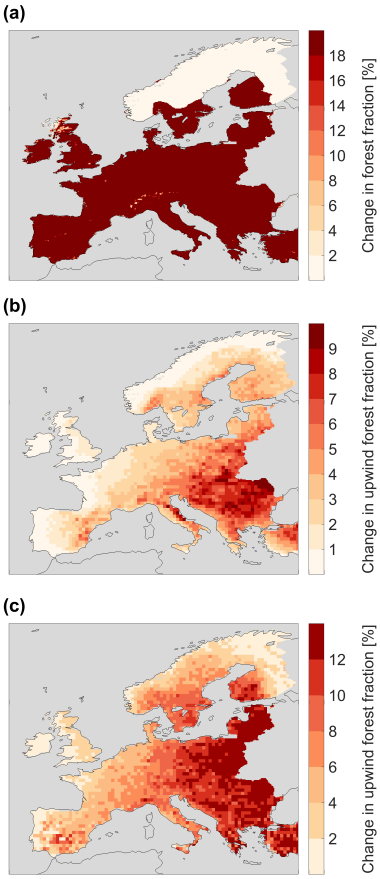


Figure 4.7: ***Changes in LC fractions for uniform forestation scenario.*** *Panel (a), change in local AL_r and forest fractions used to predict the local effect of forestation with the GAM (calculated from Eq. 4.3). To the right, the change in the upwind forest fraction following 20 % forestation in January (b) and July (c), retrieved from recalculating the upwind LC fractions in the trajectories with the altered LC maps.*

4.3 Results

4.3.1 Local effect of land cover on precipitation in site pairs

Among the 3481 rain gauge stations of GSDR and the 2376 stations of GHCN included in our analysis (Fig. C.10), we find a total of 1512 station pairs that differ at least by 20 % in the AL_r and forest LC fractions according to the CORINE LC data at 0.1° resolution and fulfill the additional selection criteria defined to mitigate the influence of potential confounding factors. Here, we focus on five clusters of station pairs, as approximately 100 sites pairs are required for a robust signal (Fig. 4.8 a). Precipitation is significantly higher at the sites dominated by forests in all regions, despite a considerable spread among individual station pairs, with the exception of Region 5 from April to July (green bars in Figs. 4.8 b, d and Fig. C.11). A stronger signal appears during the winter months, while the summer signal tends to be only slightly positive. Further, the precipitation differences between the sites with more forest minus the site with more AL_r (ΔP^1) is more pronounced under the oceanic climate of Regions 1 and 2 and decreases in strength as the regions become more continental. In relative terms, the median ΔP^1 correspond to 5-15 % and 0-10 % of the total precipitation during winter and summer, respectively (Fig. C.9).

In order to infer the relationship between forestation and precipitation in a spatially-explicit way across Europe, we use a GAM, which we first validate here at the site pairs. The ΔP^1 predicted by the GAM at the site pairs has a smaller spread than the one of the rain gauge pairs but gives a similar result for the magnitude and seasonality of the median ΔP^1 over the five regions, supported by an index of agreement (IA; Duveiller et al., 2016) of 0.80 (red/'all forcings' in Fig. 4.8 b-d and Fig. C.11). To ensure that the observed ΔP^1 does not originate from systematic differences in variables other than the local AL_r and forest LC fractions, we make a second GAM prediction, using the forcing data from the first station for all predictors at both sites except for these two LC fractions (orange/'LC only' in Fig. 4.8 b-d and Fig. C.11). While the signal in this GAM prediction is somewhat weaker in all regions, it shows a similar pattern as the paired sites (IA of 0.61), indicating that most of the precipitation difference at the site pairs can be explained by the difference in the AL_r and forest LC fractions. For a second confirmation, we fitted another GAM without the LC-related predictors to the MSWEP precipitation climatology. Reassuringly, the precipitation difference predicted by this GAM for the site pairs captures only a small fraction of the signal from the site pairs (blue/'no LC' in Fig. 4.8 b-d and Fig. C.11, IA of 0.052).

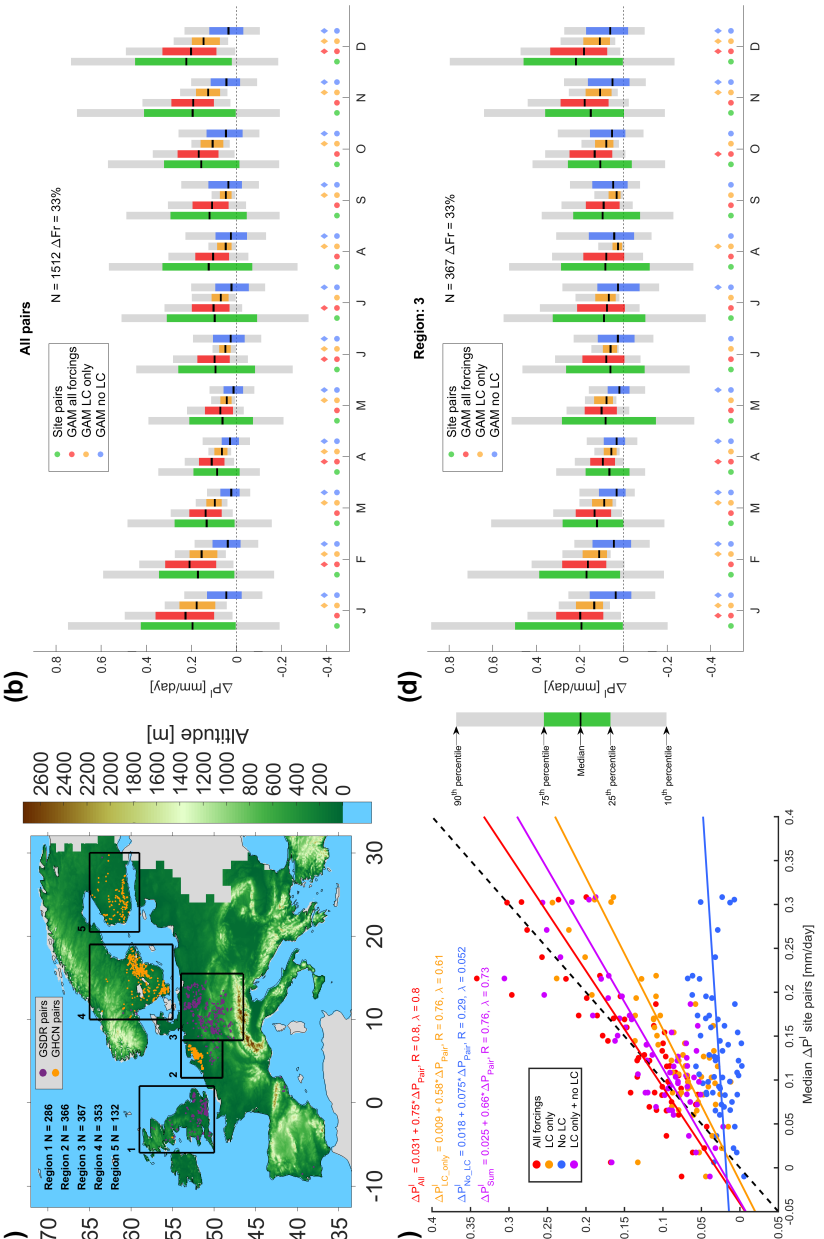

(a)
Region 1 N = 286
Region 2 N = 366
Region 3 N = 367
Region 4 N = 353
Region 5 N = 132
GSDR pairs
GHCN pairs
Altitude [m]
(b)
All pairs
N = 1512 ΔFr = 33%
Site pairs
GAM all forcings
GAM LC only
GAM no LC
ΔPI [mm/day]
(c)
Median ΔPI GAM [mm/day]
Median ΔPI site pairs [mm/day]
All forcings
LC only
No LC
LC only + no LC
90th percentile
75th percentile
Median
25th percentile
10th percentile
(d)
Region: 3
N = 367 ΔFr = 33%

Figure 4.8: ***(Previous page) Local precipitation difference between forest and agricultural land at site pairs.*** *Panel (a), locations of the site pairs in GSDR (purple) and GHCN (orange). To the right the median (black line), interquartile range (coloured shading), and range between 10^{th} and 90^{th} percentile (grey shading) of the monthly precipitation difference between the site with more forest coverage minus the site with more rainfed agricultural land (ΔP^l) over all site pairs (b) and Region 3 (d), which exhibits an intermediate signal strength (other regions are shown in Fig. C.11). Bars correspond to distribution amongst the site pairs in the rain gauge data sets (green), the generalized additive model (GAM) prediction using all the forcings at the individual stations (red), the contribution to ΔP^l in the GAM from the differences in the local AL_r and forest fractions (orange), and prediction of a GAM that was fitted without the land cover (LC) information (blue). Coloured circles indicate samples which are significantly different from zero in a t-test and diamonds GAM prediction samples which are significantly different from the sample of the site pairs in a Welch's t-test both at 5 % confidence level. N is number of site pairs and ΔFr median difference in forest/rainfed agricultural land fraction among site pairs used to estimate the contribution from those LCs in the GAM. Panel (c) shows monthly medians of ΔP^l over the five regions in the site pairs against monthly medians of the all-forcing GAM (red), LC only GAM (orange), no LC GAM (blue), and medians of LC only GAM plus medians of no LC GAM (purple). R denotes Pearson correlation coefficient and λ the index of agreement.*

4.3.2 Theoretical effect of uniform forestation on precipitation

We employ the GAM that was evaluated in the previous section to estimate the effect of a 20 % increase in forest cover uniformly across Europe, with the exception of Northern Scandinavia, where no agricultural land is available for forestation (Fig. 4.7). Locally, such a forestation increases precipitation in almost all locations with a mean of 0.087 mm/day across the area that is reforested (Fig. 4.9 c). $\Delta P^l_{20\,\%}$ is more positive during winter (mean of 0.13 mm/day) and tends to be stronger at locations close to the coast (Fig. 4.9 a). During the summer months, precipitation locally increases only weakly or even decreases slightly due to forestation (Fig. 4.9 b, mean of 0.043 mm/day). Besides the overall local increase in precipitation associated with forestation, we find substantial downwind effects from this LCC on precipitation ($\Delta P^{rem}_{20\,\%}$). Foresting 20 % of the land surface decreases winter downwind precipitation over Northern Europe, exhibits a weak signal in Central and Eastern Europe, and increases precipitation in coastal areas of Western and Southern Europe (Fig. 4.9 d, mean of 0.043 mm/day). On the other hand, the GAM predicts a downwind increase due to forestation in summer precipitation over most of Europe (Fig. 4.9 e, mean of 0.17 mm/day). The $\Delta P^{rem}_{20\,\%}$ during winter originates mostly from forestation in regions close to the coast, while the relevance of forestation in continental regions for downwind precipitation is confined to the summer season (Supplement C.3).

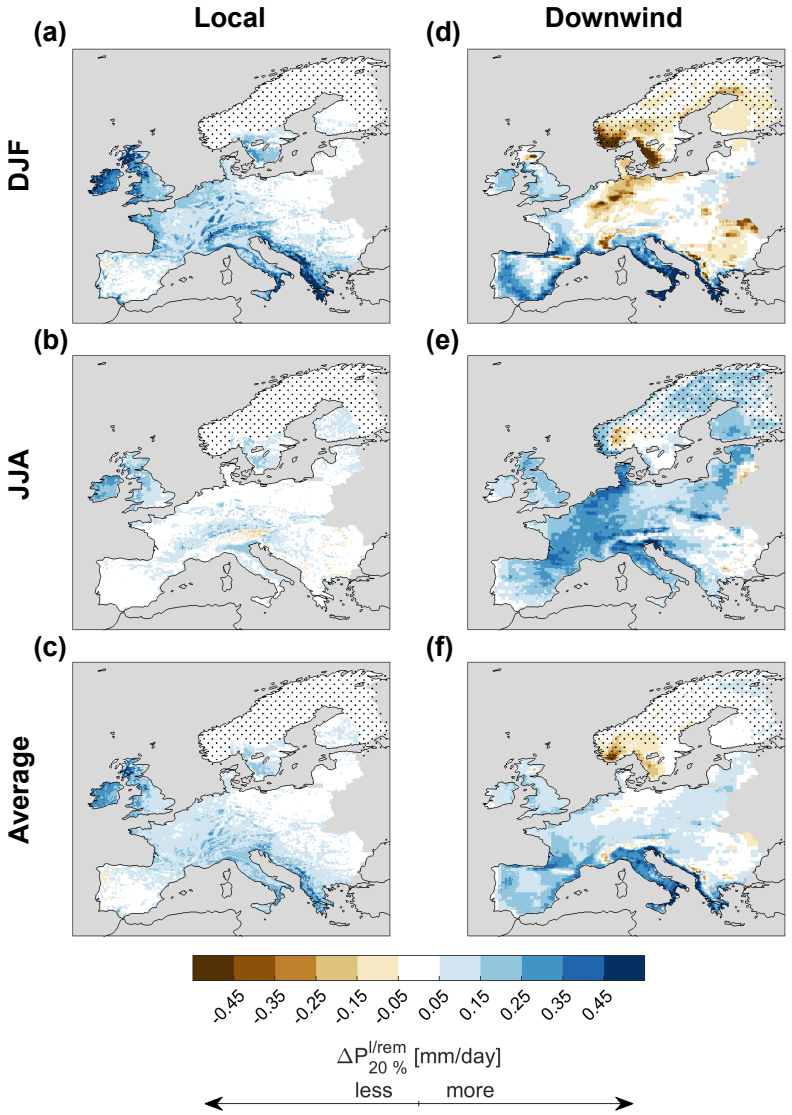


Figure 4.9: ***Theoretical precipitation change from foresting 20 % of the land surface across Europe.*** *Estimated $\Delta P^{l}_{20\,\%}$ (a-c) and $\Delta P^{rem}_{20\,\%}$ (d-f) after foresting 20 % of the land surface in the entire study domain. Top row boreal winter, middle row boreal summer, and bottom row annual mean.*

4.3.3 Implications for a realistic forestation scenario

The forestation scenarios presented so far were stylized and might not be applicable in all regions of Europe. Therefore, we estimate the changes in precipitation if the reforestation potential according to the Global Reforestation Potential Map (Griscom et al., 2017a; Griscom et al., 2017b) was realized. This map was constrained by safeguards for food and fibre security, biodiversity, and against adverse biogeophysical temperature effects (Lee et al., 2011; Li et al., 2015; Duveiller et al., 2018b). In total, 14.4 % of the land surface in the study domain is suitable for forestation, with concentrations over the British Isles, western and southern France, Portugal, Italy, and Eastern Europe (Fig. 4.10 j). The ΔP^{1}_{ref} due to this scenario often exceeds 0.1 mm/day during winter, while the summer signal rarely exceeds this value except for the British Isles. Additionally, this forestation scenario triggers substantial non-local changes in the precipitation field (Fig. 4.10 d-f). These downwind effects reach regions with no significant forestation potential, such as Scandinavia, where reforestation was prohibited due to the local biogeophysical warming effect of forests in those areas (Griscom et al., 2017a; Griscom et al., 2017b). Adding the two contributions up, precipitation is increased by 0.14 mm/day in winter, 0.13 mm/day in summer, and 0.16 mm/day for the annual mean averaged over the continent. The precipitation change amounts to more than 10 % of the annual mean precipitation for 27 % of the European area (Fig. 4.10 l).

The precipitation changes induced by realistic forestation have the potential to compensate part of the climate change signal, as estimated by ten EURO-CORDEX EUR-11 ensemble members (Fig. 4.11). During summer, realistic forestation is estimated to alleviate the precipitation decreases under RCP4.5 in the Mediterranean, Atlantic, and the Continental biogeographical regions (Fig. 4.10 k) and to compensate a considerable fraction of the precipitation signal under RCP8.5 (Fig. C.12). Therefore, forestation could play a crucial role in adapting to the increased risks of summer droughts due to climate change (Sheffield and Wood, 2008). Further, forestation is estimated to attenuate the wintertime precipitation increase in the Boreal region and even overcompensate the precipitation reduction in the Mediterranean (Fig. 4.10 k). On the other hand, it could also have potentially adverse effects by further intensifying climate change-induced precipitation increases in the Atlantic region. It needs to be noted that the forestation-induced precipitation changes presented here are determined under recent climatic conditions and may be altered under a warmer climate. Also, the EUR-11 ensemble exhibits substantial (mostly positive) biases in the simulated precipitation distribution of the recent past (Kotlarski et al., 2014; Prein et al., 2015), which could also affect the estimated precipitation changes due to climate change.

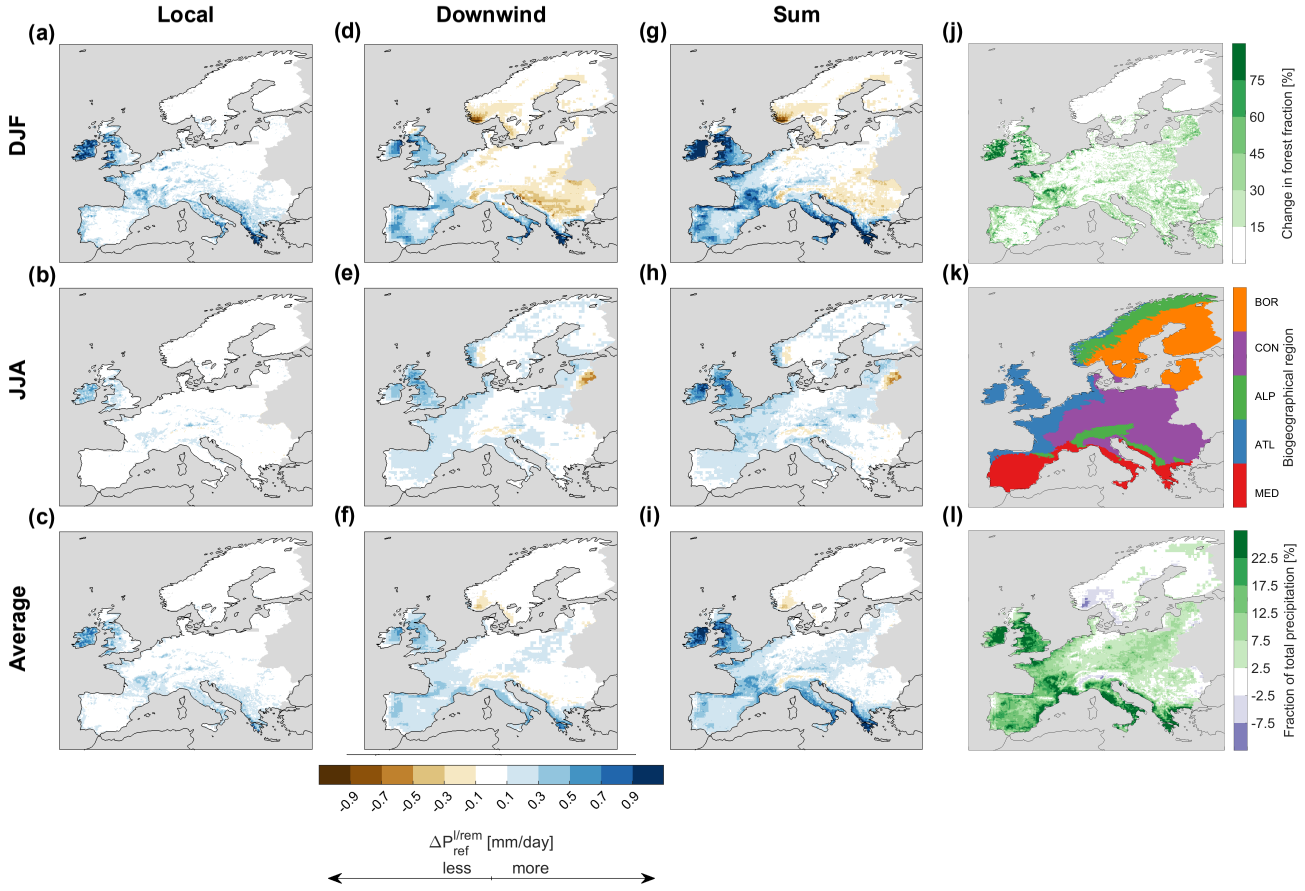
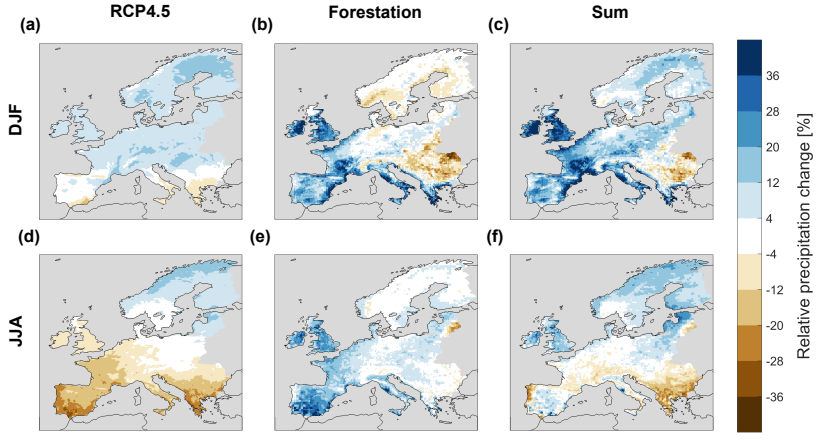


Figure 4.10: ***Potential precipitation change from realistic forestation.*** *Estimated ΔP^{l}_{for} (a-c), ΔP^{rem}_{for} (d-f), and total precipitation change (g-i) associated with forestation in panel (j). Top row boreal winter, middle row boreal summer, and bottom row annual mean. Panel (k) shows biogeographical regions over which averages are calculated in Fig. 4.11 (g) and panel (l) the precipitation change in panel (i) as a fraction of the annual mean precipitation.*



(g)

ΔP ± MAD [%]	DJF			JJA		
Region	RCP 4.5	Forestation	Sum	RCP 4.5	Forestation	Sum
Mediterranean	-0.19 ± 4.9	**23 ± 9.7**	23 ± 11	-19 ± 3.7	**19 ± 8.9**	-0.63 ± 9.1
Atlantic	6.3 ± 2.3	16 ± 11	23 ± 11	-8.7 ± 6.0	**14 ± 5.6**	5.7 ± 7.7
Alpine	7.4 ± 4.0	**-0.79 ± 4.2**	6.7 ± 5.6	3.2 ± 9.9	1.7 ± 2.5	4.9 ± 11
Continental	7.9 ± 3.2	**-1.0 ± 9.3**	6.9 ± 10	-6.0 ± 5.2	**4.8 ± 3.6**	-1.2 ± 6.4
Boreal	11 ± 2.2	**-2.1 ± 4.1**	8.6 ± 4.5	7.6 ± 3.1	2.9 ± 3.3	11 ± 5.1

Figure 4.11: ***Comparison of precipitation changes from climate change and forestation***. *Change in precipitation in EURO-CORDEX EUR-11 multi model ensemble under RCP4.5 scenario between 1986-2015 and 2071-2100 relative to modeled precipitation climatology of 1986-2015 (a, d), estimated precipitation changes due to realistic forestation scenario relative to precipitation climatology of 1986-2015 in MSWEP (b, e), and sum of the two (c, f) in boreal winter (a, b, c) and boreal summer (d, e, f). Below, area-weighted average plus/minus area-weighted mean absolute deviation from average of relative precipitation change under RCP4.5, following forestation, and sum of the two over biogeographical regions (Fig. 4.10 k). In bold numbers of forestation that reduce the climate change signal.*

4.4 Discussion and conclusions

Here, we provide a first observational estimate of forestation-induced changes in precipitation over Europe. Therefore, it is difficult to assess the robustness of our estimate with previous studies. Forest differs in several biogeophysical characteristics from agricultural land, which can explain the identified precipitation changes. Firstly, forests typically have a higher surface roughness than agricultural land. We hypothesize that this might alter the local precipitation amount by inducing more turbulence and slowing movement of precipitating air masses. The increased surface roughness over urban areas has been identified as an important contributor to the amplification of precipitation over cities and downwind of cities (Liu and Niyogi, 2019). Similarly, the high surface roughness of forests was an important contributor to the precipitation increase from afforestation in the Sahel region in a model experiment (Yosef et al., 2018). Further supporting this hypothesis, we find that higher wind speeds, as an anti-correlated proxy for the surface roughness, are associated with lower precipitation in particular during winter (Supplement C.4). The surface roughness has also implications for ΔP^{rem}. A decreased number of cyclones was found following large-scale reforestation of Europe in a regional climate model, mainly due to the high surface roughness of forests (Belušić et al., 2019). Especially forests close to the coast might therefore hinder the propagation of frontal systems into the interior of the continent. As cyclones contribute more to the total precipitation during winter than during summer (Van der Ent and Savenije, 2011; Rüdisühli et al., 2020), this can explain why forestation leads to a downwind increase of precipitation in areas close to the coast during winter, while the signal is neutral in Central and Western Europe and even negative in Northern Europe. Indeed, forestation is estimated to decrease the propagation speed of the calculated wind trajectories further away from the coasts, leading to a reduction of the rainfall amount (Supplement C.4). Secondly, observations indicate that forests typically sustain higher evapotranspiration than AL_r in particular during the summer season (Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b). We hypothesize that this is an important driver behind the downwind summertime precipitation increase from forestation in most locations of Europe, which is supported by applying a moisture source diagnostic to the wind trajectories used in this study (Supplement C.5). Higher evapotranspiration of forests was also linked to increased precipitation in the tropics (Spracklen et al., 2012; Lejeune et al., 2015) and the Sahel region (Yosef et al., 2018). Thirdly, forestation warms the land surface during winter but cools it during summer (Li et al., 2015; Schultz et al., 2017; Duveiller et al., 2018b). This could explain the seasonal cycle of the local signal we observe, as warmer temperatures at the land surface destabilize the planetary boundary layer, thereby favouring the creation of precipitation. In fact, the urban heat island effect was linked to locally increased precipitation (Liu and Niyogi, 2019). Overall, future sensitivity experiments with models could provide a valuable contribution to understand the processes underlying the signals observed here.

Despite the fact that our results are consistent with the current understanding of mechanisms at play, a number of potential caveats in the analysis should be highlighted. Our results are based on a spatial rather than a temporal comparison of the precipitation amount over different LCs. In consequence, the observed ΔP^{l} can be the result of a spatial redistribution of precipitation rather than an alteration of the overall amount of

precipitation, which would imply that the estimated ΔP^1 is not directly attributable to the effect of forestation at larger scales. Further, causality is inherently uncertain in such observation-based approaches. By choosing precipitation as a response variable and LC as a predictor, we implicitly assume that the latter influences the former. In reality, the distribution of the different LC types may also be influenced by precipitation amounts. However, the results of the site pair analysis are robust against the stringency of the selection criteria (Supplement C.1). As for the GAM analysis, the influence of potential confounding factors is minimized by including them as predictors. While the two estimates of ΔP^1 based on different data sets and methodologies show good agreement, they are not fully independent, as MSWEP v2.2 incorporates rain gauge measurements besides reanalysis products and remote sensing data sets. Hence, we cannot exclude that systematic biases in the rain gauge measurements translate to the GAM analysis. The undercatch of rain gauges measurements is exacerbated by high wind speeds (Pollock et al., 2018). As forests typically have a higher surface roughness than AL_r, lower surface wind speeds occur at the sites in regions with more forest (Supplement C.4). This could result in weaker undercatch at those sites compared to the sites with more AL_r in their proximity. Further, the ΔP^1 in the GAM could not be evaluated in the southern half of Europe, due to the scarcity of rain gauge data in those areas. Thus, the compensatory effect of forestation on the projected reduction of precipitation due to climate change in Southern Europe is prone to larger uncertainties than the precipitation changes from forestation in other regions (Note however that the GAM exhibits a similar signal as the rain gauge data over the Iberian Peninsula with relaxed selection criteria; Fig. C.13). Uncertainties are larger for ΔP^{rem}, for which we rely solely on the GAM. Overall, our results are consistent with observations in the tropics (Spracklen et al., 2012) and modeling studies (Lejeune et al., 2015; Belušić et al., 2019). Nevertheless, the estimated downwind effect of forestation is sensitive to both the construction of the GAM and the starting heights of the wind trajectory calculation, although qualitatively robust (Fig. C.3 and Supplement C.2). We therefore encourage a more thorough analysis of ΔP^{rem}.

Overall, our results imply that forestation could trigger substantial changes in precipitation over Europe. Forestation-induced precipitation changes appear to be subject to spatial trade-offs due to downwind effects. While we find a local increase in precipitation due to forestation across Europe, forestation might reduce precipitation further downwind in winter. On the other hand, forestation increases precipitation downwind in summer, likely due to higher moisture supply by forests than by AL_r. This downwind enhancement of precipitation might come at the cost of a local reduction of runoff (Trabucco et al., 2008; Padrón et al., 2017), adding yet another dimension to the spatial interconnections. Overall, our results highlight that LCCs, such as forestation, can alter precipitation in the mid-latitudes considerably, both locally and further downwind. Hence, the consequences of human land-use for water availability should be considered alongside biogeochemical effects and the biogeophysical alteration of temperatures. As droughts are projected to become more severe with changing climate in Europe (Sheffield and Wood, 2008), the interplay between LC and water availability deserves more attention.

Acknowledgements We are thankful for the help and advice of Dr. Andreas Papritz, Dr. Franziska Scholder-Aemisegger, Michael Windisch, and Dr. Anina Gilgen. We thank three anonymous reviewers for their valuable comments. This work is part of the CLIMPULSE project, which was funded by the Swiss National Science Foundations (SNSF; http://p3.snf.ch/Project-172715; grant no. 200021_172715) and the Swiss Federal Office for the Environment (FOEN).
The GSDR data analyzed in this study contain measurements of the following Institutes: Service Puplic de Wallonia, Finnish Meteorological Institute, Météo-France, Deutscher Wetterdienst, Met Éireann, Meteo Trentino, Agrometeorologico Siciliano, Autonome Provinz Bozen–Südtirol, Norwegean Meteorological Institute, Sistema Nacional de Informação de Recursos Hídricos, Instituto Português do Mar e da Atmosfera, Servei Meteorologic de Catalunya, Meteo Schweiz, UK Met Office (Met Office (2006): MIDAS UK Hourly Rainfall Data. NCAS British Atmospheric Data Centre, May 5th 2020. https://catalogue.ceda.ac.uk/uuid/bbd6916225e7475514e17fdbf11141c1), Environment Agency UK, the Scottish Environment Protection Agency (Contains public sector information licensed under the Open Government Licence v3.0), and Natural Resources Wales (Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved). The post-processed model data of CORDEX are provided by the Center for Climate Systems Modeling (C2SM), ETH Zurich (Jan Rajczak, Silje Soerland, Urs Beyerle, Curdin Spirig, and Elias Zubler).

5

Conclusions and outlook

5.1 Conclusions

This thesis examines the biogeophysical effects of foresting grassland or cropland. Overall, there is growing evidence that these effects should be considered when assessing the climatic consequences of LULCCs. The biogeophysical effects of forestation exhibit complex variations in space and time. In addition, they act both at a local scale and a regional scale, which poses a challenge for modeling and observational studies. Due to these complexities, our understanding of the biogeophysical effects of forestation is still incomplete. With this thesis, I contribute to close some of the remaining knowledge gaps.

The local LST response is probably the best-constrained biogeophysical effect of forestation, besides $\Delta\alpha^{\downarrow}$. Forests cool the daily mean and maximum LST in observations with the exception of the winter months at higher latitudes when snow is present. In Chapters 2 and 3, I compare the observed LST response to the sub-grid LST difference of the forest tiles minus the grassland and cropland tiles in offline CLM simulations. By and large, CLM is able to reproduce the daily mean and maximum LST signal given a fair comparison. It is crucial to place the different PFTs on separate soil columns when extracting the sub-grid signal to suppress lateral energy exchange in the soil, which is not realistic at the spatial scales of the observational studies. CLM4.5 exhibits a positive bias for $\Delta\mathrm{LST}^{\downarrow}_{\mathrm{max}}$ compared to observations, which appears to be related to a negative bias in $\Delta\mathrm{ET}^{\downarrow}$. Such biases may be addressed by better constraining the vertical distribution of the roots, refining the dynamics of plant water uptake, and/or parameter choices for the parametrisation of photosynthesis. A combination of these measures reduces both the positive $\Delta\mathrm{LST}^{\downarrow}_{\mathrm{max}}$ bias and the negative $\Delta\mathrm{ET}^{\downarrow}$ bias in CLM4.5. It needs to be noted however, that spatially-continuous observational ET products exhibit considerable uncertainties themselves, particularly for individual LCs. Also, atmospheric feedbacks to forestation are neglected in the first two studies owned to the offline simulations, while they cannot be excluded entirely in observations, which impairs the comparability of the sub-grid signal and the remote sensing observations.

During the night, forestation appears to cool the LST only slightly in the tropics and even warm it outside of the tropics in observations. CLM exhibits a distinct negative bias in ΔLST^{1}_{min} in comparison to remote sensing estimates. Chapter 3 demonstrates that the neglect of BHS is an important contributor to this bias. The vegetation biomass absorbs and stores energy during the day, in particular the stems of trees. In offline simulations, this cools daytime temperatures at the land surface only slightly, as the energy uptake by the biomass is compensated by a reduction of the turbulent heat fluxes. During night, energy is released from the biomass. Due to the stable structure of the surface layer at night, the compensation by the sensible heat flux is weak. Therefore, the land surface warms and releases the energy radiatively. This warming exceeds 2 K frequently in forests. BHS thus appears to be a relevant process for the diurnal temperature cycle in forests.

Chapter 4 estimates the local and downwind response of precipitation to forestation in Europe using rain gauge pairs and GAMs. Locally, forestation increases precipitation in particular during the winter months. By linking the spatial distribution of precipitation to the LC overpassed prior to precipitation events, I further determine that forestation could have important implications for the precipitation further downwind. Most importantly, forestation increases precipitation downwind during the summer months. This implies that the land is an important moisture source during the warm season, which is confirmed with a moisture source diagnostic. The estimated changes in precipitation from a realistic reforestation scenario are of comparable magnitude as the climate change signal under RCP4.5 by the end of this century. Therefore, forestation has not only relevant biogeophysical effects on temperature but also on the hydrological cycle, which deserves more attention in future studies. It remains open however, whether the relations between precipitation and the (upwind) LC fractions can be upscaled directly to forestation scenarios at larger scale. In particular the local signal might be rather the result of a spatial redistribution of precipitation than an overall increase. The finding of locally increased precipitation over forests has important implications for studies employing the Budyko framework, which examines the relationship between the fraction of ET and precipitation and the fraction of potential ET and precipitation. Differences in those fractions among catchments dominated by different LCs may not only originate from the amount of ET in those catchments but also the amount of water available from precipitation, which can be influenced by LC. For example, a catchment dominated by forests may exhibit a lower fraction of ET/P than a catchment dominated by agricultural land not because of lower ET but because of higher precipitation.

Overall my thesis demonstrates that forestation has important biogeophysical implications for the local and regional climate. It is crucial to examine and understand these effects before employing forestation as a mitigation tool for climate change, as frequently proposed. However, the biogeophysical effects of forestation are still prone to considerable uncertainties both in ESMs and observations. In particular there are three aspects that still require further attention: (1) How does forestation affect different temperature metrics at the land surface and which ones are most relevant for life on earth? Several modeling studies find that ΔLST^{1} can differ in sign from $\Delta T2M^{1}$. In consequence, the cooling potential of forests found in remote-sensing studies might be restricted to LST. On the other hand, forestation might warm T2M, which could be more relevant

than LST for human thermal comfort and the functioning of ecosystems. (2) What are the remote biogeophysical effects of forestation? Especially, we need to know at which spatial scales of forestation they become relevant. Further it would be valuable to conduct more observational studies on remote effects to constrain them in models. (3) More attention should go towards the implications of forestation for the hydrological cycle, as water is the essence of life and therefore crucial both for humans and wildlife. The current literature on this topic frequently exhibits contradictory results, which prevents us from drawing robust conclusions on how forestation affects water availability.

5.2 Outlook

5.2.1 LULCC beyond forestation

Conversions between forests and grassland/cropland were and still are the most prominent anthropogenic LULCC. Nevertheless, humans alter the properties of the land surface in other ways too. This sections provides a brief insight into the literature on the biogeophysical effects of some LULCC beyond de-, re-, and afforestation. Broadly, LUCs can be divided into land use conversions and land management (Pongratz et al., 2018). The first category corresponds to a change in human land use (e.g., abandonment of agricultural land) and is often associated with complete conversion of the LC type (e.g., conversion of cropland to natural forests). The second category refers to the way the land is managed to generate its benefit for humans (e.g., whether cropland is tilled and which species are cultivated on it).

Other LU conversions Urbanization is probably the most drastic anthropogenic LCC, altering the biogeophysical land properties completely. Despite covering only about 0.5 % of the land surface (Hurtt et al., 2020), urban areas still deserve attention due to their pronounced biogeophysical effect and their importance as a habitat for humans. Observed air temperatures are generally several degrees warmer in urban areas compared to nearby rural areas (Tzavali et al., 2015), a phenomena called the urban heat island effect (Fig. 5.1). This warming is most pronounced during the night. Similar as in BHS, the buildings in a city form an energy storage reservoir, which releases energy during the night that has been absorbed during day (Rizwan et al., 2008). In addition, the emitted longwave radiation of urban surfaces is frequently re-absorbed due to the low sky-view factor in cities, trapping the stored energy (Zhang et al., 2019). The fraction of the world population living in urban

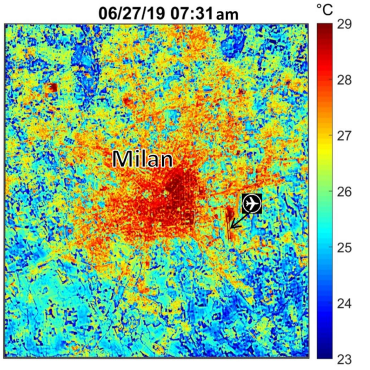


Figure 5.1: *LST in Milan during a heat wave. The city center appears much warmer than the surrounding areas. Credit: NASA/PL-Caltech.*

areas has increased strongly over the last decades, reaching more than 50 % today, and is expected to increase further (World Urbanization Prospects, 2018). The matter of the urban heat island effect and in particular strategies to mitigate this effect deserve therefore certainly attention. Unsurprisingly, a number of mitigation strategies have been proposed in the literature, often including the introduction of more vegetation in urban areas (e.g., Rizwan et al., 2008; Zhang et al., 2017).

Land management Besides the LUCs that incorporate a complete conversion from one LC type to another, such as deforestation or urbanization, humans also alter climate through land management, which can have a biogeophysical effect comparable to the one of a complete LC conversion (Luyssaert et al., 2014). LCCs associated with European forests since 1750 exerted a warming effect on summertime temperatures, mainly owned to forest management rather de-, re-, and afforestation (Naudts et al., 2016). The biogeophysical effect of species conversion from broadleaf to needleleaf trees was the dominant contributor to this signal, highlighting the importance of both biogeophysical effects and land management at least for the regional climate. Similarly, Luyssaert et al. (2018) find that the forest management strategy that minimizes the T2M in Europe differs substantially from the strategy that maximizes the land carbon sink. In detail, increasing the fraction of broadleaf trees in European forests would be beneficial for the local temperatures. Similarly, we find in an observational study that was part of the CLIMPULSE project that replacing needleleaf with broadleaf trees mitigates warm temperature extremes (Box 2).

Land management has not only been an important climatic driver in the past and present, but provides also an opportunity to mitigate and adapt to climate change while maintaining the services provided by land to humans. Besides forest management and measures to mitigate the urban heat island effect, management practices for cropland are another promising tool. Increasing the albedo of cropland reduces the mean temperature and even more so temperature extremes (Seneviratne et al., 2018). This could be achieved for example through no-tillage farming (Davin et al., 2014), cover crops, or low-chlorophyll crops (Lugato et al., 2020). Irrigation is another management practice that reduces temperatures in regions where it is applied, by repartitioning more of the available energy into latent heating (Thiery et al., 2017).

Box 2: Increasing the broad-leaved tree fraction in European forests mitigates hot temperature extremes

Introduction Not only the presence of forests but also the way they are managed affects the local climatic conditions. In this study, we estimate the impact of increasing the broadleaf tree fraction (BTF) in European forests on LST. I here provide a brief synthesis of the results from this study. For the full publication refer to Schwaab et al. (2020).

Methods Generalized additive models are fitted to the LST climatologies for a given month and time of day from MODIS (1 km resolution, up to four samples per day; Wan, 2014) and SEVIRI (Spinning Enhanced Visible and InfraRed Imager, 3-5 km resolution, 15 min temporal resolution; Trigo et al., 2008). The models are calibrated for five biogeographical regions individually, using as predictors the forest type from the Copernicus Land Monitoring Service (Langanke, 2018) to distinguish between broadleaf and needleleaf forest, 15 LC categories from CORINE Land Cover (Kosztra et al., 2019), elevation, slope, exposition, TRI, and TPI based on EU-DEM v1.0 (EEA, 2014), and the geographic coordinates. Further, GAMs are produced for the 0.1, 0.2, ..., 0.9 quantiles of the daily maximum LST in SEVIRI during summer. Finally, the effect of increasing the BTF by 80 % is calculated from the difference of a GAM prediction with a BTF of 90 % minus a prediction with a BTF of 10 %.

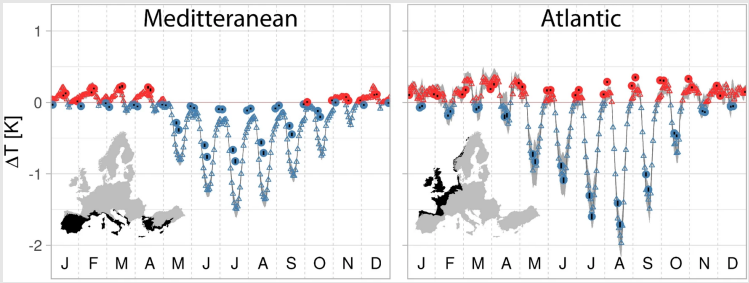


Figure 5.2: *Monthly diurnal cycle of LST change induced by an increase of the BTF by 80 % in five biogeographical regions. For every month denoted as J (January), . . . , D (December) on the x-axis the mean diurnal cycle of the LST change induced by an 80 % increase of the broad-leaved tree fraction is shown. For SEVIRI (triangles) hourly mean values are available (i.e. 24 values for each month). For MODIS (circles), observations are available at four different time steps per day which are approximately at 01:30, 10:30, 13:30 and 22:30 local solar time. Confidence intervals for SEVIRI are shown in grey. Confidence intervals for the MODIS observations are denoted as black lines through the colored dots. Red markers stand for a warming due to an increase in the BTF and blue markers for a cooling. Adapted from Schwaab et al. (2020).*

Results The LST response to increasing the BTF exhibits a distinct diurnal and seasonal cycle (Fig. 5.2). Broadleaf trees strongly decrease the LST during the day and slightly increase the LST during night. The daytime cooling is most distinct during the summer months, while there is only a slight cooling or even a warming by broadleaf trees during winter. The summertime cooling effect on the daily maximum LST due to an increase in the BTF increases in strength for the higher temperature percentiles (Fig. 5.3). As such, broadleaf trees appear to be particularly beneficial during heat waves.

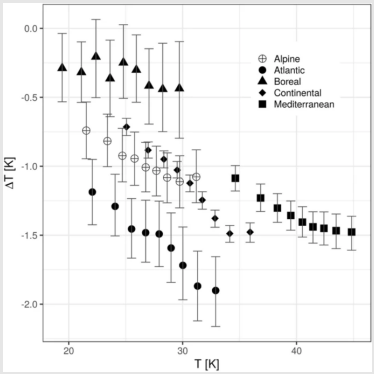


Figure 5.3: *LST change (based on SEVIRI) related to an increase of 80 % in BTF for different temperature quantiles, going from the LST change at moderate background temperatures on the left (0.1 quantile) to very high background temperatures on the right (0.9 quantile). The error bars show the 95 % confidence interval. From Schwaab et al. (2020).*

5.2.2 Future avenues for modeling studies

Improvements for aboveground energy storage scheme in CLM In Chapter 3, I demonstrated that BHS is a relevant process for the diurnal LST evolution in forests. While this first implementation was able to capture the broad nature of BHS, several aspects regarding the aboveground energy storage could still be improved. In detail, the simulated BHF could be evaluated more thoroughly with field site studies. Besides using the model parameters associated with the current implementation of BHS, a more complex BHS scheme could improve the realism of the simulated BHF. The current version assumes a uniform temperature within the stem, which is of course a strong simplification. Haverd et al. (2007) successfully applied the force-restore method to simulate BHS at a specific site. This method was originally designed to represent the soil energy storage in general circulation models (Bhumralkar, 1975). Another option could be to represent several stem layers, similar to the representation of the ground. Herrington (1969) provides a detailed proposition on how this could be done including choices for model parameters. Finally, the sensible and latent heat stored in the canopy air space is another energy storage component currently missing in CLM. Bonan et al. (2018) developed a multi-layer canopy version of CLM4.5 including a new parametrisation for the roughness sub-layer. This development brings along the representation of the energy stored in the canopy air, but also increases the complexity of the canopy model substantially compared to the current big-leaf approach. Heidkamp et al. (2018) use the

canopy height to estimate the size of the energy reservoir in the canopy air space, which is compatible with the big-leaf approach. During my PhD, I have implemented both the force-restore method and the energy storage in the canopy air space after the example of Heidkamp et al. (2018) (Box 3). However, these model developments require more testing and evaluation with data from in-situ sites.

Box 3: Further refinement of canopy energy storage in CLM5.0

Introduction The implementation of the canopy energy storage presented in Swenson et al. (2019) and Meier et al. (2019) can be refined on several aspects. Here, I present preliminary results of a model version that uses the force-restore method for the energy storage in the stem after Haverd et al. (2007) and the simple method representing the sensible and latent heat storage in the canopy air of Heidkamp et al. (2018).

Methods In the force-restore method, the stem is presented by the temperature of the bark and the temperature of the core. The calculation of the core temperature tendency remains similar to the original scheme (Swenson et al., 2019). However, the emitted longwave radiation and the sensible heat flux from the stem are calculated from the bark temperature. The bark temperature tendency is computed after Eq. (31) in Haverd et al. (2007). For the resistance to heat transfer per meter of stem diameter (r_w in Chapter 3), $300\,\mathrm{s\,m^{-2}}$ is chosen for the force-restore method instead of $1000\,\mathrm{s\,m^{-2}}$ in the original scheme. It is assumed that the air volume can be calculated from the canopy height to estimate the sensible and latent heat storage in the canopy air space. The sensible and latent heat storage terms are then computed from the change in the canopy air temperature and canopy specific humidity, respectively.

Four point scale simulations are run in the satellite phenology mode for 1997-2010, where the first five years correspond to the spin up period: (1) a simulation without any storage terms in the canopy corresponding to CLM-CTL in Chapter 3; (2) a BHS simulation with the original scheme (CLM-BHS); (3) a simulation using the force-restore method to compute the energy storage in the stem; (4) a simulation using both the force-restore method and the energy storage terms in the canopy air space. The simulations are conducted at 2.75° S/60.25° W, corresponding to the location of Michiles and Gielow (2008). This location is dominated by a dense broadleaf evergreen forests with an aboveground stem biomass of $104\,\mathrm{kg\,m^{-2}}$ and a canopy height of 29.2 m.

Results The energy storage in the stem with the force-restore method tends to be smaller than with the original scheme and both the positive and negative peak occur earlier during the day (Fig. 5.4 b). The sensible and latent energy storage in the canopy air space make a substantial contribution to the overall energy storage in the canopy during the day. They might be overestimated however, as the current implementation is rather rudimentary. All simulations that include BHS exhibit a similar diurnal LST cycle with considerably warmer temperatures than the control simulation at night

(Fig. 5.4 c). Accounting for the energy stored in the canopy air reduces the LST between 07:00 and 11:00 am. Interestingly, the temperature in the canopy air space is colder in the simulations employing the force-restore method than in the simulation with the original scheme throughout the day (Fig. 5.4 d).

Both the way the energy storage in the stem is implemented and the energy storage in the canopy air appear to affect the diurnal temperature cycles at the land surface. Further, the new energy storage schemes exhibit short-term fluctuations, which should be addressed when developing those schemes further. Overall, it would be desirable to thoroughly evaluate the canopy energy storage fluxes with several in-situ studies.

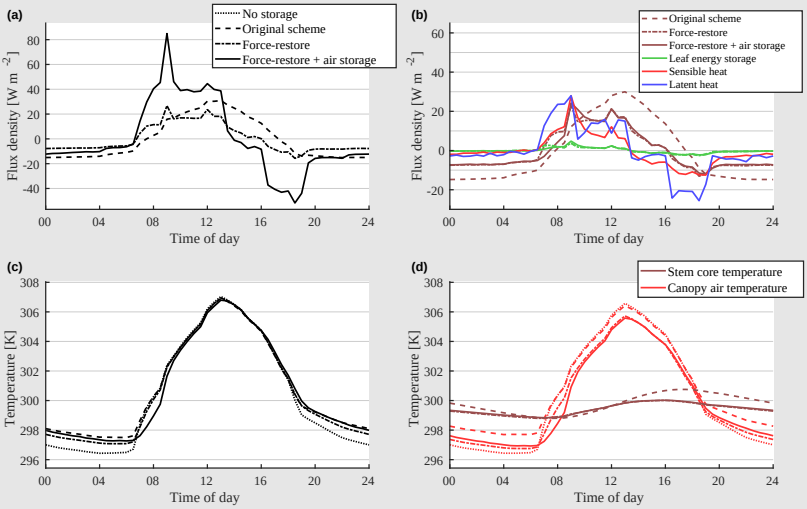


Figure 5.4: *Panel (a), the average diurnal cycle of total energy storage in canopy during boreal winter (DJF) at 2.75° S/60.25° W (location of Michiles and Gielow, 2008) of different schemes. Panel (b), average diurnal cycle of energy stored in stem (brown), leafs (green), and in canopy air space as sensible heat (red) and latent heat (blue). Average diurnal cycle in LST (c), canopy air temperature (red, d), and stem core temperature (brown, d). Dotted lines scheme without canopy energy storage, dashed lines original BHS scheme, dashed-dotted lines force-restore scheme, and continuous lines force-restore scheme with energy storage in canopy air.*

Biogeophysical effects of forestation in convection-resolving simulations Climate and weather models are able to represent deep convection at roughly a horizontal resolution of 4 km or higher, which has important implications particularly for the simulated precipitation fields (Weisman et al., 1997; Panosetti et al., 2020). It is therefore not surprising that the simulated biogeophysical effect of LCC is sensitive to the choice of the horizontal and vertical resolution of the atmosphere (Devanand et al., 2020) and even

agrees better with observational constraints if the model resolution is high enough to resolve convection (Vanden Broucke and Van Lipzig, 2017). Computational capacities have been increasing over the last decades, thereby allowing for higher model resolutions. As of today, there are even ambitions to run global climate simulations with resolved convection (Fuhrer et al., 2018). Therefore, studies employing convection resolving resolutions to study the biogeophysical effect of LCC will likely become more common in future.

5.2.3 Dynamics of forest-related LCC

The estimated biogeophysical effect of forestation from most studies in the field (including the studies in this thesis) build upon a spatial comparison of established forests to other LC types, disregarding the fact that it can take decades to (re-)establish forest. Disturbances of forests, such as fires or bark beetles, cause detectable alterations of the exchange in chemical species and energy over decades (Fig. 5.5; Liu et al., 2005; Randerson et al., 2006; Williams et al., 2012b; Vanderhoof and Williams, 2015). On the other hand, grassland or cropland reaches maturity much faster following deforestation. As such, forest-related LCCs exhibit hysteresis in the sense that the equilibrium state of the vegetation establishes quicker in one direction (deforestation) than in the other (forestation). This has important implications. Firstly, forestation as a mitigation/adaptation tool for climate change requires decades to realize its full benefit. Secondly, foresting the same area to compensate for deforestation elsewhere results in a temporary biogeophysical and biogeochemical effect, even if the two locations exhibit the same climate (i.e., the equilibrium states of the forest and agricultural land at the two locations would be the same). In consequence, ignoring the physiological evolution of forests with their age structure or only simulating net LCCs in ESMs neglects part of the climatic consequences of the imposed LCCs. This is certainly relevant, as gross LCC exceed net LCC considerably (Fuchs et al., 2014; Fuchs et al., 2018).

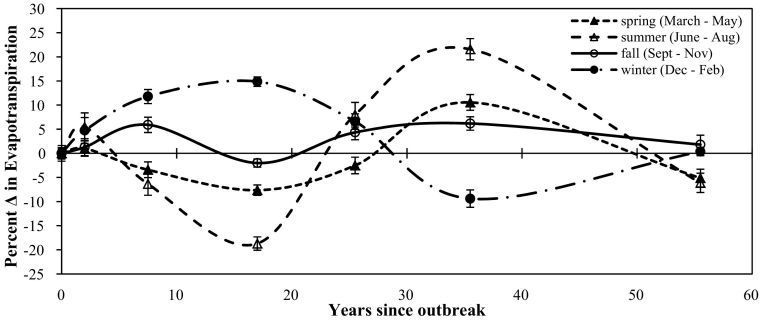


Figure 5.5: *Percent change in seasonal evapotranspiration with years since bark beetle outbreak for 2012. Change is calculated relative to local, for each plot, non-attacked forest. Error bars are plus and minus standard error. From Vanderhoof and Williams (2015).*

Such dynamics could be considered best in LSMs by simulating the age structure of the trees forming the forest instead of using one average tree to represent the entire forest at a given location. Several approaches to simulate vegetation demographic processes in ESMs have therefore been developed over the last two decades (Fisher et al., 2018) and will enable to consider these dynamics when investigating the biogeophysical effects of LCCs.

A

Appendix to Chapter 2

A.1 Sensitivity of CLM4.5 to individual modifications

Here we present a more detailed description and discussion of the individual modifications described in Section 2.3.2. In order to isolate the effect of the individual modifications three additional sensitivity experiments are presented: CLM - ROOT, CLM - 10PER, and CLM - LIGHT. Table A.4 shows which modifications of CLM4.5 are incorporated in the different sensitivity experiments.

A.1.1 Sensitivity to root distribution

In CLM4.5 ET is strongly and positively correlated to Soil Moisture (SM) at most locations, indicating that SM limitation exerts a strong control on the magnitude of ET (not shown). In CLM - DFLT, where SM is the same for all PFTs within a grid cell, forest mostly experiences higher SM stress except for the northern high-latitude winter (Fig. A.4 a). Once the separated soil columns are introduced in CLM - BASE, the differences in the SM stress are also influenced by the differences in SM, which in term are affected by the various ET rates over forest and open land. In other terms, it is possible that forests experience less SM stress than open land but only because they evaporate less water and vice versa (Fig. A.4 b). We argue that the difference in the SM stress of forest minus open land in CLM - DFLT is more representative, because it is unaffected by the ET rates of the individual PFTs in this model configuration. Under this assumption, forests are often more SM-limited than open land in CLM4.5. In contrast, two observational studies comparing SM profiles of forest and nearby pasture sites in the Amazon reveal that forests have a considerably higher capacity to access water from the soil below a depth of 2 m (Jipp et al., 1998; von Randow et al., 2004). Further, there are a number of studies reporting increased forest ET during the dry season due to the higher amount of incoming shortwave radiation, whilst the response is the opposite over pasture (Jipp et al., 1998; da Rocha et al., 2004; von Randow et al., 2004; Huete et al., 2006; Saleska et al., 2007). Altogether these studies indicate that forest ET should be less SM-limited

than open land ET. It is thus possible that forests experience too high and/or open land too little SM stress in CLM4.5.

CLM4.5 accounts for SM stress on vegetation transpiration through a stress function β_t, which ranges from zero (when soil moisture limitation completely suppresses vegetation transpiration) to one (corresponding to no SM limitation on vegetation transpiration). This function is calculated as the sum of the root fraction in each soil layer (r_i) multiplied by a PFT-dependent wilting factor (w_i):

$$\beta_t = \sum_i w_i r_i \tag{A.1}$$

The original root distributions in CLM4.5 were adapted from Zeng (2001) and are rather similar for all PFTs, especially for needleleaf trees, broadleaf deciduous trees, and grassland in the lower part of the soil (Fig. A.5). Therefore, there is no considerable difference in the default configuration of CLM4.5 regarding the ability to extract water from the lower part of the soil between forests and open land PFTs (except for broadleaf evergreen trees). Furthermore, all tree PFTs have a less negative soil matrix potential at which the stomata are fully closed and opened than the open land ones, i.e., tree PFTs have their permanent wilting point at a higher SM content than open land and hence use water more conservatively. In order to increase SM limitation for open land PFTs and thus reduce their ability to extract water from the lower part of the soil, we conduct a sensitivity experiment, called CLM - ROOT, with a much shallower root distribution for open land PFTs. The new values for the root distribution factors (r_a and r_b) are shown in Table A.1 and the resulting root distribution in Fig. A.5.

The modified root distributions strongly reduce the ET of non-arctic open land PFTs, especially ET of C_4 grass (Table A.5). Also, the ET of grassland at the location of the pasture site in the Amazon in the study of von Randow et al. (2004) is considerably reduced during the dry period, even overcompensating the positive bias in CLM - BASE (Table A.6). On the other hand, it does not affect ET during the wet season, when ET is not SM limited. Overall, this experiment reveals that modifying the root distribution has high potential to alleviate biases of CLM4.5 in ET, except for the arctic region where likely temperature and incoming shortwave radiation are the main factors limiting vegetation transpiration.

A.1.2 Sensitivity to dynamic plant water uptake

In the tropics forests often exhibit increased ET during dry periods, due to increased light availability (da Rocha et al., 2004; Huete et al., 2006; Saleska et al., 2007), even though the upper soil is dry, as they still have sufficient water supply from the lower part of the soil (Jipp et al., 1998; von Randow et al., 2004). We aim to allow a similar behaviour in CLM4.5 by introducing a dynamic plant water uptake, where plants only extract water from the 10 % of the roots with the highest wilting factor (i.e., best access to SM) for the calculation of the β_t-factor and the extraction of soil water (example in Fig. A.6). The resulting model simulation, called CLM - 10PER, was conducted by adding this modification to the configuration from the CLM - ROOT experiment.

This modification generally reduces SM stress for plants and hence increases ET for all non-arctic PFTs (Table A.5). Its impact is limited for arctic PFTs where temperature

and shortwave radiation are more important limiting factors of vegetation transpiration than water availability. A notable improvement can be observed for tropical deciduous broadleaf trees for which average ET is increased by 0.11 tmm/day, thereby alleviating the negative bias compared to GETA. Furthermore, it improves the seasonal dynamics of forest ET in the tropics. With the 10 % modification forests show increased ET during the dry period at the forest site of da Rocha et al. (2004). This is the case as trees are now less SM-limited during the dry period than in CLM - BASE, since they have a significant fraction of their roots in the still-moist lower part of the soil, allowing them to exploit the increase in incoming shortwave radiation. On the other hand, ET at the pasture site of von Randow et al. (2004) remains largely unaffected, as grassland has only limited access to SM from the lower part of the soil due to the shallow root distribution introduced in CLM - ROOT. It hence appears that a dynamic plant water uptake could be crucial for the representation of the seasonal dynamics of ET (and possibly photosynthetic activity in general) in the tropics.

A.1.3 Sensitivity to light limitation

As arctic PFTs are only weakly affected by the previously introduced modifications of SM stress as well as the maximum rate of carboxylation described in the next section, we performed a sensitivity experiment with altered light limitation, which is called CLM - LIGHT. Since ET values are strongly negatively biased for boreal deciduous broadleaf trees and C_3 arctic grass (Table A.5), the light limitation of photosynthesis for C_3 plants was lessened by increasing the factor 0.5 in Eq. 8.7 of Oleson et al. (2013) to 0.6. Because ET of C_4 grass exhibits a strong positive bias, their quantum efficiency was reduced from 0.05 to 0.025 mol CO_2 mol^{-1} photon, thereby increasing their light limitation.

Altering the light limitation of photosynthesis impacts ET in all climate zones (Table A.5). Its impact is strongest in the tropics and remains small in boreal regions. Of the C_3 PFTs tropical evergreen broadleaf trees are impacted strongest. The implemented modification alleviates the negative ET bias for evergreen broadleaf trees during the dry season but slightly increases the positive bias during the wet season, overall still leading to a further improvement of the difference between the two seasons (Table A.6). Additionally, the increased light limitation reduces ET of C_4 grass during the wet season similar to the observations over the grassland site in von Randow et al. (2004). This is likely responsible for the increased ET during the dry season as well, since the reduced SM consumption during the wet season is carried over to the following dry season, therefore reducing the SM stress.

A.1.4 Sensitivity to the maximum rate of carboxylation

V_{cmax} appears to be a suitable parameter to tune vegetation transpiration values, since it is not well constrained from observations and vegetation transpiration in models is highly sensitive to this parameter (Bonan et al., 2011). In CLM4.5 the values reported by Kattge et al. (2009) are used except for tropical evergreen broadleaf trees, for which a higher value was chosen to alleviate model biases (Bonan et al., 2012; Oleson et al., 2013). In order to test the sensitivity of the PFT-specific ET values to V_{cmax}, we

conduct a final sensitivity experiment with new values of this parameter in addition to the other modifications presented beforehand, with the aim to alleviate the biases to GETA (Table A.1). Additionally, the minimum stomatal conductance of C_4 plants, which is by default four times larger than that of C_3 plants, was reduced from 40 000 μmol m^{-2} s^{-1} to 20 000 μmol m^{-2} s^{-1} (see Eq. 8.1 in Oleson et al. 2013) in this sensitivity experiment, which we call CLM - PLUS.

As already shown by Bonan et al. (2011), photosynthetic activity of C_3 PFTs is strongly influenced by the choice of V_{cmax}, except for the boreal ones where light or temperature are more important limiting factors of photosynthesis. The CLM - PLUS simulation alleviates biases in ET averaged for the individual PFTs compared to GETA, in particular by reducing ET over temperate evergreen needleleaf trees, both temperate and tropical evergreen broadleaf trees, and C_4 grass, as well as by increasing ET of tropical deciduous broadleaf trees (Table A.5). The mismatch between results of CLM4.5 and the in-situ measurements of von Randow et al. (2004) and da Rocha et al. (2004) in the Amazon region are reduced in this new configuration during the wet season, but enhanced during the dry one (Tables A.6). As in the CLM - LIGHT experiment the reduction of C_4 grass ET during the wet season at the pasture site of von Randow et al. (2004) is partly compensated by an ET increase during the dry period. Overall, ET of C_4 grass compares well with the mean value of GETA. The in-situ observations of von Randow et al. (2004) on the other hand support a stronger tuning for this particular PFT in order to further reduce its ET.

A.2 Appendix figures and tables

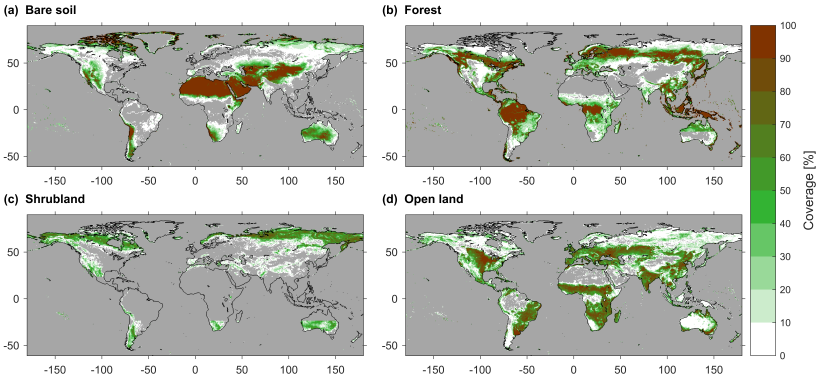


Figure A.1: *The fraction of the CLM4.5 grid cells covered by (a) bare soil, (b) forest, (c) shrubland, and (d) open land.*

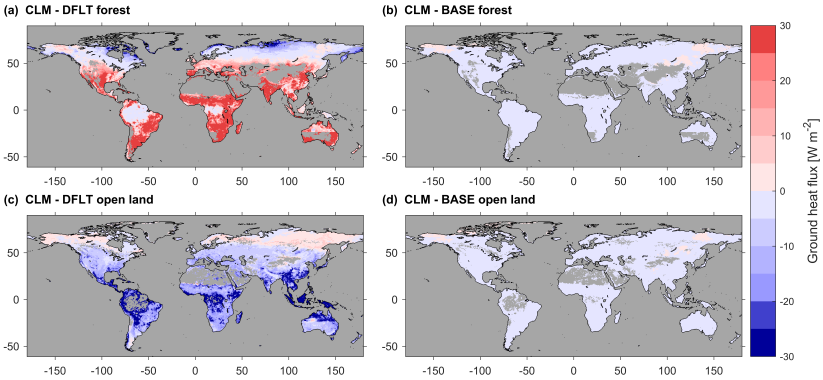


Figure A.2: *Ground heat flux for forests (a and b) and open land (c and d) in CLM - DFLT (a and c) and CLM - BASE (b and d). Positive values correspond to a heat flux out of the soil.*

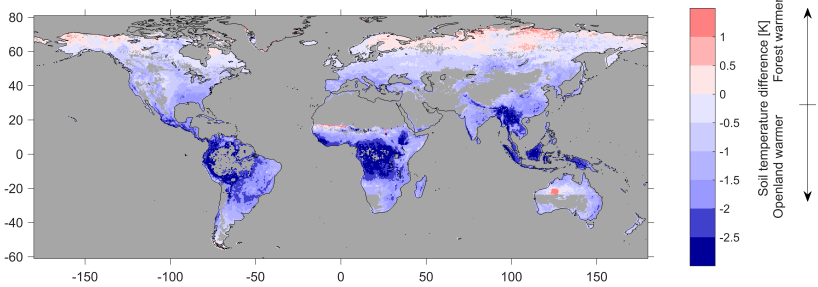


Figure A.3: *Difference in vertically-averaged annual mean soil temperature of forest minus open land in CLM - BASE.*

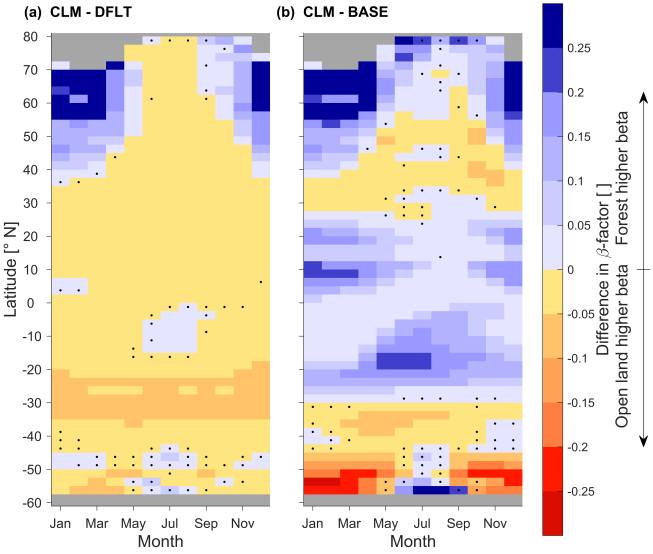


Figure A.4: *Seasonal and latitudinal variations of β_t -factor differences of forest minus open land in (a) CLM - DFLT and (b) CLM - BASE. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot.*

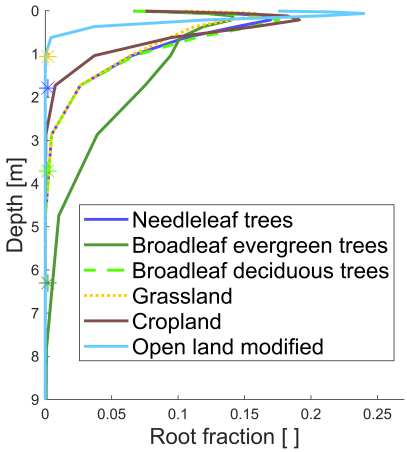


Figure A.5: *Vertical root fraction distribution of the different PFTs in the default version of CLM4.5 and in light blue the modified root fraction distribution of open land PFTs used in CLM - PLUS. The asterisks mark the reported maximum rooting depths of Fan et al. (2017) for annual grass (yellow), evergreen needleleaf trees (dark blue), deciduous broadleaf trees (ligth green), and evergreen broadleaf trees (dark green).*

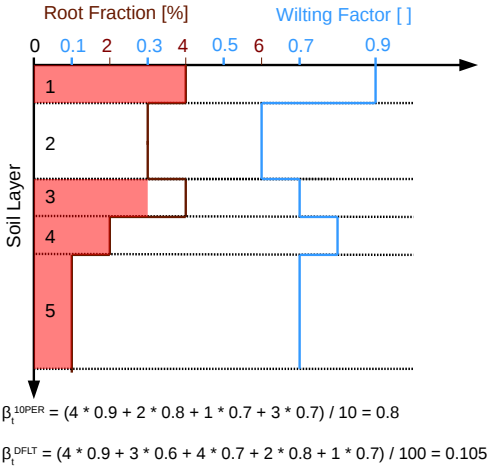


Figure A.6: ***Example of the calculation of the β_t-factor with the 10 % modification.*** *Shown are five soil layers with the fraction of the roots in these layers in brown and the wilting factor in blue. On the bottom the calculation of β_t for this particular example with the 10 % modification (β_t^{10PER}) and the default calculation in CLM4.5 (β_t^{DFLT}), assuming the roots not shown have a wilting factor of zero. The root fractions eventually used to calculate β_t^{10PER} are shaded in red.*

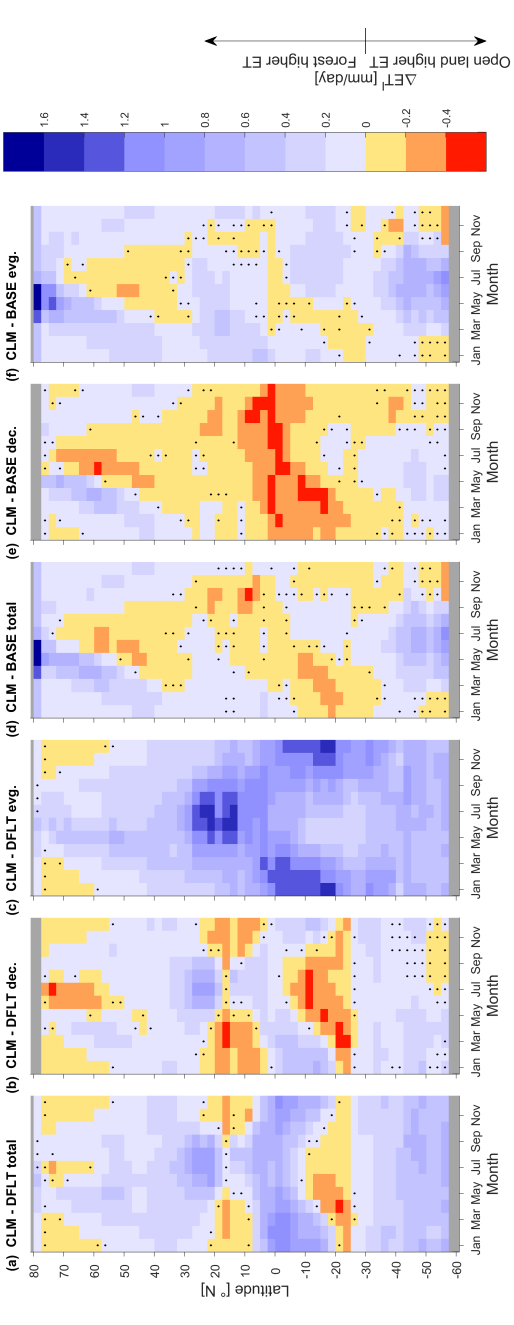


Figure A.7: *Seasonal and latitudinal variations of ΔET^{l} in CLM - DFLT for (a) all tree PFTs minus open land, (b) deciduous tree PFTs only minus open land, (c) evergreen tree PFTs only minus open land. Points with a mean which is insignificantly different from zero in a two-sided t-test at 95 % confidence level are marked with a black dot. All data from the 2002-2010 analysis period corresponding to a given latitude and a given month are pooled to derive the sample set for the test. The same for CLM - BASE in panels (d), (e), and (f).*

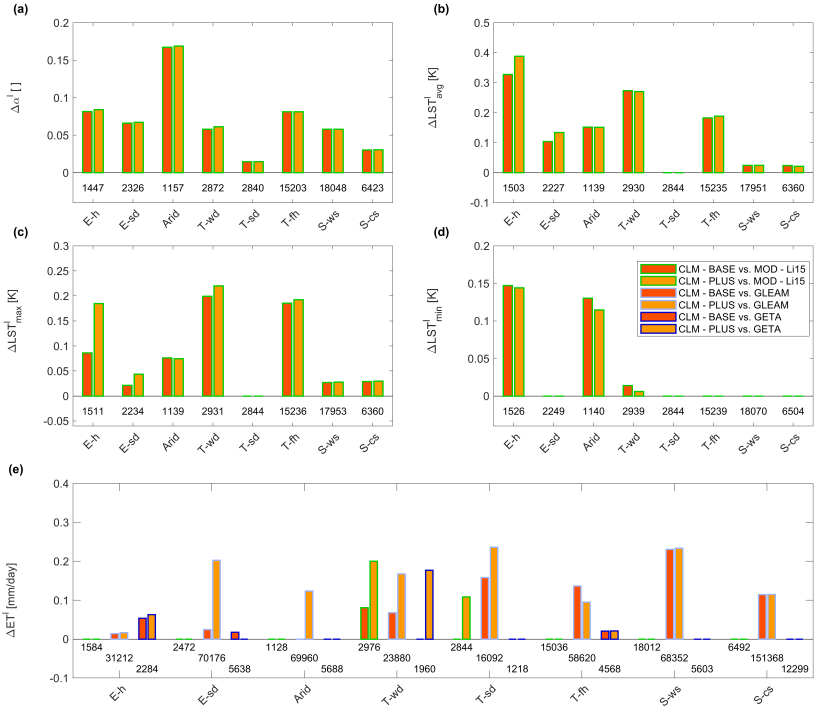


Figure A.8: *IA (Duveiller et al., 2016) of CLM - BASE (red), and CLM - PLUS (orange) with forestation observations over Köppen-Geiger climate zones (Kottek et al., 2006, ; Fig. 2.1) for monthly (a) $\Delta\alpha^l$, (b) ΔLST^l_{avg}, (c) ΔLST^l_{max}, and (d) ΔLST^l_{min}. Panel (e) shows the IA over the Köppen-Geiger climate zone for ΔET^l of CLM - BASE (red), and CLM - PLUS (orange) with forestation (green edge), GLEAM (light blue edge), GETA (dark blue edge). The numbers indicate the size of the data samples used for the calculation of IA.*

Table A.1: *The PFT-specific values of V_{cmax} [$\mu mol\, m^{-2}\, s^{-1}$], r_a, and r_b [ ] in default of CLM4.5 and in CLM - PLUS.*

PFT name	Default			CLM - PLUS		
	r_a	r_b	V_{cmax}	r_a	r_b	V_{cmax}
NET - temperate	7.0	2.0	62.5	default		50
NET - boreal	7.0	2.0	62.6	default		default
NDT - boreal	7.0	2.0	39.1	default		default
EBT - tropical	7.0	1.0	55.0	default		35
EBT - temperate	7.0	1.0	61.5	default		50
DBT - tropical	6.0	2.0	41.0	default		65
DBT - temperate	6.0	2.0	57.7	default		default
DBT - boreal	6.0	2.0	57.7	default		70
C_3 arctic grass	11.0	2.0	78.2	11.0	11.0	90
C_3 grass	11.0	2.0	78.2	11.0	11.0	60
C_4 grass	11.0	2.0	51.6	11.0	11.0	default
Crop	6.0	3.0	100.7	11.0	11.0	90

Table A.2: *The default PFT classification in CLM4.5.*

No.	Abbr.	Full name	Class
1	Bare soil		–
2	NET - temperate	Temperate evergreen needleleaf tree	Forest
3	NET - boreal	Boreal evergreen needleleaf tree	Forest
4	NDT - boreal	Boreal deciduous needleleaf tree	Forest
5	BET - tropical	Tropical evergreen broadleaf tree	Forest
6	BET - temperate	Temperate evergreen broadleaf tree	Forest
7	BDT - tropical	Tropical deciduous broadleaf tree	Forest
8	BDT - temperate	Temperate deciduous broadleaf tree	Forest
9	BDT - boreal	Boreal deciduous broadleaf tree	Forest
10	BES - temperate	Temperate evergreen broadleaf shrub	–
11	BDS - temperate	Temperate deciduous broadleaf shrub	–
12	BDS - boreal	Boreal deciduous broadleaf shrub	–
13	C_3 arctic grass		Open land
14	C_3 grass		Open land
15	C_4 grass		Open land
16	Crop	Unmanaged rainfed C_3 crop	Open land

Table A.3: *The LC types from Ambrose and Sterling (2014) (GETA) used in this study and the numbers of the respective PFTs in CLM4.5 applied to the different LC types (Table A.2).*

Abbr. GETA	Full name GETA	PFTs of CLM4.5
ENF	Evergreen needleleaf forest	2, 3
DNF	Deciduous needleleaf forest	4
EBF	Evergreen broadleaf forest	5, 6
DBF	Deciduous broadleaf forest	7, 8, 9
GRS	Grassland	13, 14, 15
CRN	Non-irrigated cropland	16

Table A.4: *Overview of the different modifications of CLM4.5 incorporated in the simulations presented this study.*

Run	SeSCs	Shallow roots	10 %	Light limitation	V_{cmax}
CLM - DFLT	–	–	–	–	–
CLM - BASE	✓	–	–	–	–
CLM - ROOT	✓	✓	–	–	–
CLM - 10PER	✓	✓	✓	–	–
CLM - LIGHT	✓	✓	✓	✓	–
CLM - PLUS	✓	✓	✓	✓	✓

Table A.5: *Area-weighted annual mean ET for each PFT analysed in this study according to the GETA data and in the different configurations of CLM4.5 and fraction of the land surface covered by the different PFTs. On the bottom is listed the global integral of annual ET.*

Abbr.	Full name	Frac. [%]	ET [mm/day]					
			GETA	BASE	ROOT	10PER	LIGHT	PLUS
NET - temperate	Needleleaf evergreen tree - temperate	3.2	1.74	1.78	1.78	1.81	1.84	1.75
NET - boreal	Needleleaf evergreen tree - boreal	6.9	1.00	0.97	0.97	0.98	1.00	1.00
NDT - boreal	Needleleaf deciduous tree - boreal	1.0	0.72	0.72	0.72	0.72	0.73	0.73
EBT - tropical	Broadleaf evergreen tree - tropical	9.5	3.47	3.70	3.70	3.78	3.87	3.52
EBT - temperate	Broadleaf evergreen tree - temperate	1.5	2.58	2.61	2.61	2.66	2.70	2.60
DBT - tropical	Broadleaf deciduous tree - tropical	8.0	2.65	2.31	2.31	2.42	2.44	2.62
DBT - temperate	Broadleaf deciduous tree - temperate	3.1	1.78	1.74	1.74	1.76	1.79	1.79
DBT - boreal	Broadleaf deciduous tree - boreal	1.3	1.23	1.08	1.08	1.08	1.10	1.13
C_3 arctic grass		3.1	0.81	0.66	0.65	0.65	0.66	0.67
C_3 grass		8.8	1.48	1.60	1.53	1.56	1.57	1.53
C_4 grass		8.0	2.06	2.32	2.18	2.22	2.12	2.04
Crop	C_3 unmanaged rainfed crop	10	1.90	1.76	1.70	1.73	1.74	1.73
Total ET [$km^3\ yr^{-1}$]				70223	69059	70322	70649	69023

Table A.6: *ET and latent heat flux in-situ observations from various studies and the values of the different CLM4.5 sensitivity tests at the respective locations.*

Study	Region	PFTs	Unit	Season	Obs.	BASE	ROOT	10PER	LIGHT	PLUS
da Rocha et al. (2004)	Amazon	EBT	mm/day	Dry	3.96	3.49	3.49	3.90	4.06	3.48
				Wet	3.18	3.57	3.57	3.57	3.64	3.37
				All	3.51	3.54	3.54	3.68	3.79	3.40
von Randow et al. (2004)	Amazon	EBT	$W\,m^{-2}$	Dry	108.6	82.9	82.9	100.6	105.3	90.8
				Wet	104.5	113.9	113.9	113.8	116.2	108.9
		Grass		Dry	63.9	81.2	56.0	60.2	62.7	64.7
				Wet	83.0	113.9	113.9	113.9	106.1	100.1
Liu et al. (2005)	Alaska	Grass	$W\,m^{-2}$	All	16.1	16.4	16.8	16.8	16.8	16.8
		DBT		All	22.5	13.7	13.7	14.0	14.0	14.1
		ENF		All	23.9	18.0	18.0	18.4	18.4	18.4

B

Appendix to Chapter 3

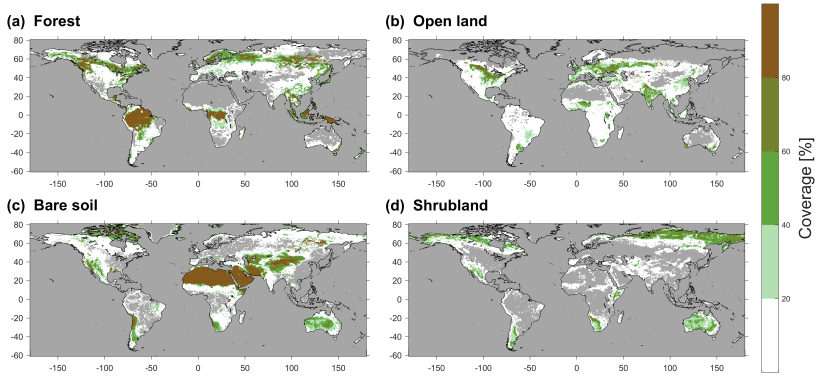


Figure B.1: *The fraction of the CLM5.0 grid cells covered by (a) forest, (b) open land, (c) bare soil, and (d) shrubland.*

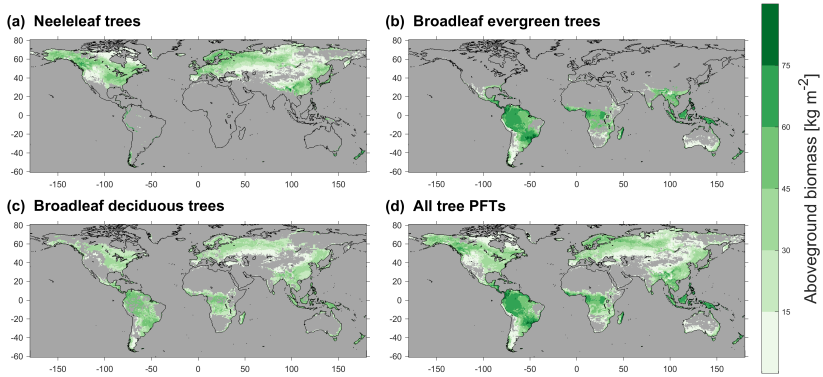


Figure B.2: *Aboveground biomass averaged over (a) all needleleaf plant functional types (PFTs), (b) all broadleaf evergreen PFTs, (c) all broadleaf deciduous PFTs, and (d) all tree PFTs.*

Table B.1: *Comparison of AboveGround Biomass (AGB) [*$\mathrm{kg\,m^{-3}}$*] and diurnal range of the Biomass Heat Flux (BHF) [*$\mathrm{W\,m^{-2}}$*] in in-situ observations and the CLM - BHS simulation. The months during which measurements were made go from January (1) to December (12). Values marked with a asterisk were estimated by reading from graphs displaying the diurnal cycle of biomass heat fluxes. For some studies the respective plant functional type was not available at the location of the observational study. Therefore, the values of* [1] *temperate broadleaf evergreen forest,* [2] *tropical broadleaf evergreen forest,* [3] *tropical broadleaf evergreen forest, and* [4] *temperate needleleaf evergreen forest at 38.25° N/122.25° W were used instead when comparing to the observational studies of Haverd et al. (2007), Silberstein et al. (2001), Meesters and Vugts (1996), and Garai et al. (2010), respectively.*

Study	Lat.	Lon.	Months	Forest type	Observations		Model	
					AGB	Range BHF	AGB	Range BHF
Kilinc et al. (2012)	37.25° S	145.25° E	1, 2, 12 3, 4, 5 6, 7, 8 9, 10, 11	Broadleaf evergreen	112.5	43.0* 28.0* 22.0* 35.0*	26.2	55.7 47.6 38.6 49.1
Haverd et al. (2007)[1]	35.75° S	148.25° E	3	Broadleaf deciduous	66.0	105*	32.9	68.3
Silberstein et al. (2001)[2]	32.75° S	116.25° E	3 10	Broadleaf deciduous	37.8	125* 115*	33.4	66.6 56.5
Meesters and Vugts (1996)[1]	17.75° S	177.25° E	8	Needle (plant.)	18.4	54.0*	76.3	63.8
Moore and Fisch (1986)	2.75° S	59.75° W	5, 6, 7, 8	Broadleaf evergreen	60.3	45.0*	58.8	76.2
Michiles and Gielow (2008)	2.75° S	60.25° W	4 8	Broadleaf evergreen	66.8	45.0* 53.0*	60.0	62.3 79.1
Garai et al. (2010)[3]	38.25° N	121.75° W	6	Broadleaf deciduous	7.2	17.8*	9.2	33.1
Oliphant et al. (2004)	39.25° N	86.25° W	all	Broadleaf deciduous	19.5	24.8*	35.5	46.2
Burns et al. (2015)	40.25° N	105.75° W	5, 6, 7, 8, 9	Needleleaf evergreen	?	120	4.0	25.0
McCaughey and Saxton (1988)	45.75° N	77.25° W	5, 6, 7, 8	Mixed forest	20.9	17.5*	32.3	53.0
Lindroth et al. (2010)	60.25° N	17.25° E	6, 7	Needleleaf evergreen	43.4	60.0*	53.5	55.8

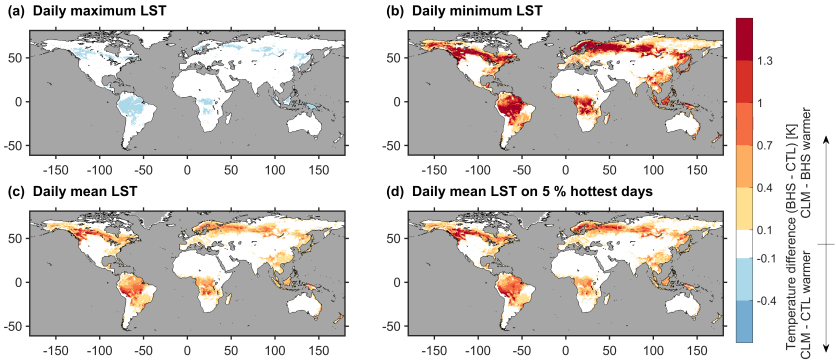


Figure B.3: *As Fig. 3.3 but LST during boreal summer (JJA).*

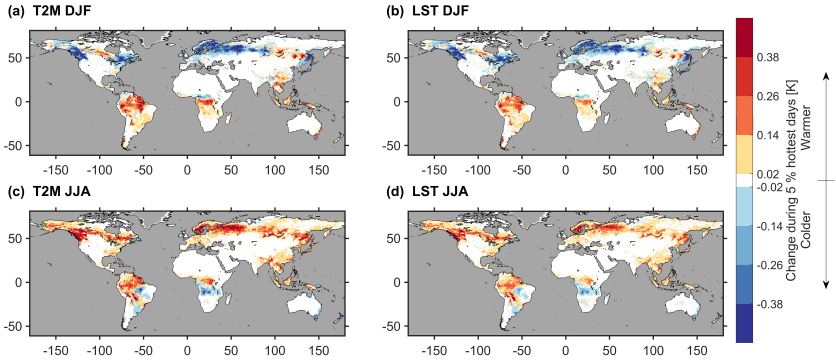


Figure B.4: *Temperature change due to BHS averaged over the days when the atmospheric temperature exceeded its 95 % percentile in comparison to the mean temperature change over all days for (a) 2 m air temperature (T2M) during boreal winter (DJF; Fig. B.5 d - c), (b) LST during DJF (Fig. B.6 d - c), (c) T2M during boreal summer (JJA; Fig. 3.3 d - c), and (d) LST during JJA (Fig. B.3 d - c).*

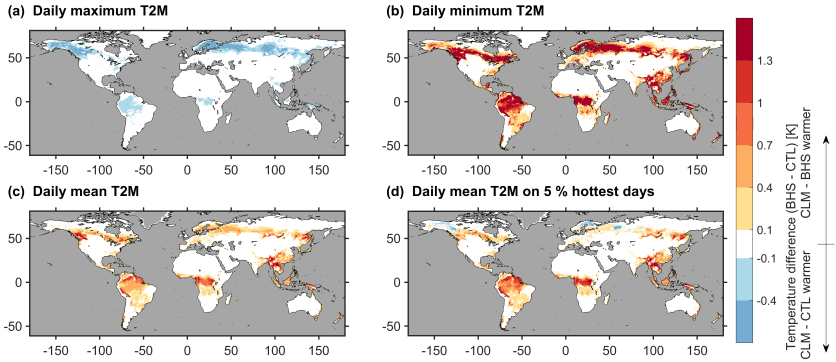


Figure B.5: *As Fig. 3.3 but for boreal winter (DJF).*

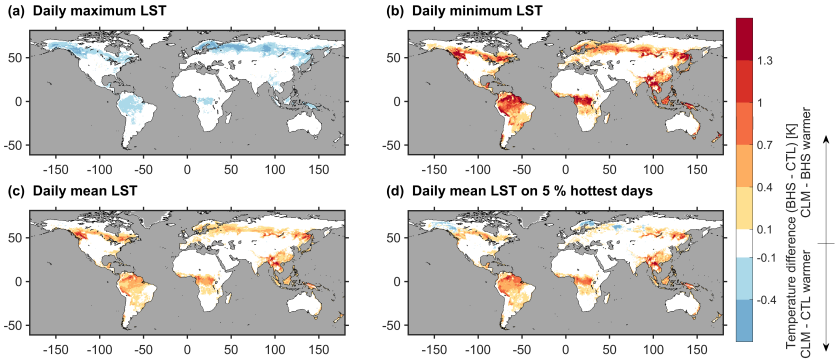


Figure B.6: *As Fig. 3.3 but for LST during boreal winter (DJF).*

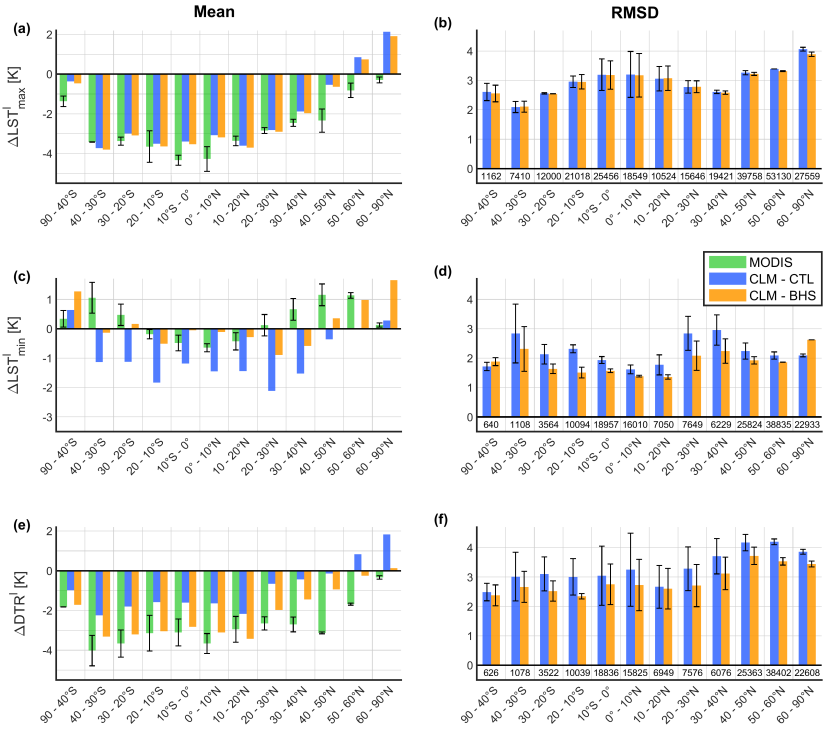


Figure B.7: *Area-weighted annual mean over different latitudinal bands of the average of the two MODIS products (green; Li et al., 2015; Duveiller et al., 2018b), CLM - CTL (blue), and CLM - BHS (orange) in (a) ΔLST^{l}_{max}, (c) ΔLST^{l}_{min}, and (e) ΔDTR^{l}. The black whiskers depict the range between the two observational data sets. Panels (b), (d), and (f) display the area-weighted spatiotemporal root-mean-squared deviation (RMSD) of CLM - CTL (blue), and CLM - BHS (orange) against the two MODIS data sets for (b) ΔLST^{l}_{max}, (d) ΔLST^{l}_{min}, and (f) ΔDTR^{l}, respectively. The numbers below the bar indicate the size of the sample used to calculate the RMSDs and the whiskers the range between the RMSD against MOD - Li15 and the RMSD against MOD - Du18.*

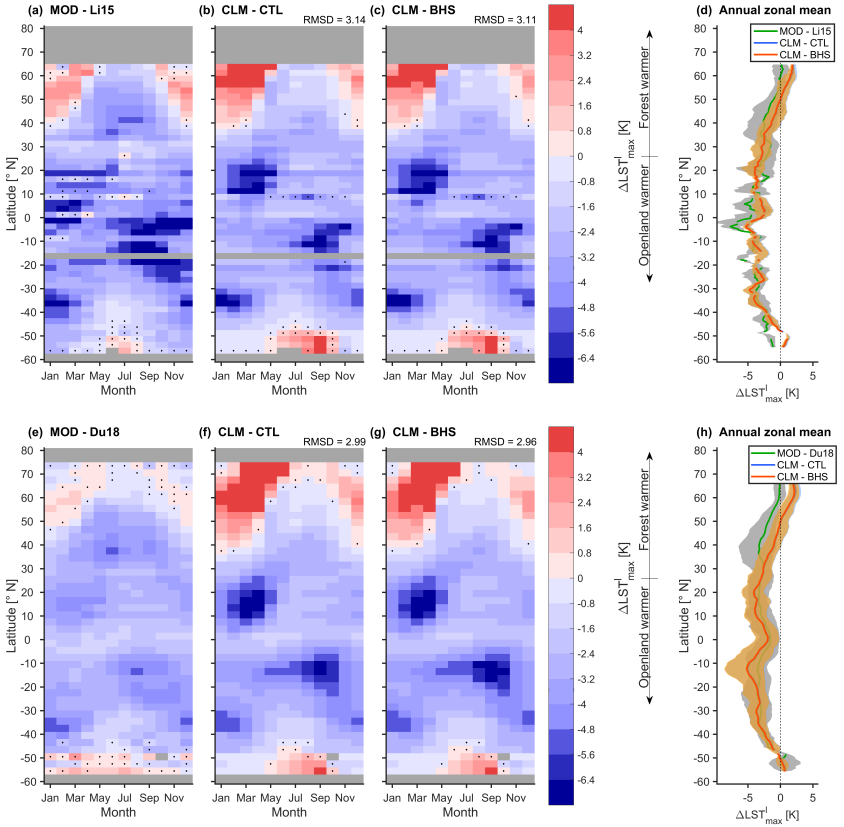


Figure B.8: *As Fig. 3.6 but for* ΔLST^{l}_{max}.

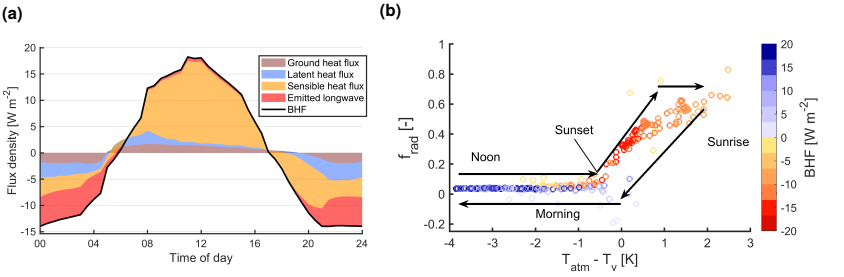


Figure B.9: *As Fig. 3.8 but for boreal summer (JJA) at 60.25° N/17.25° E (location of Lindroth et al., 2010). Panel (b) displays the fraction the BHF compensated by emitted longwave radiation based on the monthly-averaged diurnal cycle from May to September.*

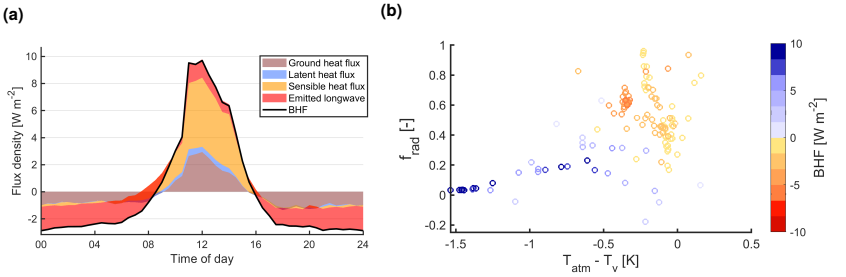


Figure B.10: *As Fig. B.9 bur for boreal winter (DJF) in both panels.*

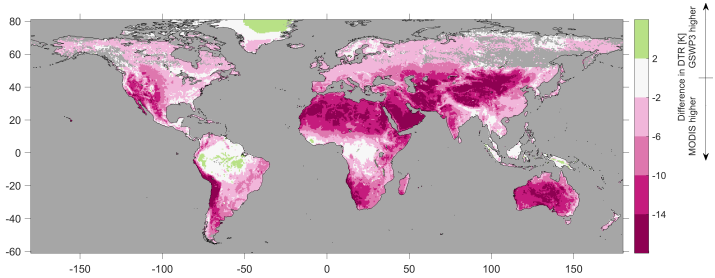


Figure B.11: *Difference of the Diurnal Temperature Range (DTR) in GSWP3 data set (used to force CLM5.0) minus the DTR of the MODIS LST. Forcing data where extracted at 1:30 am/pm solar time for comparison with MODIS.*

C

Appendix to Chapter 4

C.1 Sensitivity of site pair analysis to selection criteria

Here, we present sensitivity tests of the site pair analysis to the selection criteria of the site pairs, which are summarized in Table C.1. Overall, the site pair analysis is affected by a trade-off between the necessity of obtaining large enough sample size to reduce the influence of noise and excluding confounding factors by applying stricter selection criteria (Fig. C.1 a). With an increasing number of site pairs the agreement between the full GAM and the rain gauges measurements increases. Synchronously however, the fraction of ΔP^1 that can be explained by other factors than LC increases, resulting in an increasing IA between the 'no LC' GAM and the rain gauges. The IA of the portion of ΔP^1 that is explained by the differences in AL_r and forest fraction in the GAM exhibits an optimum at approximately 1000 to 2000 sites.

The magnitude of the median ΔP^1 increases with an increasing minimum threshold for the differences in the fractions of AL_r and forest (Fig. C.1 b). However, this increasing tendency weakens towards higher minimum thresholds, implying that ΔP^1 is nonlinear and saturates for higher rates of forestation. ΔP^1 tends to strengthen for higher thresholds also in Region 1, with the exception of the highest threshold, for which the robustness of the signal disintegrates due to the small sample size (Fig. C.1 d). ΔP^1 also exhibits a weak increasing tendency when relaxing other selection criteria than the fractions of AL_r and forest (Fig. C.1 c). This likely originates from the increasing influence of confounding factors for the median ΔP^1, as indicated by the higher IAs for the 'no LC' GAM for larger numbers of site pairs. More distinctively, the spread among the site pairs increases, implying that relaxed selection criteria result in more noise due to other variables than the AL_r and forest fractions.

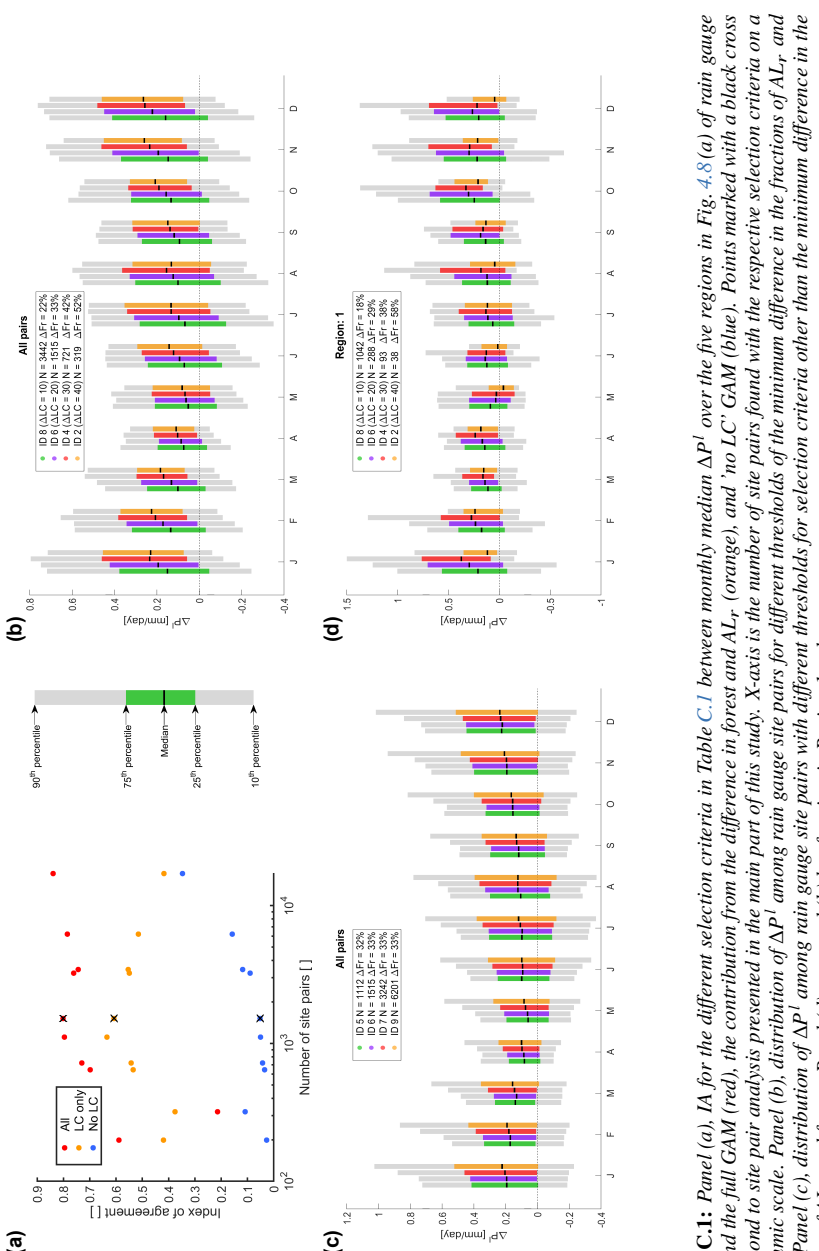


Figure C.1: *Panel (a), IA for the different selection criteria in Table C.1 between monthly median ΔP^1 over the five regions in Fig. 4.8 (a) of rain gauge pairs and the full GAM (red), the contribution from the difference in forest and AL_r (orange), and 'no LC' GAM (blue). Points marked with a black cross correspond to site pair analysis presented in the main part of this study. X-axis is the number of site pairs found with the respective selection criteria on a logarithmic scale. Panel (b), distribution of ΔP^1 among rain gauge site pairs for different thresholds of the minimum difference in the fractions of AL_r and forest. Panel (c), distribution of ΔP^1 among rain gauge site pairs with different thresholds for selection criteria other than the minimum difference in the fractions of AL_r and forest. Panel (d), as panel (b) but for sites in Region 1 only.*

Table C.1: ***Sensitivity tests for selection criteria.*** *From left to right, identification number of sensitivity test, number of site pairs found in respective test, minimum thresholds for difference in fraction of AL_r and forest, minimum threshold for sum of the AL_r and forest fractions, as well as maximum thresholds for distance between sites, altitude difference, difference in open water fraction, difference in upwind open water fraction, difference in slope, difference in TRI, and difference in annual mean upwind distance to coast. Last column, median annual mean ΔP^l across all site pairs and interquartile range of annual mean ΔP^l in brackets.*

ID	N	ΔLC [%]	Sum [%]	Dist. [°]	Alt. [m]	pct_ow [%]	uw_pct_ow [%]	Slope [°]	TRI [m]	uw_cd [°]	Median ΔP^l [mm/day]
1	199	30	50	0.50	15	15	15	4	2	0.20	0.14 (0.02-0.27)
2	319	40	50	0.75	25	20	20	5	3	0.25	0.17 (0.07-0.31)
3	645	25	50	0.60	20	15	15	5	3	0.25	0.15 (0.04-0.29)
4	721	30	50	0.75	25	20	20	5	3	0.25	0.17 (0.06-0.31)
5	1112	20	50	0.75	25	20	20	5	3	0.15	0.14 (0.02-0.27)
6*	1515	20	50	0.75	25	20	20	5	3	0.25	0.14 (0.02-0.28)
7	3242	20	40	0.75	40	20	20	8	6	0.40	0.15 (0.01-0.30)
8	3442	10	50	0.75	25	20	20	5	3	0.25	0.11 (-0.02-0.25)
9	6201	20	40	1.00	50	20	20	10	6	0.50	0.15 (0.00-0.33)
10	17053	15	40	1.00	100	20	20	12	7	0.60	0.16 (-0.01-0.38)

*Presented in main part of study.

C.2 Uncertainty due to choice in GAM structure

The results from the GAMs, in particular ΔP^{rem}, are sensitive to the structure of the GAM. Here we present three additional GAMs to illustrate this sensitivity, besides the GAMs presented in the main part of this study (Table C.2): (1) A GAM with a more flexible pure spatial term. (2) A GAM were we removed the interaction terms of the LC-related variables with latitude and longitude, which were sometimes insignificant according to the ANOVA analysis. (3) A GAM, for which the wind-related predictor variables were calculated based on trajectories that were started 200 and 300 hPa above the surface instead of starting at the two pressure levels that produced most precipitation (but still for the same precipitation events; Fig. C.16). The 'no LC' GAM was also added in Table C.2 for completeness, but is not discussed here further. To quantify the spatial autocorrelation of the residuals we compute the average Moran's I across the months for distances below 1.5° ($MI_{<1.5}$) and for distances from 1.5° to 3° ($MI_{1.5-3}$), weighting points by one divided by distance from the central point squared. For a comparison, the MSWEP precipitation fields exhibit a $MI_{<1.5}$ and $MI_{1.5-3}$ of 0.758 and 0.326, respectively.

When increasing the flexibility of the pure spatial term the resulting GAM2 represents the precipitation climatology of MSWEP better than GAM1 and reduces the spatial autocorrelation of the residuals. However, the agreement of GAM2 with the site pairs is lower than the one of GAM1, as part of $\Delta P^l_{20\,\%}$ becomes incorporated in the more flexible spatial term. Looking at GAM3, it appears that its performance is weaker than GAM1 in all aspects considered. In particular, the spatial interaction terms of the LC-related variables appear relevant for the conformity of the GAMs with the site pair analysis. Nevertheless, the overall features of $\Delta P^l_{20\,\%}$ are similar for the individual GAMs, indicating that this part of the precipitation signal from forestation is robust (Fig. C.2). Larger discrepancies emerge for $\Delta P^{rem}_{20\,\%}$, in particular regarding the strength of the signal in summer over the central parts of Europe (Fig. C.3). Notably, the selection of the starting heights when

calculating the wind trajectories has a distinct effect on the estimate of $\Delta P^{rem}_{20\,\%}$ (compare GAM1 and GAM4). GAM4 even represents the precipitation climatology of MSWEP slightly better than GAM1, even though the underlying wind trajectories are presumably less tightly linked to the precipitation events for GAM4. Based on this argument, we have decided to utilize the trajectories for which the starting heights were based on the production of precipitation to generate the results presented in the main part of this study. All things considered, the sensitivity of ΔP^{rem} on the starting heights of the wind trajectories certainly deserves further attention in future studies.

Table C.2: ***Sensitivity tests for GAM structure.*** *First column, ID used for respective GAM followed by changes made in comparison to GAM1, which is presented in main part of this study (see Section 4.2.3). Columns three to seven, adjusted R^2, RMSE, and IA of precipitation fields simulated by GAMs in comparison to MSWEP as well as average $MI_{<1.5}$ and $MI_{1.5-3}$ of the residuals. Last column, IA of monthly median $\Delta P^{l}_{20\,\%}$ for the five regions in Fig. 4.8 (a) between rain gauge site pairs and GAMs.*

ID	Changes	R^2	RMSE	IA	$MI_{<1.5}$	$MI_{1.5-3}$	IA_{PS}
GAM1	- (GAM presented in main part)	0.933	0.298	0.964	0.430	-0.0284	0.802
GAM2	k=20, sp=3 for pure spatial term instead of 15 and 5	0.942	0.279	0.969	0.407	-0.0170	0.769
GAM3	no 'prox_LC_i, lat, lon' and 'uw_pct_LC_i, lat, lon' terms	0.930	0.303	0.963	0.434	-0.0288	0.749
GAM4	Trajectories starting from 200 and 300 hPa above surface	0.936	0.293	0.966	0.417	-0.0249	- (0.80)*
no LC	Terms including prox_LC_i removed	0.921	0.319	0.957	0.448	-0.0286	0.0910

* Not directly comparable as different trajectories affects uw_pct_opwa and uw_cd criteria in site pair selection.

C.3 Sensitivity of downwind precipitation changes to location of forestation

As for the local effect, the location of forestation is important for downwind alterations of precipitation. Here, we briefly discuss ΔP^{rem} from two additional forestation scenarios besides the European-wide 20 % forestation: (1) foresting 20 % of the land in areas with an annual mean upwind coast distance (Fig. 4.5 d) less than 5° and (2) foresting 20 % of the land in areas with an annual mean upwind coast distance exceeding 5° (corresponding to 557 km). This is done by recalculating the upwind LC fractions as described in 'Estimating the theoretical effect of forestation on precipitation' in the Methods, if forestation only occurs in the respective regions of interest.

Forestation over areas with an average upwind coast distance of less than 5° explains most of the estimated ΔP^{rem} in winter, while forestation over areas further away from the coast has only a minor downwind effect on precipitation (Fig. C.4 a, d, g, and j). In summer on the other hand, forestation over the more continental regions of Europe increases precipitation over those areas, indicating that the high ET of forests can be an important source of moisture in continental climates (Fig. C.4 b, e, h, and k). This is consistent with the finding that atmospheric moisture recycling in the northern mid-latitudes is unimportant during winter but substantial during summer (Van der Ent and Savenije, 2011, ; also see Supplement C.5). It therefore appears that the downwind reduction in winter precipitation due to forestation can be avoided by prohibiting forestation in areas close to the coast, while still profiting from the downwind increase in precipitation during summer. It needs to be mentioned however that forestation over the continental parts of Europe might affect winter precipitation in regions that lie outside of our study domain.

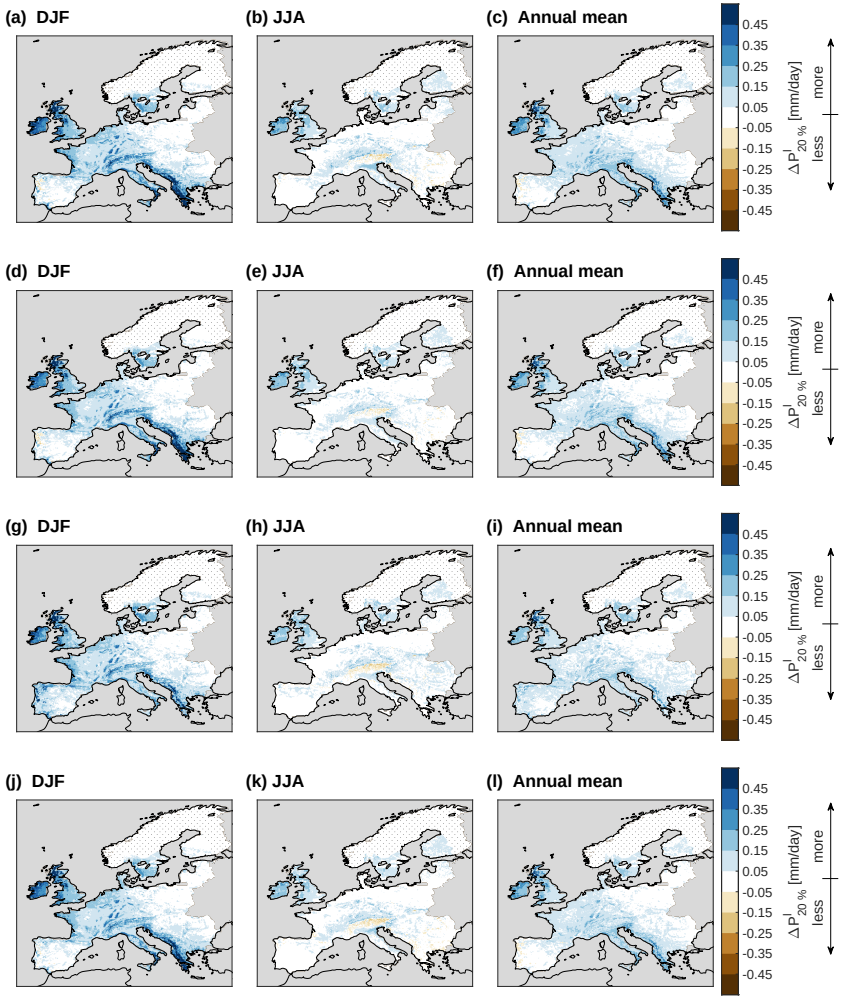


Figure C.2: *Estimated $\Delta P^{l}_{20\,\%}$ from foresting 20 % of the land surface according to GAM1 (a-c; Fig. 4.9 (a)-() repeated for convenience), GAM2 (d-f), GAM3 (g-j), and GAM4 (j-l) in Table C.2. Left column boreal winter, middle column boreal summer, and right column annual mean.*

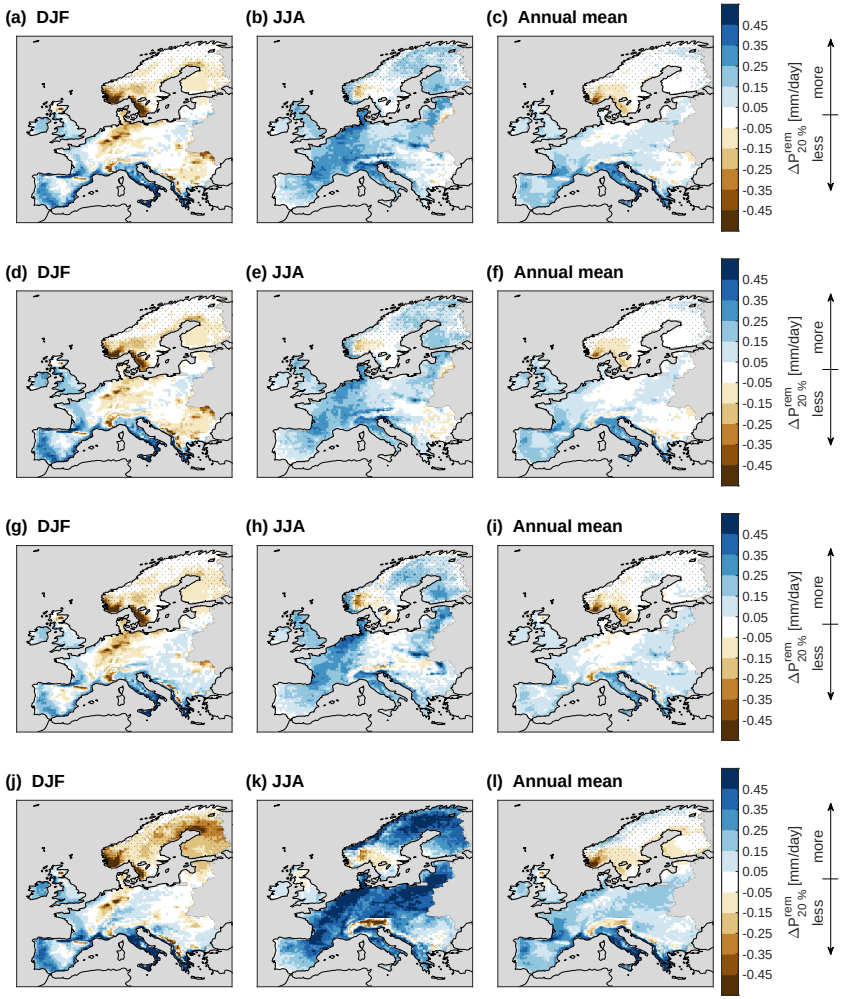


Figure C.3: *As Fig. C.2 but for* $\Delta P^{rem}_{20\,\%}$.

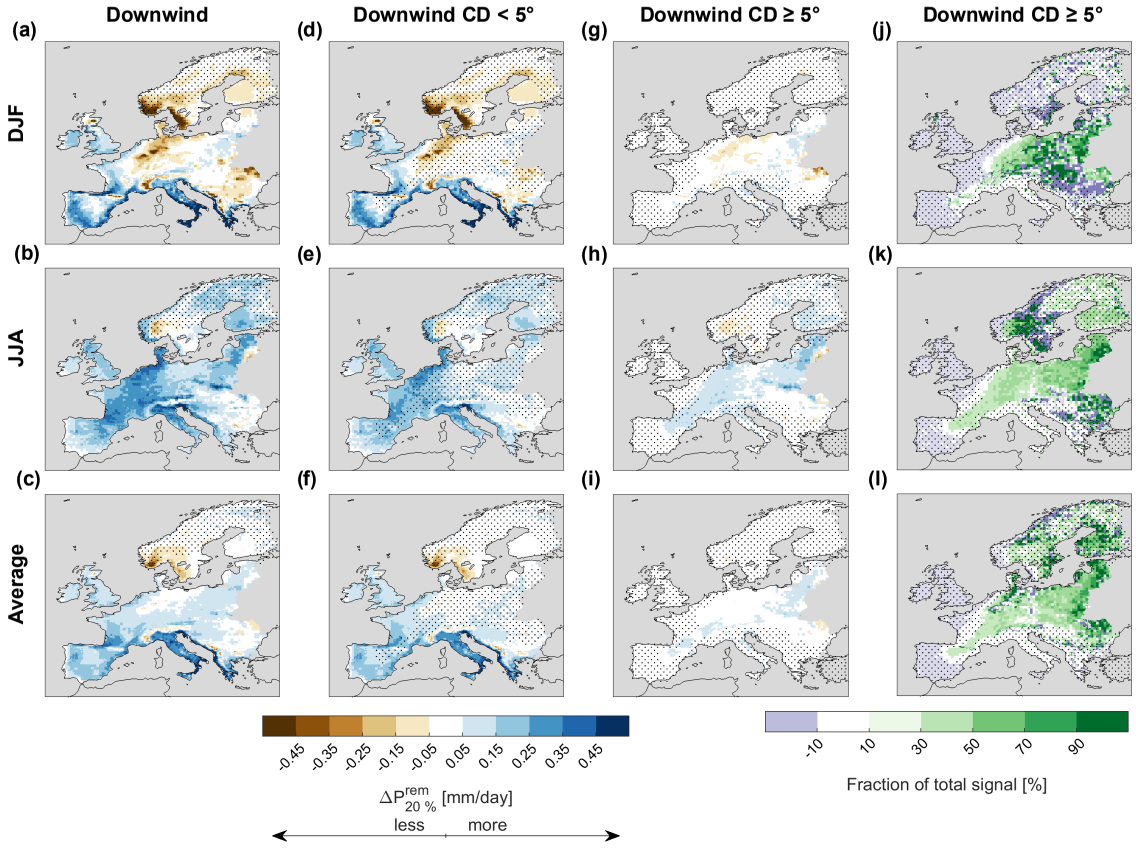


Figure C.4: *Estimated downwind precipitation change (a-c) after foresting 20 % of the land surface in the entire study domain (Fig. 4.9 (d)-(f) repeated for convenience). To the right, downwind precipitation change from foresting 20 % of the land surface over areas with an annual mean upwind coast distance smaller than 5°(d-f) and larger than 5°(g-i). Stippling depicts areas that were not forested in the respective scenario. On the very right, fraction of panels (a)-(c) explained by panels (g)-(i). Top row boreal winter (DJF), middle row boreal summer (JJA), and bottom row annual mean.*

C.4 The role of surface roughness

It was hypothesized in the main part of this study that the higher surface roughness of forests compared to AL_r is an important physical driver of the estimated precipitation changes from forestation through altering the propagation speed of precipitating air masses. Here, we analyze the connection of the monthly ERA5-Land (C3S, 2019) wind speed 10 m above the surface (WS) and the propagation speed of the wind trajectories (PS), calculated from the upwind coast distance divided by the upwind coast time (see 4.2.1), to forestation and precipitation. We use these two variables as a proxy for the local surface roughness in the case of WS and the upwind surface roughness prior to precipitation events in the case of PS. We present results from three additional GAMs in this section: Two GAMs of the same structure as the GAM of the main part (Table 4.3), but with WS and PS as the response variables instead of the MSWEP precipitation climatology. The third GAM is composed of a number of 's' terms for various physical variables that might affect precipitation, including WS and PS (Table C.3), and is fitted to the MSWEP precipitation climatology as described in Section 4.2.3. With the latter GAM, we estimate the change in precipitation associated with an 1 m/s increase in WS by subtracting a GAM prediction where WS was decreased by 0.5 m/s from a GAM prediction where WS was increased by 0.5 m/s. Using the same procedure, we estimate the change in precipitation following a 4 m/s increase in PS.

During the winter months, the latter GAM identifies a pronounced link between precipitation and wind speed (Fig. C.5 a and d). Higher surface wind speeds are associated with lower precipitation, implying that rougher surfaces receive more precipitation. According to the GAM that was fitted to WS, forestation results in a decrease of WS during this season in most regions of the study domain, in particular in Regions 1 and 2, where $\Delta P^{l}_{20\,\%}$ was strongest (Fig. C.6 a). Further, a faster propagation of the air masses from the coast to the precipitation event is associated with higher amount of precipitation during winter (Fig. C.5 d). Forestation appears to hinder the propagation of precipitation events towards the continental regions of Europe, as indicated by the estimation reduction in downwind PS due to forestation (Fig. C.6 d-f). This conforms with the positive $\Delta P^{rem}_{20\,\%}$ along the southern and western coastal regions of Europe and the neutral to negative $\Delta P^{rem}_{20\,\%}$ toward the East (Fig.4.9). During summer on the other hand, the relationship between precipitation and WS is weaker, in agreement with the less positive $\Delta P^{l}_{20\,\%}$ than during winter. Also, we find an only minor connection of PS and precipitation, implying that different processes such as evapotranspiration are more relevant for the estimated $\Delta P^{rem}_{20\,\%}$ during this season.

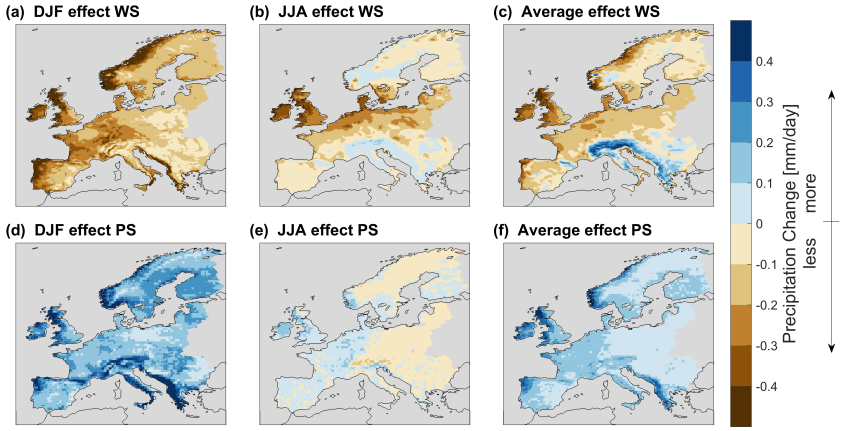


Figure C.5: *Panels (a)-(c), change in precipitation associated with a WS increase of 1 m/s. Panels (d)-(f), change in precipitation associated with a PS increase of 4 m/s. Left column boreal winter, middle column boreal summer, and right column annual mean.*

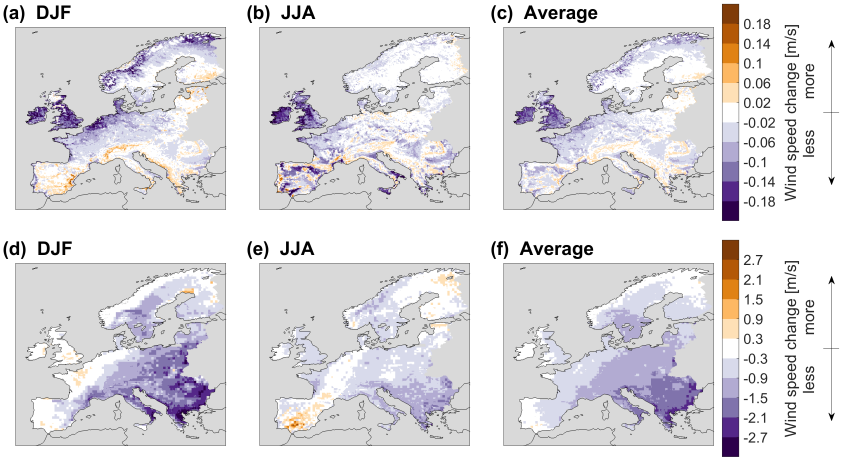


Figure C.6: *Panels (a)-(c), estimated local change in surface WS from foresting 20 % of the land surface in the entire study domain. Panels (d)-(f), estimated downwind change in PS from the same forestation scenario.*

Table C.3: *Terms used to construct the GAM with potential physical drivers of precipitation. Consult Section 4.2.3 for more information on meaning of terms and construction of GAM.*

Variable(s)	Type	k	sp	Data source
alt	s	10	estimated	EU-DEM v1.1 (Section 4.2.1)
expo	s	10	estimated	EU-DEM v1.1 (Section 4.2.1)
slope	s	10	estimated	EU-DEM v1.1 (Section 4.2.1)
TPI	s	10	estimated	EU-DEM v1.1 (Section 4.2.1)
TRI	s	10	estimated	EU-DEM v1.1 (Section 4.2.1)
dw_hd	s	10	estimated	ERA5 (Section 4.2.1)
uw_cd	s	10	estimated	ERA5 (Section 4.2.1)
uw_hd	s	10	estimated	ERA5 (Section 4.2.1)
PS	s	10	estimated	ERA5 (Section 4.2.1)
Albedo	s	10	estimated	ERA5-Land
Latent heat flux	s	10	estimated	ERA5-Land
Sensible heat flux	s	10	estimated	ERA5-Land
T2M	s	10	estimated	ERA5-Land
ws	s	10	estimated	ERA5-Land
lat, lon	ti	15	5	-

C.5 Moisture source diagnostic

In this supplement we apply the Lagrangian moisture source diagnostic (Sodemann et al., 2008) to the air parcel trajectories described in Section 4.2.1 in order to inspect the importance of the different LC types as a moisture source for precipitation. This diagnostic attributes moisture uptake along a Lagrangian trajectory to a precipitation event at the end of this trajectory by tracking the specific humidity of the air parcel along its path (for a more detailed description check Sodemann et al., 2008). We deviate from the method as described in Sodemann et al. (2008) on three aspects: (1) We apply the diagnostic at hourly instead of 6-hourly temporal resolution. (2) We remove the threshold for the minimal change in specific humidity for a moisture uptake to be identified i.e, any change in specific humidity is treated. (3) We do not check whether an identified moisture uptake occurred within the planetary boundary layer, as was done in a follow-up study using the same diagnostic (Aemisegger et al., 2014). Ultimately, the diagnostic retrieves the fraction of moisture provided from each point the trajectory overpassed to the final precipitation event. We then divide the fraction of the moisture provided by each point into fractions from the different LCs proportionally to the CLC fractions at this point. As a result, it can be quantified for each trajectory, which fraction of the final precipitation event was supplied by each LC class, by summing the fractions contributed by the individual LC types along the trajectory. These fractions are then converted into contributions (in mm/h) by multiplying the fractions with the precipitation rate at the end of the trajectory. Finally, we calculate the median contribution from each LC class for each location and month.

The moisture source diagnostic reveals that the moisture supply for precipitation is fundamentally different between summer and winter. During winter, the vast majority water vapour originates from open water (Fig. C.7 a). Along the coasts, both AL_r and forests provide hardly any moisture for precipitation (Fig. C.7 c and e), indicating that the downwind precipitation increase due to forestation in those regions (Fig. 4.9 d and Fig. 4.10 d) is rather the result of altered atmospheric dynamics than an alteration in moisture supply. On the other hand, reduced downwind precipitation due to forestation in Central Europe and Scandinavia is at least partly explained by reduced moisture supply for precipitation following forestation (Fig. C.7 g). During summer, the contribution of open water to precipitation declines sharply moving away from the coast, which exemplifies the importance of moisture supply from land during this season (Fig. C.7 b). Both AL_r and forest contribute over-proportionally to precipitation in comparison to the upwind LC fractions associated with these LCs (Fig. C.7 d and f). However, this over-proportionality is more pronounced for forest, indicating that increased ET over forests in comparison to AL_r can explain the remote increase in precipitation following forestation that is found in the main part of this study (Fig. 4.9 e and Fig. 4.10 e).

While the results from the moisture source diagnostic conform overall with the ΔP^{rem} from the main part of this study, there are some limitations to this analysis. Firstly, the moisture source diagnostic cannot capture mechanisms other than moisture supply, such as slowing down the propagation of precipitating air masses or preferential triggering of convection over a certain land cover. This is most obvious in coastal regions during winter, where moisture contributions from AL_r and forest can clearly not explain the remote precipitation increase following forestation in those areas. Secondly, we assume that the moisture uptake at a certain position is originating from the different LC types proportionally to their fractional coverage, which is an oversimplification. Finally, it cannot account for local modifications in precipitation by specific LC types. Given our results of locally increased precipitation over forests in comparison to AL_r, an underestimation of the moisture supply from forests appears likely, as part of their moisture supply is compensated by increased precipitation over forests, which leads to a reduction of the specific humidity.

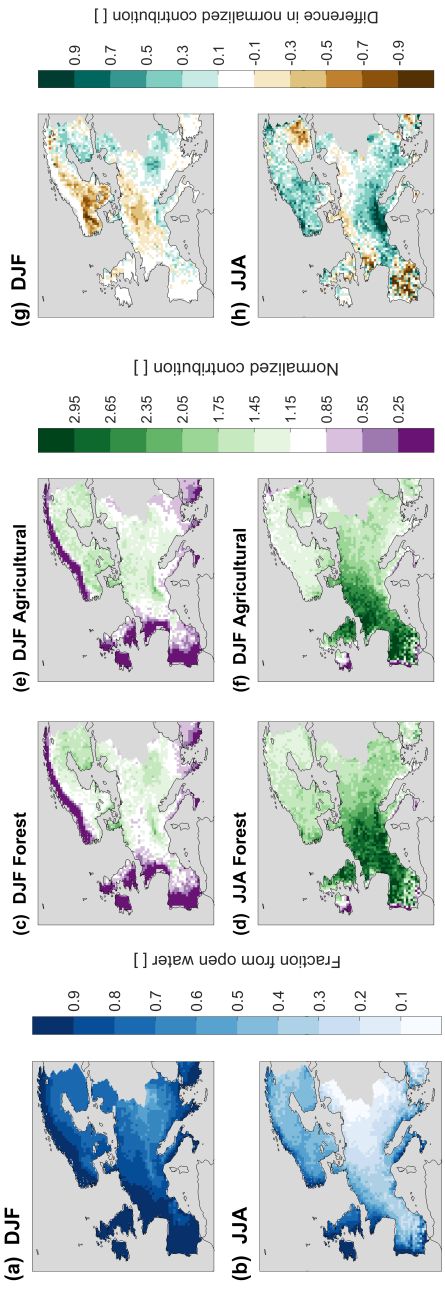


Figure C.7: *Panels (a), and (b) fraction of precipitation in ERA5 trajectories originating from open water. Panels (c)-(f) fraction of precipitation in ERA5 trajectories originating from forest (c, d) and AL_r (e, f) normalized by the upwind LC fraction of the respective LC. Values above one indicate that the respective LC is contributing over-proportionally to precipitation in comparison to its upwind LC fraction and below one under-proportionally. Panels (g) and (h) difference in normalized contribution between forest (c, d) and AL_r (e, f). Top row boreal winter and bottom row boreal summer.*

C.6 Appendix figures and tables

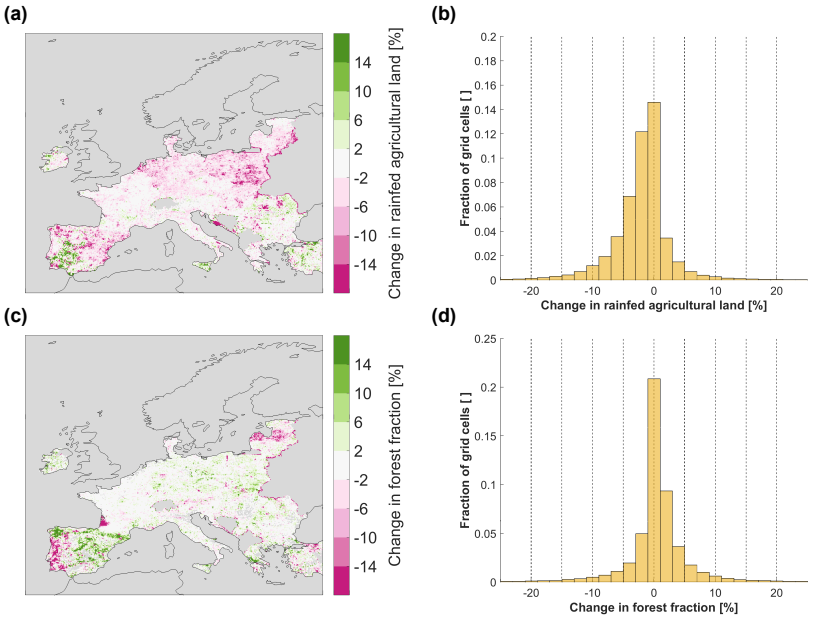


Figure C.8: ***Changes in AL_r and forest between 1990 and 2018.*** *Panel (a), spatial map of change in AL_r fraction according to CLC and panel (b) histogram of these changes. Panels (c) and (d) the same for the forest fraction. Note that fewer countries were included in CLC of 1990 resulting in missing data for some areas.*

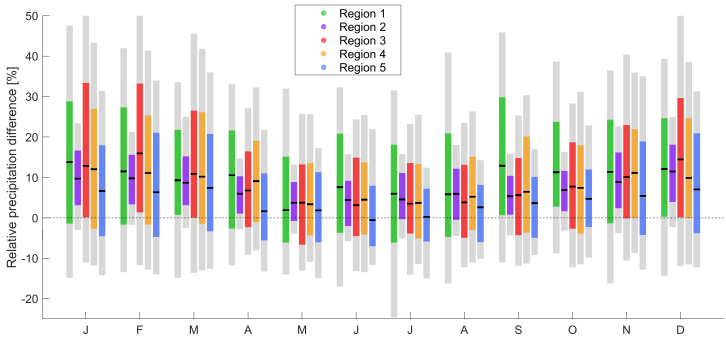


Figure C.9: *The median (black line), interquartile range (coloured shading), and range between 10^{th} and 90^{th} percentile (grey shading) ΔP^{l} in the rain gauge data sets as a fraction of the precipitation at the site with more AL_r over the five regions in Fig. 4.8 (a).*

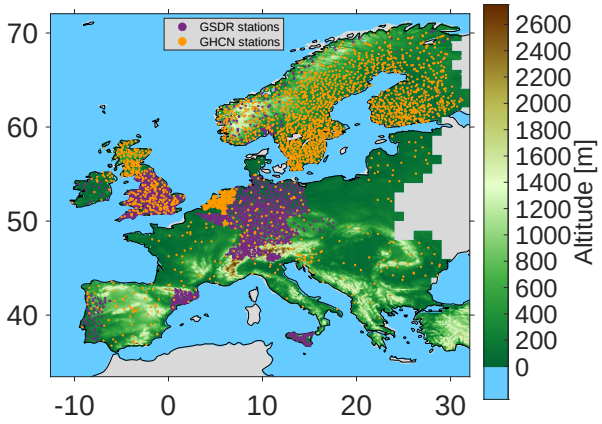


Figure C.10: *Location of the GSDR (purple) and GHCN (orange) rain gauge stations included in the site pair analysis.*

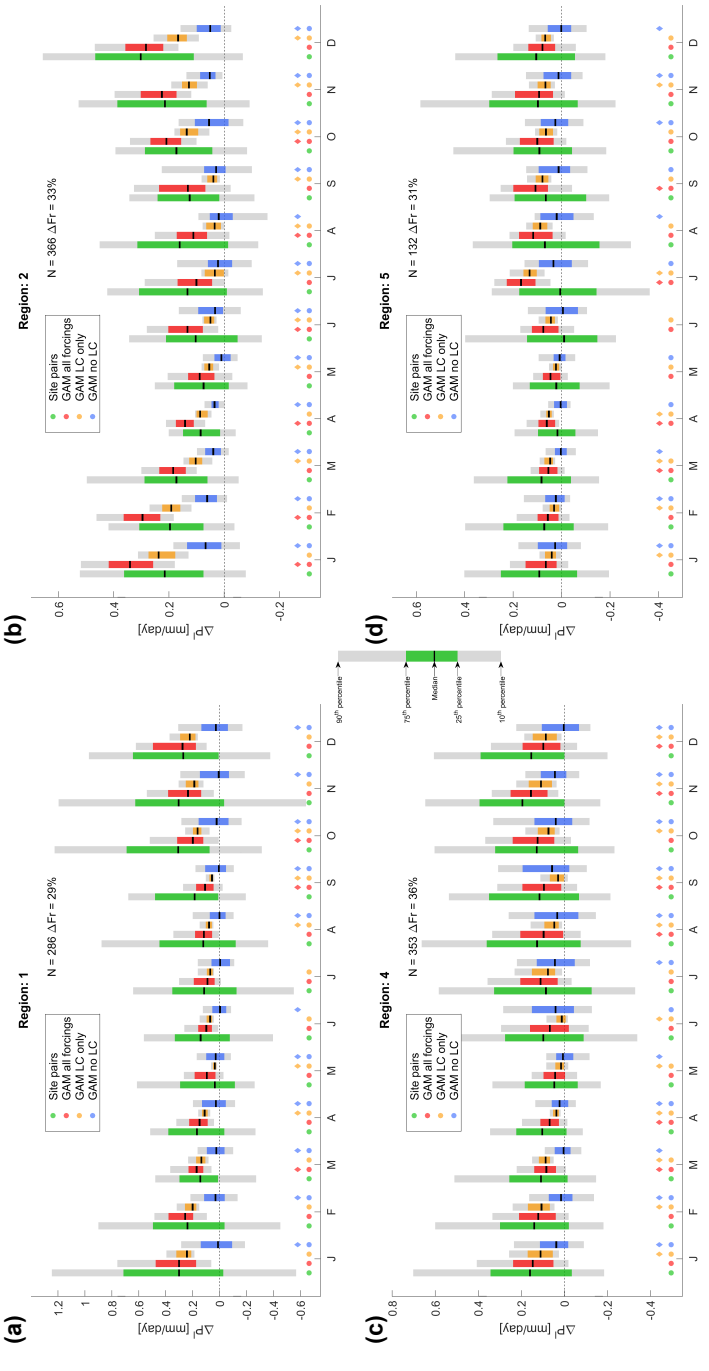
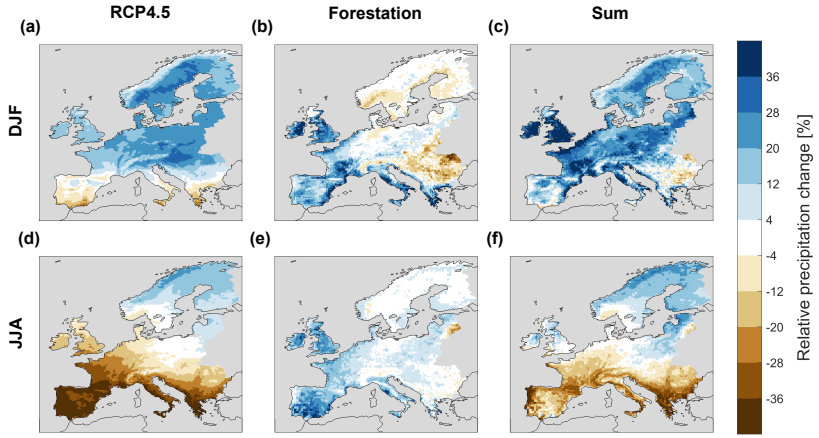


Figure C.11: *As Fig.4.8 (d) but for Regions 1, 2, 4, and 5.*



(g)

ΔP ± MAD [%]	DJF			JJA		
Region	RCP 4.5	Forestation	Sum	RCP 4.5	Forestation	Sum
Mediterranean	-0.19 ± 4.9	**23 ± 9.7**	23 ± 11	-19 ± 3.7	**19 ± 8.9**	-0.63 ± 9.1
Atlantic	6.3 ± 2.3	16 ± 11	23 ± 11	-8.7 ± 6.0	**14 ± 5.6**	5.7 ± 7.7
Alpine	7.4 ± 4.0	**-0.79 ± 4.2**	6.7 ± 5.6	3.2 ± 9.9	1.7 ± 2.5	4.9 ± 11
Continental	7.9 ± 3.2	**-1.0 ± 9.3**	6.9 ± 10	-6.0 ± 5.2	**4.8 ± 3.6**	-1.2 ± 6.4
Boreal	11 ± 2.2	**-2.1 ± 4.1**	8.6 ± 4.5	7.6 ± 3.1	2.9 ± 3.3	11 ± 5.1

Figure C.12: *As Fig. 4.11 but for RCP8.5.*

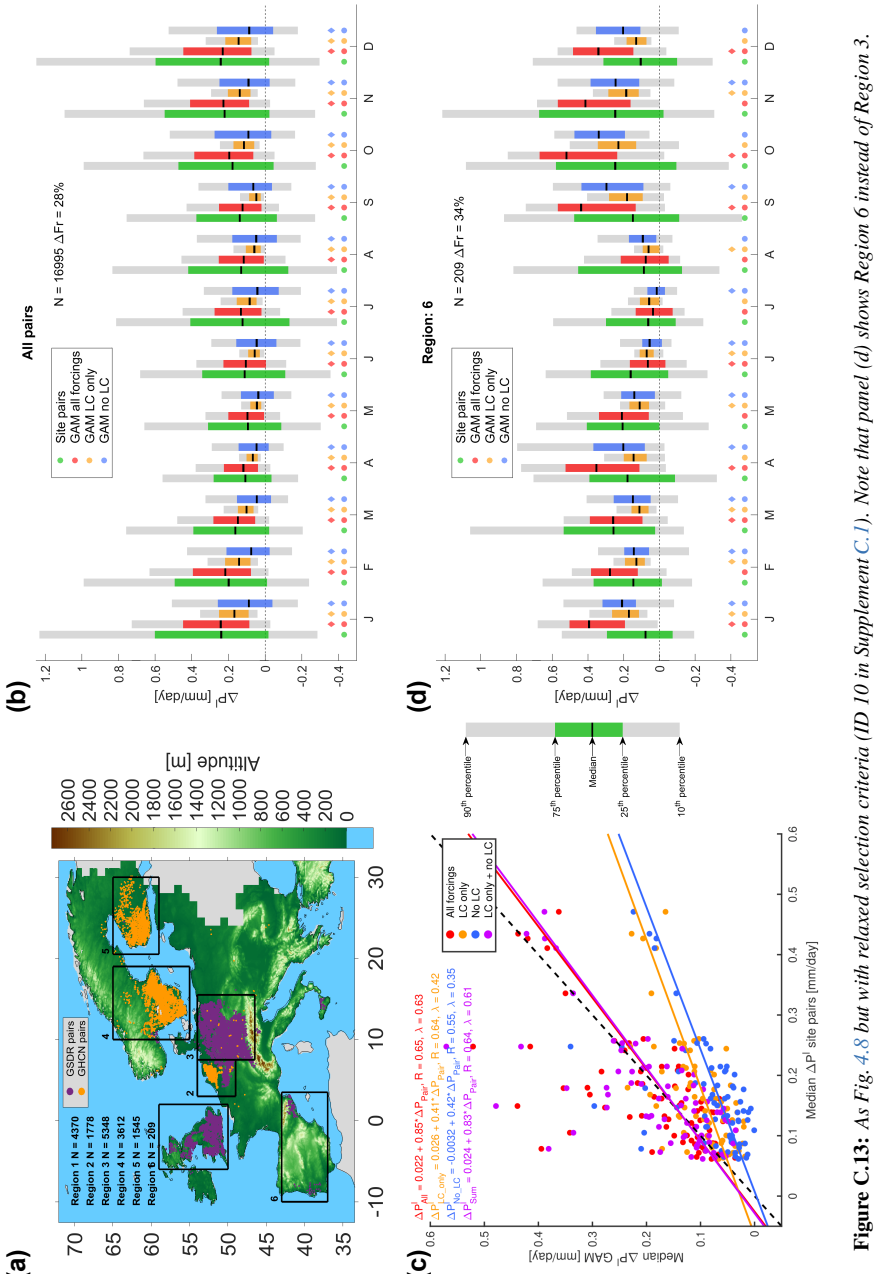


Figure C.13: *As Fig. 4.8 but with relaxed selection criteria (ID 10 in Supplement C.1). Note that panel (d) shows Region 6 instead of Region 3.*

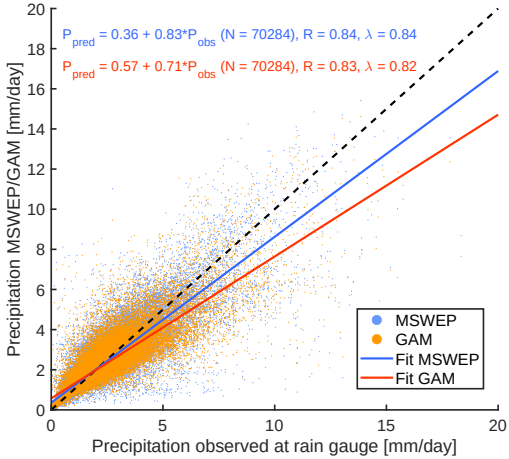


Figure C.14: *Comparison of monthly precipitation climatologies between the rain gauge data and original MSWEP precipitation climatology at the respective locations (green dots) and the GAM fit of MSWEP (orange dots). Black, green, and red lines show the 1:1 line, a linear fit of the original MSWEP data, and a linear fit of the GAM data, respectively. On top is shown the formula of the linear fit, the Pearson correlation coefficient, and the IA.*

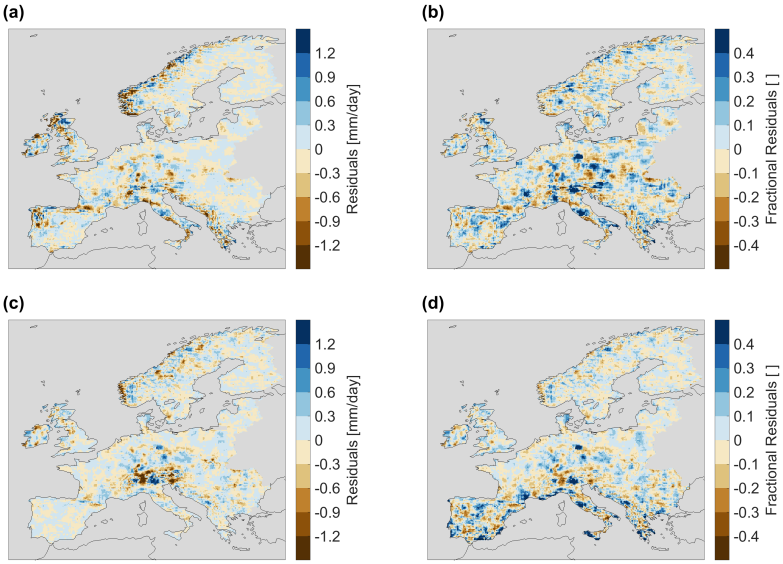


Figure C.15: *Residuals of GAM-modeled precipitation fields in January (a) and July (c). To the right the residuals as a fraction of the precipitation climatology in MSWEP in January (b) and July (d).*

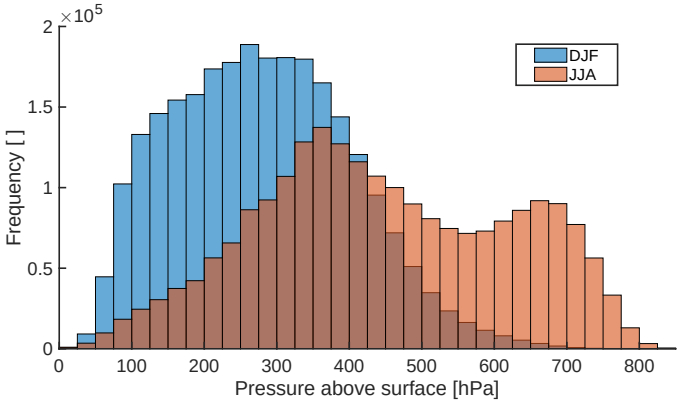


Figure C.16: *Histogram of the two vertical levels in ERA5 reanalysis data that produced most precipitation. Blue bars boreal winter (DJF) and red bars boreal summer (JJA). X-axis is pressure above the surface pressure.*

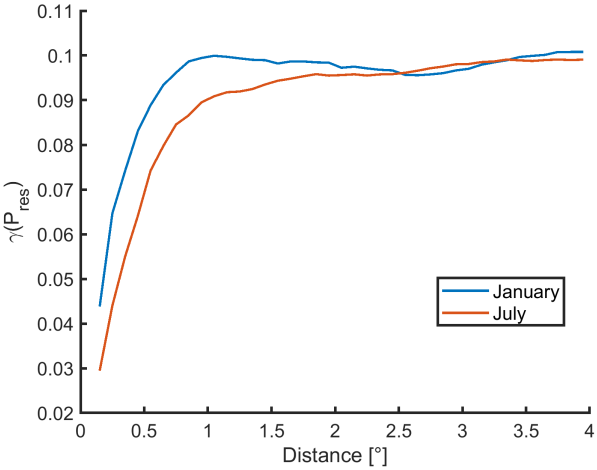


Figure C.17: *Area-weighted semivariogram of the GAM residuals in January (blue) and July (orange).*

List of symbols and abbreviations

Symbol/Abbr.	Long name/description [unit]
α	Albedo [ ]
β_t	Soil moisture stress function for photosynthesis and vegetation transpiration in CLM [ ]
$\Delta\alpha^{\,l}$	Local albedo change due to forestation [ ]
$\Delta\mathrm{DTR}^{\,l}$	Local diurnal temperature range change due to forestation [K]
$\Delta\mathrm{ET}^{\,l}$	Local ET change due to forestation [mm/day]
$\Delta\mathrm{LST}^{\,l}$	Local land surface temperature change due to forestation [K]
$\Delta\mathrm{LST}^{\,l}_{\mathrm{avg}}$	Daily average local land surface temperature change due to forestation [K]
$\Delta\mathrm{LST}^{\,l}_{\mathrm{max}}$	Daily maximum land surface temperature change due to forestation [K]
$\Delta\mathrm{LST}^{\,l}_{\mathrm{min}}$	Daily minimum local land surface temperature change due to forestation [K]
$\Delta\mathrm{LST}^{\,\mathrm{rem}}$	Remote land surface temperature change due to forestation [K]
$\Delta\mathrm{P}^{\,l}$	Local precipitation difference of sites with more forest minus site with more AL_r [mm/day]
$\Delta\mathrm{P}^{\,l}_{20\,\%}$	Local change in precipitation due to foresting 20 % rainfed agricultural land [mm/day]
$\Delta\mathrm{P}^{\,l}_{\mathrm{ref}}$	Local change in precipitation due to realistic reforestation scenario [mm/day]
$\Delta\mathrm{P}^{\,\mathrm{rem}}_{20\,\%}$	Downwind change in precipitation due to foresting 20 % rainfed agricultural land [mm/day]
$\Delta\mathrm{P}^{\,\mathrm{rem}}_{20\,\%}$	Downwind change in precipitation due to realistic reforestation scenario [mm/day]
$\Delta\mathrm{T2M}^{\,l}$	Local 2 m temperature change due to forestation [K]
$\Delta\mathrm{X}^{\,l}$	Local change in variable X due to biogeophysical effects of LCC
$\Delta\mathrm{X}^{\,\mathrm{rem}}$	Remote change in variable X due to biogeophysical effects of LCC
$\Delta\mathrm{X}^{\,\mathrm{tot}}$	Total change in variable X due to biogeophysical effects of LCC
AGB	Aboveground biomass [$\mathrm{kg\,m^{-2}}$]
AL_r	Rainfed agricultural land
ALB	Aboveground leaf biomass [$\mathrm{kg\,m^{-2}}$]
ANOVA	Analysis of variance
ASB	Aboveground stem biomass [$\mathrm{kg\,m^{-2}}$]
alt	Altitude [m]
BDT	Deciduous broadleaf tree
BET	Evergreen broadleaf tree
BHF	Biomass heat flux [$\mathrm{W\,m^{-2}}$]
BHS	Biomass heat storage

BTF	Broadleaf tree fraction [%]
CESM	Community Earth System Model
CLC	CORINE Land Cover 2000
CLM	Community Land Model
CMIP5	Coupled Model Intercomparison Project Phase 5
DJF	December/January/February (boreal winter)
DTR	Diurnal temperature range [k]
dw_hd	Downwind height difference [m]
ESM	Earth system model
ET	Evapotranspiration [mm/day]
expo	Exposition [° from North]
FAO	Food and Agriculture Organization of the United Nations
f_{rad}	Fraction of biomass heat flux that is compensated by decreased/increased emission of longwave radiation by the land surface [ ]
fREML	Fast restricted maximum likelihood method
GAM	Generalized additive model
GCM	General circulation model
GDAL	Geospatial Data Abstraction software Library 2019
GETA	Global Evapotranspiration Assembly
GHCN	Global Historical Climatology Network
GHG	Greenhouse gas
GLEAM	Global Land Evaporation Amsterdam Model
GMIA5	Global Map of Irrigation Areas version 5
GSDR	Global Sub-Daily Rainfall Dataset
GSWP3	Global Soil Wetness Project
IA	Index of agreement
IPCC	Intergovernmental Panel on Climate Change
JJA	June/July/August (boreal summer)
k	Maximum number of nodes (in a smooth/spline)
LAGRANTO	Lagrangian Analysis Tool
LC	Land cover
LCC	Land cover change
LU	Land use
LUC	Land use change
LUCAS	Land Use & Climate Across Scales (model intercomparison project)
LUCID	Land-use and Climate, Identification of Robust Impacts (model intercomparison project)
LULCC	Land use and land cover change
LSM	Land surface model
LST	Land surface temperature [K]

mgcv	Mixed GAM Computation Vehicle with Automatic Smoothness Estimation
MI	Moran's I [ ]
MODIS	Moderate Resolution Imaging Spectroradiometer
MSWEP	Multi-Source Weighted-Ensemble Precipitation version 2.2
NDT	Deciduous needleleaf tree
NET	Evergreen needleleaf tree
P	Precipitation [mm/day]
pct_LC$_i$	Fraction of land cover i [%]
PFT	Plant functional type
prox_LC$_i$	Proximity fraction of land cover i [%]
PS	Propagation speed of wind trajectories [m/s]
RCM	Regional climate model
RMSD	Root mean squared deviation
TPI	Topographic position index [m]
TRI	Terrain ruggedness index [m]
s	Thin plate regression spline
SEVIRI	Spinning Enhanced Visible and Infrared Imager
sp	Smoothing parameter
T2M	2 m air temperature [K]
ti	Tensor product smooth
uw_cd	Upwind distance to coast [° from coast]
uw_ct	Upwind time to coast [h]
uw_hd	Upwind height difference [m]
uw_pct_LC$_i$	Upwind fraction of land cover i [%]
V_{cmax}	Maximum rate of carboxylation [µmol m^{-2} s^{-1}]
WRPM	World Reforestation Potential Map
WS	Wind speed 10 m above surface [m/s]

Bibliography

Aemisegger, F., S. Pfahl, H. Sodemann, I. Lehner, S. I. Seneviratne, and H. Wernli (2014). "Deuterium excess as a proxy for continental moisture recycling and plant transpiration". *Atmos. Chem. Phys.*, 14.8, 4029–4054. DOI: 10.5194/acp-14-4029-2014.

Alkama, R. and A. Cescatti (2016). "Biophysical climate impacts of recent changes in global forest cover". *Science*, 351, 600–604. DOI: 10.1126/science.aac8083.

Ambrose, S. M. and S. M. Sterling (2014). "Global patterns of annual actual evapotranspiration with land-cover type: Knowledge gained from a new observation-based database". *Hydrol. Earth Syst. Sc. Discuss.*, 11, 12103–12135. DOI: 10.5194/hessd-11-12103-2014.

Aston, A. (1985). "Heat storage in a young eucalypt forest". *Agr. Forest Meteorol.*, 35.1, 281–297. DOI: 10.1016/0168-1923(85)90090-5.

Azar, C., K. Lindgren, M. Obersteiner, et al. (2010). "The feasibility of low CO2 concentration targets and the role of bio-energy with carbon capture and storage (BECCS)". *Climatic Change*, 100, 195–202. DOI: 10.1007/s10584-010-9832-7.

Bala, G., K. Caldeira, M. Wickett, T. J. Phillips, D. B. Lobell, C. Delire, and A. Mirin (2007). "Combined climate and carbon-cycle effects of large-scale deforestation". *P. Natl. Acad. Sci. USA*, 104.16, 6550–6555. DOI: 10.1073/pnas.0608998104.

Baldocchi, D. and S. Ma (2013). "How will land use affect air temperature in the surface boundary layer? Lessons learned from a comparative study on the energy balance of an oak savanna and annual grassland in California, USA". *Tellus B*, 65.1, 19994. DOI: 10.3402/tellusb.v65i0.19994.

Bathiany, S., M. Claussen, V. Brovkin, T. Raddatz, and V. Gayler (2010). "Combined biogeophysical and biogeochemical effects of large-scale forest cover changes in the MPI earth system model". *Biogeosciences*, 7.5, 1383–1399. DOI: 10.5194/bg-7-1383-2010.

Beck, H. E., M. Pan, T. Roy, G. P. Weedon, F. Pappenberger, A. I. J. M. van Dijk, G. J. Huffman, R. F. Adler, and E. F. Wood (2019a). "Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS". *Hydrol. Earth Syst. Sc.*, 23.1, 207–224. DOI: 10.5194/hess-23-207-2019.

Beck, H. E., A. I. J. M. van Dijk, V. Levizzani, J. Schellekens, D. G. Miralles, B. Martens, and A. de Roo (2017a). "MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data". *Hydrol. Earth Syst. Sc.*, 21.1, 589–615. DOI: 10.5194/hess-21-589-2017.

Beck, H. E., N. Vergopolan, M. Pan, V. Levizzani, A. I. J. M. van Dijk, G. P. Weedon, L. Brocca, F. Pappenberger, G. J. Huffman, and E. F. Wood (2017b). "Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling". *Hydrology and Earth System Sciences*, 21.12, 6201–6217. DOI: 10.5194/hess-21-6201-2017.

Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler (2019b). "MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment". *B. Am. Meteorol. Soc.*, 100.3, 473–500. DOI: 10.1175/BAMS-D-17-0138.1.

Belušić, D., R. Fuentes-Franco, G. Strandberg, and A. Jukimenko (2019). "Afforestation reduces cyclone intensity and precipitation extremes over Europe". *Environ. Res. Lett.*, 14.7, 074009. DOI: 10.1088/1748-9326/ab23b2.

Betts, R. A. (2001). "Biogeophysical impacts of land use on present-day climate: near-surface temperature change and radiative forcing". *Atmos. Sci. Lett.*, 2.1, 39–51. DOI: 10.1006/asle.2001.0023.

Bhumralkar, C. M. (1975). "Numerical experiments on the computation of ground surface temperature in an atmospheric general circulation model". *J. Appl. Meteorol.*, 14, 1246–1258. DOI: 10.1175/1520-0450(1975)014<1246:NEOTCO>2.0.CO;2.

Blenkinsop, S., E. Lewis, S. C. Chan, and H. J. Fowler (2017). "Quality-control of an hourly rainfall dataset and climatology of extremes for the UK". *International Journal of Climatology*, 37.2, 722–740. DOI: 10.1002/joc.4735.

Boenisch, G. and J. Kattge (2017). *TRY Plant Trait Database*.

Bonan, G. (2008). "Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests". *Science*, 320, 1444–1449. DOI: 10.1126/science.1155121.

Bonan, G. B., E. G. Patton, I. N. Harman, K. W. Oleson, J. J. Finnigan, Y. Lu, and E. A. Burakowski (2018). "Modeling canopy-induced turbulence in the Earth system: a unified parameterization of turbulent exchange within plant canopies and the roughness sublayer (CLM-ml v0)". *Geosci. Model Dev.*, 11.4, 1467–1496. DOI: 10.5194/gmd-11-1467-2018.

Bonan, G. B., D. Pollar, and S. Thompson (1992). "Effects of boreal forest vegetation on global climate". *Nature*, 359, 716–718. DOI: 10.1038/359716a0.

Bonan, G., P. J. Lawrence, K. W. Oleson, S. Levis, M. Jung, M. Reichstein, D. M. Lawrence, and S. C. Swenson (2011). "Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data". *J. Geophys. Res.-Biogeo.*, 116.G02014. DOI: 10.1029/2010JG001593.

Bonan, G., K. W. Oleson, R. A. Fisher, G. Lasslop, and M. Reichstein (2012). "Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4". *J. Geophys. Res.-Biogeo.*, 117.G02026. DOI: 10.1029/2011JG001913.

Boone, A., P. Samuelsson, S. Gollvik, A. Napoly, L. Jarlan, E. Brun, and B. Decharme (2017). "The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8 – Part 1: Model description". *Geosci. Model Dev.*, 10.2, 843–872. DOI: 10.5194/gmd-10-843-2017.

Bosman, P. J. M., C. C. van Heerwaarden, and A. J. Teuling (2019). "Sensible heating as a potential mechanism for enhanced cloud formation over temperate forest". *Q. J. Roy. Meteor. Soc.*, 145.719, 450–468. DOI: 10.1002/qj.3441.

Breil, M., D. Rechid, E. Davin, N. de Noblet-Ducoudré, E. Katragkou, R. Cardoso, P. Hoffmann, L. Jach, P. Soares, G. Sofiadis, S. Strada8, G. Strandberg, M. Tölle, and K. Warrach-Sagi (2020). "The opposing effects of re/af-forestation on the diurnal temperature cycle at the surface and in the lowest atmospheric model level in the European summer". *J. Climate*, 1–58. DOI: 10.1175/JCLI-D-19-0624.1.

Bright, R. M., E. L. Davin, T. O'Halloran, J. Pongratz, K. Zhao, and A. Cescatti (2017). "Local temperature response to land cover and management change driven by non-radiative processes". *Nat. Clim. Change*, 7, 296–302. DOI: 10.1038/nclimate3250.

Brown, S., J. Sathaye, M. Cannell, and P. E. Kauppi (1996). "Mitigation of carbon emissions to the atmosphere by forest management". *Commonw. Forest. Rev.*, 75, 80–91.

Buendia, C., R. J. Batalla, S. Sabater, A. Palau, and R. Marcé (2016). "Runoff Trends Driven by Climate and Afforestation in a Pyrenean Basin". *Land Degradation & Development*, 27.3, 823–838. DOI: 10.1002/ldr.2384.

Burakowski, E., A. Tawfik, A. Ouimette, L. Lepine, K. Novick, S. Ollinger, C. Zarzycki, and G. Bonan (2018). "The role of surface roughness, albedo, and Bowen ratio on ecosystem energy balance in the Eastern United States". *Agr. Forest Meteorol.*, 249, 367–376. DOI: 10.1016/j.agrformet.2017.11.030.

Burns, S. P., P. D. Blanken, A. A. Turnipseed, J. Hu, and R. K. Monson (2015). "The influence of warm-season precipitation on the diel cycle of the surface energy balance and carbon dioxide at a Colorado subalpine forest site". *Biogeosciences*, 12.23, 7349–7377. DOI: 10.5194/bg-12-7349-2015.

Canadell, J., R. B. Jackson, J. B. Ehleringer, H. A. Mooney, O. E. Sala, and E.-D. Schulze (1996). "Maximum rooting depth of vegetation types at the global scale". *Oecologia*, 108.4, 583–595. DOI: 10.1007/BF00329030.

Cescatti, A., B. Marcolla, S. K. S. Vannan, J. Y. Pan, M. O. Román, X. Yang, P. Ciais, R. B. Cook, B. E. Law, G. Matteucci, M. Migliavacca, E. Moors, A. D. Richardson, G. Seufert, and C. B. Schaaf (2012). "Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network". *Remote Sens. Environ.*, 121, 323–334. DOI: 10.1016/j.rse.2012.02.019.

CH2018 (2018). *2018 Climate scenarios for Switzerland*. Tech. rep. Zurich: National Centre for Climate Services.

Chan, K. M. and R. Wood (2013). "The seasonal cycle of planetary boundary layer depth determined using COSMIC radio occultation data". *J. Geophys. Res.-Atmos.*, 118.22, 12, 422–12, 434. DOI: 10.1002/2013JD020147.

Chen, C., T. Park, X. Wang, et al. (2019). "China and India lead in greening of the world through land-use management." *Nat. Sustain.*, 2, 122–129. DOI: 10.1038/s41893-019-0220-7.

Chen, L. and P. A. Dirmeyer (2020). "Reconciling the disagreement between observed and simulated temperature responses to deforestation". *Nat. Commun.*, 11.202. DOI: 10.1038/s41467-019-14017-0.

Chen, L., P. A. Dirmeyer, Z. Guo, and N. M. Schultz (2018). "Pairing FLUXNET sites to validate model representations of land-use/land-cover change". *Hydrol. Earth Syst. Sc. Discuss.*, 22.1, 111–125. DOI: 10.5194/hess-22-111-2018.

Chen, L. and P. A. Dirmeyer (2016). "Adapting observationally based metrics of biogeophysical feedbacks from land cover/land use change to climate modeling". *Environ. Res. Lett.*, 11.3, 034002. DOI: 10.1088/1748-9326/11/3/034002.

Chen, L. and P. A. Dirmeyer (2019a). "Differing Responses of the Diurnal Cycle of Land Surface and Air Temperatures to Deforestation". *J. Climate*, 32.20, 7067–7079. DOI: 10.1175/JCLI-D-19-0002.1.

— (2019b). "Global observed and modelled impacts of irrigation on surface temperature". *Int. J. Climatol.*, 39.5, 2587–2600. DOI: 10.1002/joc.5973.

Cherubini, F., B. Huang, X. Hu, M. H. Tölle, and A. H. Strømman (2018a). "Quantifying the climate response to extreme land cover changes in Europe with a regional model". *Environ. Res. Lett.*, 13.074002. DOI: 10.1088/1748-9326/aac794.

Cherubini, F., F. Santaniello, X. Hu, J. Sonesson, A. H. Strømman, J. Weslien, L. B. Djupström, and T. Ranius (2018b). "Climate impacts of retention forestry in a Swedish boreal pine forest". *J. Land Use Sci.*, 13.3, 301–318. DOI: 10.1080/1747423X.2018.1529831.

Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, C. A., R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Quéré, R. B. Myneni, S. Piao, and P. Thornton (2013). *Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panle on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A.*

Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)] Cambridge, United Kingdom and New York, NY, USA.

Claussen, M., V. Brovkin, and A. Ganopolski (2001). "Biogeophysical versus biogeochemical feedbacks of large-scale land cover change". *Geophys. Res. Lett.*, 28.6, 1011–1014. DOI: 10.1029/2000GL012471.

Cohn, A. S., N. Bhattarai, J. Campolo, O. Crompton, D. Dralle, J. Duncan, and S. Thompson (2019). "Forest loss in Brazil increases maximum temperatures within 50 km". *Environ. Res. Lett.*, 14.8, 084047. DOI: 10.1088/1748-9326/ab31fb.

Copernicus Climate Change Service (C3S) (2017). *ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate*.

— (2019). *C3S ERA5-Land reanalysis*.

Crowther, T., H. Glick, K. Covey, C. Bettigole, D. Maynard, S. Thomas, J. Smith, G. Hintler, M. Duguid, G. Amatulli, M.-N. Tuanmu, W. Jetz, C. Salas-Eljatib, C. Stam, D. Piotto, R. Tavani, S. Green, G. Bruce, S. Williams, and M. Bradford (2015). "Mapping tree density at a global scale". *Nature*, advance online publication. DOI: 10.1038/nature14967.

da Rocha, H. R., M. L. Goulden, S. D. Miller, M. C. Menton, L. D. V. O. Pinto, H. C. de Freitas, and A. M. e Silva Figueira (2004). "Seasonality of Water and Heat Fluxes Over a Tropical Forest in Eastern Amazonia". *Ecol. Appl.*, 14.sp4, 22–32. DOI: 10.1890/02-6001.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008). "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States". *Int. J. Climatol.*, 28.15, 2031–2064. DOI: 10.1002/joc.1688.

Davin, E. L. and N. de Noblet-Ducoudré (2010). "Climatic impact of global-scale deforestation: Radiative versus nonradiative processes". *J. Climate*, 23, 97–112. DOI: 10.1175/2009JCLI3102.1.

Davin, E. L., N. de Noblet-Ducoudré, and P. Friedlingstein (2007). "Impact of land cover change on surface climate: Relevance of the radiative forcing concept". *Geophys. Res. Lett.*, 34.L13702. DOI: 10.1029/2007GL029678.

Davin, E. L., D. Rechid, M. Breil, R. M. Cardoso, E. Coppola, P. Hoffmann, L. L. Jach, E. Katragkou, N. de Noblet-Ducoudré, K. Radtke, M. Raffa, P. M. M. Soares, G. Sofiadis, S. Strada, G. Strandberg, M. H. Tölle, K. Warrach-Sagi, and V. Wulfmeyer (2020). "Biogeophysical impacts of forestation in Europe: First results from the LUCAS Regional Climate Model intercomparison". *Earth Syst. Dynam.*, 11, 183–200. DOI: 10.5194/esd-11-183-2020.

Davin, E. L. and S. I. Seneviratne (2012). "Role of land surface processes and diffuse/direct radiation partitioning in simulating the European climate". *Biogeosciences*, 9.5, 1695–1707. DOI: 10.5194/bg-9-1695-2012.

Davin, E. L., E. Maisonnave, and S. I. Seneviratne (2016). "Is land surface processes representation a possible weak link in current Regional Climate Models?" *Environ. Res. Lett.*, 11.7, 074027.

Davin, E. L., S. I. Seneviratne, P. Ciais, A. Olioso, and T. Wang (2014). "Preferential cooling of hot extremes from cropland albedo management". *P. Natl. Acad. Sci. USA*, 111.27, 9757–9761. DOI: 10.1073/pnas.1317323111.

Davin, E. L., R. Stöckli, E. B. Jaeger, S. Levis, and S. I. Seneviratne (2011). "COSMO-CLM2: a new version of the COSMO-CLM model coupled to the Community Land Model". *Clim. Dynam.*, 37.9, 1889–1907. DOI: 10.1007/s00382-011-1019-z.

de Noblet-Ducoudré, N., J.-P. Boisier, A. Pitman, G. B. Bonan, V. Brovkin, F. Cruz, C. Delire, V. Gayler, B. J. J. M. van den Hurk, P. J. Lawrence, M. K. van der Molen, C. Müller, C. H. Reick,

B. J. Strengers, and A. Voldoire (2012). "Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: Results from the first set of LUCID experiments". *J. Climate*, 25, 3261–3181. DOI: 10.1175/JCLI-D-11-00338.1.

de Oliveira, G., N. A. Brunsell, E. C. Moraes, Y. E. Shimabukuro, T. V. dos Santos, C. von Randow, R. G. de Aguiar, and L. E. Aragao (2019). "Effects of land-cover changes on the partitioning of surface energy and water fluxes in Amazonia using high-resolution satellite imagery". *Ecohydrology*, 12.6. e2126 ECO-18-0089.R1, e2126. DOI: 10.1002/eco.2126.

DeAngelis, A., F. Dominguez, Y. Fan, A. Robock, M. D. Kustu, and D. Robinson (2010). "Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States". *J. Geophys. Res.-Atmos.*, 115.D15. DOI: 10.1029/2010JD013892.

Devanand, A., M. Huang, D. M. Lawrence, C. M. Zarzycki, Z. Feng, P. J. Lawrence, Y. Qian, and Z. Yang (2020). "Land use and land cover change strongly modulates land-atmosphere coupling and warm-season precipitation over the Central United States in CESM2-VR". *J. Adv. Model Earth Sy.*, n/a.n/a. e2019MS001925 2019MS001925, e2019MS001925. DOI: 10.1029/2019MS001925.

Devaraju, N., G. Bala, and R. Nemani (2015). "Modelling the influence of land-use changes on biophysical and biochemical interactions at regional and global scales". *Plant Cell Environ.*, 38.9, 1931–1946. DOI: 10.1111/pce.12488.

Devaraju, N., N. de Noblet-Ducoudré, B. Quesada, and G. Bala (2018). "Quantifying the Relative Importance of Direct and Indirect Biophysical Effects of Deforestation on Surface Temperature and Teleconnections". *J. Climate*, 31.10, 3811–3829. DOI: 10.1175/JCLI-D-17-0563.1.

Dorman, J. L. and P. J. Sellers (1989). "A Global Climatology of Albedo, Roughness Length and Stomatal Resistance for Atmospheric General Circulation Models as Represented by the Simple Biosphere Model (SiB)". *J. Appl. Meteorol.*, 28.9, 833–855. DOI: 10.1175/1520-0450(1989)028<0833:AGCOAR>2.0.CO;2.

Dormann, C., J. McPherson, M. Araújo, R. Bivand, J. Bolliger, G. Carl, R. Davies, A. Hirzel, W. Jetz, W. Kissling, I. Kühn, R. Ohlemüller, P. Peres-Neto, B. Reineking, B. Schröder, F. Schurr, and R. Wilson (2007). "Methods to Account for Spatial Autocorrelation in the Analysis of Species Distributional Data: A Review". *Ecography*, 30, 609–628. DOI: 10.1111/j.2007.0906-7590.05171.x.

Duveiller, G., G. Forzieri, E. Robertson, W. Li, G. Georgievski, P. Lawrence, A. Wiltshire, P. Ciais, J. Pongratz, S. Sitch, A. Arneth, and A. Cescatti (2018a). "Biophysics and vegetation cover change: a process-based evaluation framework for confronting land surface models with satellite observations". *Earth System Science Data Discussions*, 1–24. DOI: 10.5194/essd-2018-24.

Duveiller, G., J. Hooker, and A. Cescatti (2018b). "The mark of vegetation change on Earth's surface energy balance". *Nat. Commun.*, 9.679. DOI: 10.5194/essd-2018-24.

Duveiller, G., D. Fasbender, and M. Meroni (2016). "Revisiting the concept of a symmetric index of agreement for continuous datasets". *Sci. Rep.-UK*, 6, 19401. DOI: 10.1038/srep19401.

Edburg, S. L., J. A. Hicke, P. D. Brooks, E. G. Pendall, B. E. Ewers, U. Norton, D. Gochis, E. D. Gutmann, and A. J. Meddens (2012). "Cascading impacts of bark beetle-caused tree mortality on coupled biogeophysical and biogeochemical processes". *Front. Ecol. Environ.*, 10.8, 416–424. DOI: 10.1890/110173.

EEA (2014). *EU-DEM Statistical Validation*. Copenhagen, Denmark.

Ellison, D., C. E. Morris, B. Locatelli, D. Sheil, J. Cohen, D. Murdiyarso, V. Gutierrez, M. van Noordwijk, I. F. Creed, J. Pokorny, D. Gaveau, D. V. Spracklen, A. B. Tobella, U. Ilstedt, A. J. Teuling, S. G. Gebrehiwot, D. C. Sands, B. Muys, B. Verbist, E. Springgay, Y. Sugandi, and

C. A. Sullivan (2017). "Trees, forests and water: Cool insights for a hot world". *Global Environ. Chang.*, 43, 51–61. DOI: 10.1016/j.gloenvcha.2017.01.002.

Fan, Y., G. Miguez-Macho, E. G. Jobbágy, R. B. Jackson, and C. Otero-Casal (2017). "Hydrologic regulation of plant rooting depth". *P. Natl. Acad. Sci. USA*, 114.40, 10572–10577. DOI: 10.1073/pnas.1712381114.

FAO 2017: The future of food and agriculture Trends and challenges. (2017). Tech. rep.

Fatichi, S., M. J. Zeeman, J. Fuhrer, and P. Burlando (2014). "Ecohydrological effects of management on subalpine grasslands: From local to catchment scale". *Water Resour. Res.*, 50.1, 148–164. DOI: 10.1002/2013WR014535.

Findell, K. L., E. Shevliakova, P. C. D. Milly, and R.-J. Stouffer (2007). "Modeled impact of anthropogenic land cover change on climate". *J. Climate*, 20, 3621–3634. DOI: 10.1175/JCLI4185.1.

Fisher, R. A., C. D. Koven, W. R. L. Anderegg, B. O. Christoffersen, M. C. Dietze, C. E. Farrior, J. A. Holm, G. C. Hurtt, R. G. Knox, P. J. Lawrence, J. W. Lichstein, M. Longo, A. M. Matheny, D. Medvigy, H. C. Muller-Landau, T. L. Powell, S. P. Serbin, H. Sato, J. K. Shuman, B. Smith, A. T. Trugman, T. Viskari, H. Verbeeck, E. Weng, C. Xu, X. Xu, T. Zhang, and P. R. Moorcroft (2018). "Vegetation demographics in Earth System Models: A review of progress and priorities". *Glob. Change Biol.*, 24.1, 35–54. DOI: 10.1111/gcb.13910.

Foken, T. (2008). "The energy balance closure problem: An overview". *Ecol. Appl.*, 18.6, 1351–1367. DOI: 10.1890/06-0922.1.

Forzieri, G., D. G. Miralles, P. Ciais, et al. (2020). "Increased control of vegetation on global terrestrial energy fluxes". *Nat. Clim. Change*, 10, 356–362. DOI: 10.1038/s41558-020-0717-0.

Friedlingstein, P., M. O'Sullivan, M. W. Jones, R. M. Andrew, J. Hauck, A. Olsen, G. P. Peters, W. Peters, J. Pongratz, S. Sitch, C. Le Quéré, J. G. Canadell, P. Ciais, R. B. Jackson, S. Alin, L. E. O. C. Aragão, A. Arneth, V. Arora, N. R. Bates, M. Becker, A. Benoit-Cattin, H. C. Bittig, L. Bopp, S. Bultan, N. Chandra, F. Chevallier, L. P. Chini, W. Evans, L. Florentie, P. M. Forster, T. Gasser, M. Gehlen, D. Gilfillan, T. Gkritzalis, L. Gregor, N. Gruber, I. Harris, K. Hartung, V. Haverd, R. A. Houghton, T. Ilyina, A. K. Jain, E. Joetzjer, K. Kadono, E. Kato, V. Kitidis, J. I. Korsbakken, P. Landschützer, N. Lefèvre, A. Lenton, S. Lienert, Z. Liu, D. Lombardozzi, G. Marland, N. Metzl, D. R. Munro, J. E. M. S. Nabel, S.-I. Nakaoka, Y. Niwa, K. O'Brien, T. Ono, P. I. Palmer, D. Pierrot, B. Poulter, L. Resplandy, E. Robertson, C. Rödenbeck, J. Schwinger, R. Séférian, I. Skjelvan, A. J. P. Smith, A. J. Sutton, T. Tanhua, P. P. Tans, H. Tian, B. Tilbrook, G. van der Werf, N. Vuichard, A. P. Walker, R. Wanninkhof, A. J. Watson, D. Willis, A. J. Wiltshire, W. Yuan, X. Yue, and S. Zaehle (2020). "Global Carbon Budget 2020". *Earth Syst. Sci. Data*, 12.4, 3269–3340. DOI: 10.5194/essd-12-3269-2020.

Fuchs, R., M. Herold, P. H. Verburg, and J. G. P. W. Clevers (2013). "A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe". *Biogeosciences*, 10.3, 1543–1559. DOI: 10.5194/bg-10-1543-2013.

Fuchs, R., R. Prestele, and P. H. Verburg (2018). "A global assessment of gross and net land change dynamics for current conditions and future scenarios". *Earth Syst. Dynam.*, 9.2, 441–458. DOI: 10.5194/esd-9-441-2018.

Fuchs, R., M. Herold, P. H. Verburg, J. G. Clevers, and J. Eberle (2014). "Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010". *Glob. Change Biol.*, 21.1, 299–313. DOI: 10.1111/gcb.12714.

Fuchs, R., P. H. Verburg, J. G. Clevers, and M. Herold (2015). "The potential of old maps and encyclopaedias for reconstructing historic European land cover/use change". *Applied Geography*, 59, 43–55. DOI: 10.1016/j.apgeog.2015.02.013.

Fuhrer, O., T. Chadha, T. Hoefler, G. Kwasniewski, X. Lapillonne, D. Leutwyler, D. Lüthi, C. Osuna, C. Schär, T. C. Schulthess, and H. Vogt (2018). "Near-global climate simulation at 1 km resolution: establishing a performance baseline on 4888 GPUs with COSMO 5.0". *Geosci. Model Dev.*, 11.4, 1665–1681. DOI: 10.5194/gmd-11-1665-2018.

Garai, A., J. Kleissl, and S. G. Llewellyn Smith (2010). "Estimation of Biomass Heat Storage Using Thermal Infrared Imagery: Application to a Walnut Orchard". *Boundary-Lay. Meteorol.*, 137.2, 333–342. DOI: 10.1007/s10546-010-9524-x.

Gash, J. H. C. and J. B. Stewart (1979). "An analytical model of rainfall interception by forests". *Q. J. Roy. Meteor. Soc.*, 105, 43–55.

Ge, J., W. Guo, A. J. Pitman, M. G. De Kauwe, X. Chen, and C. Fu (2019). "The Nonradiative Effect Dominates Local Surface Temperature Change Caused by Afforestation in China". *J. Climate*, 32.14, 4445–4471. DOI: 10.1175/JCLI-D-18-0772.1.

Ghimire, B., C. A. Williams, G. J. Collatz, and M. Vanderhoof (2012). "Fire-induced carbon emissions and regrowth uptake in western U.S. forests: Documenting variation across forest types, fire severity, and climate regions". *J. Geophys. Res.-Biogeo.*, 117.G3. DOI: 10.1029/2011JG001935.

Gibbs, H. and J. Salmon (2015). "Mapping the world's degraded lands". *Applied Geography*, 57, 12–21. DOI: 10.1016/j.apgeog.2014.11.024.

Griscom, B. W., J. Adams, P. W. Ellis, R. A. Houghton, G. Lomax, D. A. Miteva, W. H. Schlesinger, D. Shoch, J. V. Siikamäki, P. Smith, P. Woodbury, C. Zganjar, A. Blackman, J. Campari, R. T. Conant, C. Delgado, P. Elias, T. Gopalakrishna, M. R. Hamsik, M. Herrero, J. Kiesecker, E. Landis, L. Laestadius, S. M. Leavitt, S. Minnemeyer, S. Polasky, P. Potapov, F. E. Putz, J. Sanderman, M. Silvius, E. Wollenberg, and J. Fargione (2017a). "Natural climate solutions". *P. Natl. Acad. Sci. USA*, 114.44, 11645–11650. DOI: 10.1073/pnas.1710465114.

Griscom, B. W., J. Adams, P. W. Ellis, R. A. Houghton, G. Lomax, D. A. Miteva, W. H. Schlesinger, D. Shoch, J. V. Siikamäki, P. Smith, P. Woodbury, C. Zganjar, A. Blackman, J. Campari, R. T. Conant, C. Delgado, P. Elias, T. Gopalakrishna, M. R. Hamsik, J. K. Mario Herrero, E. Landis, L. Laestadius, S. M. Leavitt, S. Minnemeyer, S. Polasky, P. Potapov, F. E. Putz, J. Sanderman, M. Silvius, E. Wollenberg, and J. Fargione (2017b). *Global Reforestation Potential Map*. DOI: 10.5281/zenodo.883444.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). "High-Resolution Global Maps of 21st-Century Forest Cover Change". *Science*, 342, 850–853. DOI: 10.1126/science.1244693.

Harding, R. J., S.-E. Gryning, S. Halldin, and C. R. Lloyd (2001). "Progress in understanding of land surface/atmosphere exchanges at high latitudes". *Theor. Appl. Climatol.*, 70.1, 5–18. DOI: 10.1007/s007040170002.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2014). "Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset". *Int. J. Climatol.*, 34.3, 623–642. DOI: 10.1002/joc.3711.

Hauser, M., W. Thiery, and S. I. Seneviratne (2019). "Potential of global land water recycling to mitigate local temperature extremes". *Earth Syst. Dynam.*, 10.1, 157–169. DOI: 10.5194/esd-10-157-2019.

Haverd, V., M. Cuntz, R. Leuning, and H. Keith (2007). "Air and biomass heat storage fluxes in a forest canopy: Calculation within a soil vegetation atmosphere transfer model". *Agr. Forest Meteorol.*, 147, 125–139. DOI: 10.1016/j.agrformet.2007.07.006.

Heidkamp, M., A. Chlond, and F. Ament (2018). "Closing the energy balance using a canopy heat capacity and storage concept – a physically based approach for the land component JSBACHv3.11". *Geosci. Model Dev.*, 11.8, 3465–3479. DOI: 10.5194/gmd-11-3465-2018.

Hentgen, L., N. Ban, N. Kröner, D. Leutwyler, and C. Schär (2019). "Clouds in Convection-Resolving Climate Simulations Over Europe". *J. Geophys. Res.-Atmos.*, 124.7, 3849–3870. DOI: 10.1029/2018JD030150.

Herrington, L. P. (1969). "On Temperature and Heat Flow in Tree Stems". *Yale School of Forestry & Environmental Studies Bulletin Series*, 76.

Hicke, J. A., A. J. Meddens, and C. A. Kolden (2015). "Recent Tree Mortality in the Western United States from Bark Beetles and Forest Fires". *Forest Science*, 62.2, 141–153. DOI: 10.5849/forsci.15-086.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär (2009). "The Soil Moisture–Precipitation Feedback in Simulations with Explicit and Parameterized Convection". *Journal of Climate*, 22.19, 5003–5020. DOI: 10.1175/2009JCLI2604.1.

Houze Jr., R. A. (2012). "Orographic effects on precipitating clouds". *Rev. Geophys.*, 50.1. DOI: 10.1029/2011RG000365.

Hu, X., B. Huang, and F. Cherubini (2019). "Impacts of idealized land cover changes on climate extremes in Europe". *Ecol. Indic.*, 104, 626–635. DOI: 10.1016/j.ecolind.2019.05.037.

Huang, C., S. Kim, K. Song, J. R. Townshend, P. Davis, A. Altstatt, O. Rodas, A. Yanosky, R. Clay, C. J. Tucker, and J. Musinsky (2009). "Assessment of Paraguay's forest cover change using Landsat observations". *Global Planet. Change*, 67, 1–12. DOI: 10.1016/j.gloplacha.2008.12.009.

Huete, A. R., K. Didan, Y. E. Shimabukuro, P. Ratana, S. R. Saleska, L. R. Hutyra, W. Yang, R. R. Nemani, and R. Myneni (2006). "Amazon rainforests green-up with sunlight in dry season". *Geophys. Res. Lett.*, 33.L06405. DOI: 10.1029/2005GL025583.

Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall (2013). "The Community Earth System Model: A Framework for Collaborative Research". *B. Am. Meteorol. Soc.*, 94.9, 1339–1360. DOI: 10.1175/BAMS-D-12-00121.1.

Hurtt, G. C., 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 (2011). "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, 117. DOI: 10.1007/s10584-011-0153-2.

Hurtt, G. C., L. Chini, R. Sahajpal, S. Frolking, B. L. Bodirsky, K. Calvin, J. C. Doelman, J. Fisk, S. Fujimori, K. Klein Goldewijk, T. Hasegawa, P. Havlik, A. Heinimann, F. Humpenöder, J. Jungclaus, J. O. Kaplan, J. Kennedy, T. Krisztin, D. Lawrence, P. Lawrence, L. Ma, O. Mertz, J. Pongratz, A. Popp, B. Poulter, K. Riahi, E. Shevliakova, E. Stehfest, P. Thornton, F. N. Tubiello, D. P. van Vuuren, and X. Zhang (2020). "Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6". *Geosci. Model Dev.*, 13.11, 5425–5464. DOI: 10.5194/gmd-13-5425-2020.

Hurtt, G. C., S. Frolking, M. G. Fearon, B. Moore, E. Shevliakova, S. Malyshev, S. W. Pacala, and R. A. Houghton (2006). "The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands". *Glob. Change Biol.*, 12.7, 1208–1229. DOI: 10.1111/j.1365-2486.2006.01150.x.

IPCC, 2014: IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC.* Geneva, Switzerland.

IPCC 2018: Rogelj, J., D. Shindell, K. Jiang, S. Fifita, P. Forster, V. Ginzburg, C. Handa, H. Kheshgi, S. Kobayashi, E. Kriegler, L. Mundaca, R. Séférian, and M.V. Vilariño (2018). *Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press.*

IPCC 2019a: Shukla, P., J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. Robers, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferra, E. Hughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, and J. Malley (2019). *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*.

IPCC 2019b: Arneth, A., F. Denton, F. Agus, A. Elbehri, K. Erb, B. E. Osman, M. Rahimi, M. Rounsevell, A. Spence, and R. Valentini (2019). *Framing and Context. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosyste [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.*

IPCC 2019c: Jia, G., E. Shevlioakov, P. Artaxo, N. De Noblet-Ducoudré, R. Houghton, J. House, K. Kitajima, C. Lennard, A. Popp, A. Sirin, R. Sukumar, and L. Verchot (2019). *Landclimate interactions. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M, Belkacemi, J. Malley, (eds.)]. In press*.

IPCC 2019d: P.R. Shukla, J. Skea, R. Slade, R. van Diemen, E. Haughey, J. Malley, M. Pathak, J. Portugal Pereira (eds.) (2019). *Technical Summary. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M, Belkacemi, J. Malley, (eds.)]. In press.*

IPCC 2019e: IPCC (2019). *Summary for Policymakers. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.- O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.*

Jacob, D. et al. (2014). "EURO-CORDEX: new high-resolution climate change projections for European impact research". *Reg. Environ. Change*, 14, 563–578. DOI: 10.1007/s10113-013-0499-2.

Jacobson, A. P., J. M. Riggio, and A. Tait (2019). "Global areas of low human impact ('Low Impact Areas') and fragmentation of the natural world." *Sci. Rep.*, 9.14179. DOI: 10.1038/s41598-019-50558-6.

Jipp, P. H., D. C. Nepstad, D. K. Cassel, and C. R. de Carvalho (1998). "Deep Soil Moisture Storage and Transpiration in Forests and Pastures of Seasonally-dry Amazonia". *Climate Change*, 39.2-3, 395–412. DOI: 10.1023/A:1005308930871.

Juang, J.-Y., G. Katul, M. Siqueira, P. Stoy, and K. Novick (2007). "Separating the effects of albedo from eco-physiological changes on surface temperature along a successional chronosequence in the southeastern United States". *Geophys. Res. Lett.*, 34.L21408. DOI: 10.1029/2007GL031296.

Jung, M., M. Reichstein, H. A. Margolis, A. Cescatti, A. D. Richardson, M. A. Arain, A. Arneth, C. Bernhofer, D. Bonal, J. Chen, D. Gianelle, N. Gobron, G. Kiely, W. Kutsch, G. Lasslop, B. E. Law, A. Lindroth, L. Merbold, L. Montagnani, E. J. Moors, D. Papale, M. Sottocornola, F. Vaccari, and C. Williams (2011). "Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations". *J. Geophys. Res.-Biogeo.*, 116.G3. DOI: 10.1029/2010JG001566.

Kaplan, J. O., K. M. Krumhardt, E. C. Ellis, W. F. Ruddiman, C. Lemmen, and K. K. Goldewijk (2011). "Holocene carbon emissions as a result of anthropogenic land cover change". *The Holocene*, 21.5, 775–791. DOI: 10.1177/0959683610386983.

Kattge, J., W. Knorr, T. Raddatz, and C. Wirth (2009). "Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models". *Glob. Change Biol.*, 15.4, 976–991. DOI: 10.1111/j.1365-2486.2008.01744.x.

Khanna, J., D. Medvigy, S. Fueglistaler, and R. Walko (2017). "Regional dry-season climate changes due to three decades of Amazonian deforestation." *Nat. Clim. Change*, 7, 200–204. DOI: 10.1038/nclimate3226.

Kilinc, M., J. Beringer, L. B. Hutley, V. Haverd, and N. Tapper (2012). "An analysis of the surface energy budget above the world's tallest angiosperm forest". *Agr. Forest Meteorol.*, 166, 23–31. DOI: 10.1016/j.agrformet.2012.05.014.

Kim, H. (2017). *Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1)*. DOI: 10.20783/DIAS.501.

Kim, J. E., M. M. Laguë, S. Pennypacker, E. Dawson, and A. L. S. Swann (2020). "Evaporative Resistance is of Equal Importance as Surface Albedo in High-Latitude Surface Temperatures Due to Cloud Feedbacks". *Geophysical Research Letters*, 47.4. e2019GL085663 10.1029/2019GL085663, e2019GL085663. DOI: 10.1029/2019GL085663.

Kosztra, B., G. Büttner, G. Hazeu, and S. Arnold (2019). *Updated CLC illustrated nomenclature guidelines*. European Environment Agency.

Kotlarski, S., K. Keuler, O. B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen, D. Jacob, D. Lüthi, E. van Meijgaard, G. Nikulin, C. Schär, C. Teichmann, R. Vautard, K. Warrach-Sagi, and V. Wulfmeyer (2014). "Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble". *Geosci. Model Dev.*, 7.4, 1297–1333. DOI: 10.5194/gmd-7-1297-2014.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006). "World Map of the Köppen-Geiger climate classification updated". *Meteorol. Z.*, 15.3, 259–263. DOI: 10.1127/0941-2948/2006/0130.

Kumagai, T., H. Kanamori, and T. Yasunari (2013). "Deforestation-induced reduction in rainfall". *Hydrol. Process.*, 27.25, 3811–3814. DOI: 10.1002/hyp.10060.

Kumar, S., P. A. Dirmeyer, V. Merwade, T. DelSole, J. M. Adams, and D. Niyogi (2013). "Land use/cover change impacts in CMIP5 climate simulations: A new methodology and 21st century challenges". *J. Geophys. Res.*, 118, 6337–6353. DOI: 10.1002/jgrd.50463.

Laguë, M. M., G. B. Bonan, and A. L. S. Swann (2019). "Separating the Impact of Individual Land Surface Properties on the Terrestrial Surface Energy Budget in both the Coupled and Uncoupled Land–Atmosphere System". *J. Climate*, 32.18, 5725–5744. DOI: 10.1175/JCLI-D-18-0812.1.

Langanke, T. (2018). *Copernicus Land Monitoring Service: High Resolution Layer Forest: Product Specification Document*. Tech. rep. Copenhagen, Denmark.

Lawrence, D., R. Fisher, C. Koven, K. W. Oleson, S. C. Swenson, M. Vertenstein, B. Andre, G. B. Bonan, B. Ghimire, L. Van Kampenhout, D. Kennedy, E. Kluzek, R. Knox, P. Lawrence, F. Li, H. Li, D. Lombardozzi, Y. Lu, J. Perket, W. Riley, W. Sacks, M. Shi, W. Wieder, and C. Xu (2018). *Technical Description of version 5.0 of the Community Land Model (CLM)*. Boulder, Colorado.

Lawrence, D. M., R. A. Fisher, C. D. Koven, K. W. Oleson, S. C. Swenson, G. Bonan, N. Collier, and et al. (2019). "The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty". *J. Adv. Model Earth Sy.*, 11.12, 4245–4287. DOI: 10.1029/2018MS001583.

Lawrence, P. J. and T. N. Chase (2007). "Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0)". *J. Geophys. Res.-Biogeo.*, 112.G01023. DOI: 10.1029/2006JG000168.

Lee, X., M. L. Goulden, D. Y. Hollinger, A. Barr, T. A. Black, G. Bohrer, R. Bracho, B. Drake, A. Goldstein, L. Gu, G. Katul, T. Kolb, B. E. Law, L. H. Margolis, T. Meyers, R. Monson, W. Munger, R. Oren, K. T. Paw U, A. D. Richardson, R. Schmid H. P. Staebler, S. Wofsy, and L. Zhao (2011). "Observed increase in local cooling effect of deforestation at higher latitude". *Nature*, 479, 384–387. DOI: 10.1038/nature10588.

Lejeune, Q., E. L. Davin, G. Duveiller, B. Crezee, R. Meier, A. Cescatti, and S. I. Seneviratne (2020). "Biases in the albedo sensitivity to deforestation in CMIP5 models and their impacts on the associated historical Radiative Forcing". *Earth Syst. Dynam.*, 1–34. DOI: 10.5194/esd-2019-94.

Lejeune, Q., E. L. Davin, B. P. Guillod, and S. I. Seneviratne (2015). "Influence of Amazonian deforestation on the future evolution of regional surface fluxes, circulation, surface temperature and precipitation". *Clim. Dynam.*, 44, 2769–2786. DOI: 10.1007/s00382-014-2203-8.

Lejeune, Q., E. L. Davin, L. Gudmundsson, J. Winckler, and S. I. Seneviratne (2018). "Historical deforestation locally increased the intensity of hot days in northern mid-latitudes". *Nat. Clim. Change*, 8.5, 386–390. DOI: 10.1038/s41558-018-0131-z;10.1038/s41558-018-0131-z.

Lejeune, Q., S. I. Seneviratne, and E. L. Davin (2017). "Historical Land-Cover Change Impacts on Climate: Comparative Assessment of LUCID and CMIP5 Multimodel Experiments". *J. Climate*, 30, 1439–1459. DOI: 10.1175/JCLI-D-16-0213.1.

Lewis, E., H. Fowler, L. Alexander, R. Dunn, F. McClean, R. Barbero, S. Guerreiro, X.-F. Li, and S. Blenkinsop (2019). "GSDR: A Global Sub-Daily Rainfall Dataset". *J. Climate*, 32.15, 4715–4729. DOI: 10.1175/JCLI-D-18-0143.1.

Li, X., H. Chen, J. Wei, W. Hua, S. Sun, H. Ma, X. Li, and J. Li (2018). "Inconsistent Responses of Hot Extremes to Historical Land Use and Cover Change Among the Selected CMIP5 Models". *Journal of Geophysical Research: Atmospheres*, 123.7, 3497–3512. DOI: 10.1002/2017JD028161.

Li, Y., N. De Noblet-Decoudré, E. L. Davin, S. Motesharrei, N. Zeng, L. Shuangcheng, and E. Kalnay (2016a). "The role of spatial scale and background climate in the latitudinal temperature response to deforestation". *Earth Syst. Dynam.*, 7, 167–181. DOI: 10.5194/esd-7-167-2016.

Li, Y., M. Zhao, S. Motesharrei, Q. Mu, E. Kalnay, and S. Li (2015). "Local cooling and warming effects of forests based on satellite observations". *Nat. Commun.*, 6.6603. DOI: 10.1038/ncomms7603.

Li, Y., M. Zhao, D. J. Mildrexler, S. Motesharrei, Q. Mu, E. Kalnay, F. Zhao, S. Li, and K. Wang (2016b). "Potential and Actual impacts of deforestation and afforestation on land surface temperature". *J. Geophys. Res.*, 121.24, 14372–14386. DOI: 10.1002/2016JD024969.

Li, Z. and S. N. Wood (2020). "Faster model matrix crossproducts for large generalized linear models with discretized covariates". *Stat. Comput.*, 30, 19–25. DOI: 10.1007/s11222-019-09864-2.

Liao, W., A. J. Rigden, and D. Li (2018). "Attribution of Local Temperature Response to Deforestation". *J. Geophys. Res.-Biogeo.*, 123.5, 1572–1587. DOI: 10.1029/2018JG004401.

Lindroth, A., M. Mölder, and F. Lagergren (2010). "Heat storage in forest biomass improves energy balance closure". *Biogeoscience*, 7.1. DOI: 10.5194/bg-7-301-2010.

Liu, H., J. T. Randerson, J. Lindfors, and F. S. Chapin III (2005). "Changes in the surface energy budget after fire in boreal ecosystems of interior Alaska: An annual perspective". *J. Geophys. Res.*, 110.D13101. DOI: 10.1029/2004JD005158.

Liu, J. and D. Niyogi (2019). "Meta-analysis of urbanization impact on rainfall modification". *Sci. Rep.*, 9.7301. DOI: 10.1038/s41598-019-42494-2.

Lu, N. (2019). "Scale effects of topographic ruggedness on precipitation over Qinghai-Tibet Plateau". *Atmos. Sci. Lett.*, 20.6, e904. DOI: 10.1002/asl.904.

Lugato, E., A. Cescatti, A. Jones, G. Ceccherini, and G. Duveiller (2020). "Maximising climate mitigation potential by carbon and radiative agricultural land management with cover crops". *Environ. Res. Lett.*, DOI: 10.1088/1748-9326/aba137.

Luyssaert, S., M. Jammet, P. C. Stoy, S. Estel, J. Pongratz, E. Ceschia, G. Churkina, A. Don, K.-H. Erb, M. Ferlicoq, B. Gielen, T. Grünwald, R. A. Houghton, K. K., A. Knohl, T. Kolb, T. Kuemmerle, T. Laurila, A. Lohila, D. Loustau, M. J. McGrath, P. Meyfroidt, E. J. Moors, K. Naudts, K. Novick, J. Otto, K. Pilegaard, C. A. Pio, S. Rambal, R. C., J. Ryder, A. E. Suyker, A. Varlagin, M. Wattenback, and A. J. Dolman (2014). "Land management and land-cover change have impacts of similar magnitude on surface temperature". *Nat. Clim. Change*, 4, 389–393. DOI: 10.1038/nclimate2196.

Luyssaert, S., G. Marie, A. Valade, Y. Chen, S. N. Dojomo, J. Ryder, J. Otto, K. Naudts, A. S. Lanso, J. Ghattas, and M. J. McGrath (2018). "Trade-offs in using European forests to meet climate objectives". *Nature*, 562, 259–262. DOI: 10.1038/s41586-018-0577-1.

Ma, L., G. C. Hurtt, L. P. Chini, R. Sahajpal, J. Pongratz, S. Frolking, E. Stehfest, K. Klein Goldewijk, D. O'Leary, and J. C. Doelman (2020). "Global rules for translating land-use change (LUH2) to land-cover change for CMIP6 using GLM2". *Geosci. Model Dev.*, 13.7, 3203–3220. DOI: 10.5194/gmd-13-3203-2020.

Malyshev, S., E. Shevliakova, R. J. Stouffer, and S. W. Pacala (2015). "Contrasting Local versus Regional Effects of Land-Use-Change-Induced Heterogeneity on Historical Climate: Analysis with the GFDL Earth System Model". *J. Climate*, 28, 5448–5469. DOI: 10.1175/JCLI-D-14-00586.1.

Maness, H., P. Kushner, and I. Fung (2013). "Summertime climate response to mountain pine beetle disturbance in British Columbia". *Nat. Geosci.*, 6, 65–70. DOI: 10.1038/ngeo1642.

Margono, B. A., P. V. Potapov, S. A. Turubanova, F. Stolle, and M. C. Hansen (2014). "Primary forest cover loss in Indonesia over 2000–2012". *Nat. Clim. Change*, 4, 730–735. DOI: 10.1038/nclimate2277.

Martens, B. and D. G. Miralles (2017). *Global Land Evaporation Amsterdam Model*.

Martens, B., D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. C. Verhoest (2017). "GLEAM v3: satellite-based land evaporation and root-zone soil moisture". *Geosci. Model Dev.*, 10.5. DOI: 10.5194/gmd-2016-162.

McCaughey, J. and W. Saxton (1988). "Energy balance storage terms in a mixed forest". *Agr. Forest Meteorol.*, 44.1, 1–18. DOI: 10.1016/0168-1923(88)90029-9.

McGrath, M. J., S. Luyssaert, P. Meyfroidt, J. O. Kaplan, M. Bürgi, Y. Chen, K. Erb, U. Gimmi, D. McInerney, K. Naudts, J. Otto, F. Pasztor, J. Ryder, M.-J. Schelhaas, and A. Valade (2015). "Reconstructing European forest management from 1600 to 2010". *Biogeosciences*, 12.14, 4291–4316. DOI: 10.5194/bg-12-4291-2015.

Meesters, A. and H. Vugts (1996). "Calculation of heat storage in stems". *Agr. Forest Meteorol.*, 78.3, 181–202. DOI: 10.1016/0168-1923(95)02251-1.

Meier, R., E. L. Davin, Q. Lejeune, M. Hauser, Y. Li, B. Martens, N. M. Schultz, S. Sterling, and W. Thiery (2018). "Evaluating and improving the Community Land Model's sensitivity to land cover". *Biogeosciences*, 15.15, 4731–4757. DOI: 10.5194/bg-15-4731-2018.

Meier, R., J. Schwaab, S. I. Seneviratne, M. Sprenger, E. Lewis, and E. L. Davin (2020). "Empirical estimate of forestation-induced precipitation changes in Europe". *Nat. Geosci.*,

Meier, R., E. L. Davin, S. C. Swenson, D. M. Lawrence, and J. Schwaab (2019). "Biomass heat storage dampens diurnal temperature variations in forests". *Environ. Res. Lett.*, 14.8, 084026. DOI: 10.1088/1748-9326/ab2b4e.

Meiyappan, P. and A. K. Jain (2012). "Three distinct global estimates of historical land-cover change and land-use conversions for over 200 years". *Front. Earth Sci.*, 6, 122–139. DOI: 10.1007/s11707-012-0314-2.

Menne, M. J., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, Y. X., S. Anthony, R. Ray, V. R. S., G. B. E., and T. G. Houston (2012a). *Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.20*. DOI: 10.7289/V5D21VHZ.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012b). "An Overview of the Global Historical Climatology Network-Daily Database". *Journal of Atmospheric and Oceanic Technology*, 29.7, 897–910. DOI: 10.1175/JTECH-D-11-00103.1.

Meyers, T. P. and S. E. Hollinger (2004). "An assessment of storage terms in the surface energy balance of maize and soybean". *Agr. Forest Meteorol.*, 125.1, 105–115. DOI: 10.1016/j.agrformet.2004.03.001.

Michel, D., C. Jiménez, D. G. Miralles, M. Jung, M. Hirschi, A. Ershadi, B. Martens, M. F. McCabe, J. B. Fisher, Q. Mu, S. I. Seneviratne, E. F. Wood, and D. Fernández-Prieto (2016). "The WACMOS-ET project - Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms". *Hydrol. Earth Syst. Sc.*, 20.2, 803–822. DOI: 10.5194/hess-20-803-2016.

Michiles, A. A. D. and R. Gielow (2008). "Above-ground thermal energy storage rates, trunk heat fluxes and surface energy balance in a central Amazonian rainforest". *Agr. Forest Meteorol.*, 148.6, 917–930. DOI: 10.1016/j.agrformet.2008.01.001.

Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2011). "Global land-surface evaporation estimated from satellite-based observations". *Hydrol. Earth Syst. Sc.*, 15.2, 454–469. DOI: 10.5194/hess-15-453-2011.

Miralles, D. G., C. Jiménez, M. Jung, D. Michel, A. Ershadi, M. F. McCabe, M. Hirschi, B. Martens, A. J. Dolman, J. B. Fisher, Q. Mu, S. I. Seneviratne, E. F. Wood, and D. Fernández-Prieto (2016). "The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets". *Hydrol. Earth Syst. Sc.*, 20.2, 823–842. DOI: 10.5194/hess-20-823-2016.

Moore, C. and G. Fisch (1986). "Estimating heat storage in Amazonian tropical forest". *Agr. Forest Meteorol.*, 38.1, 147–168. DOI: 10.1016/0168-1923(86)90055-9.

Mu, Q. Z., M. S. Zhao, and S. W. Running (2011). "Improvements to a MODIS global terrestrial evapotranspiration algorithm". *Remote Sens. Environ.*, 115, 1781–1800. DOI: 10.1016/j.rse.2011.02.019.

Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011). "Evaluation of global boservations-based evapotranspiration datasets and IPCC AR4 simulations". *Geophys. Res. Lett.*, 38.L06402. DOI: 10.1029/2010GL046230.

Mueller, B., M. Hirschi, C. Jimenez, P. Ciais, P. A. Dirmeyer, A. J. Dolman, J. B. Fisher, M. Jung, F. Ludwig, F. Maignan, D. G. Miralles, M. F. McCabe, M. Reichstein, J. Sheffield, K. Wang, E. F. Wood, Y. Zhang, and S. I. Seneviratne (2013). "Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis". *Hydrol. Earth Syst. Sc.*, 17, 3707–3720. DOI: 10.5194/hess-17-3707-2013.

Naudts, K., Y. Chen, M. J. McGrath, J. Ryder, A. Valade, J. Otto, and S. Luyssaert (2016). "Europe's forest management did not mitigate climate warming". *Science*, 351.6273, 597–600. DOI: 10.1126/science.aad7270.

Navarro, L. M. and H. M. Pereira (2015). "Rewilding Abandoned Landscapes in Europe". *Rewilding European Landscapes*. Cham: Springer International Publishing, 3–23. DOI: 10.1007/978-3-319-12039-3_1.

Oleson, K. W., G. B. Bonan, S. Levis, and M. Vertenstein (2004). "Effects of land use change on North American climate: impact of surface datasets and model biogeophysics". *Clim. Dynam.*, 23.2, 117–132. DOI: 10.1007/s00382-004-0426-9.

Oleson, K. W., D. Lawrence, G. B. Bonan, B. Drewniak, M. Huang, C. Koven, S. Levis, F. Li, W. Riley, Z. M. Subin, S. C. Swenson, P. E. Thornton, A. Bozbiyik, R. Fisher, C. Heald, E. Kluzek, J.-F. Lamarque, P. Lawrence, L. Leung, and Z.-L. Yang (2013). *Technical Description of version 4.5 fo the Community Land Model (CLM)*. Boulder, CO: National Center for Atmospheric Research. DOI: 10.5065/D6RR1W7M.

Oliphant, A., C. Grimmond, H. Zutter, H. Schmid, H.-B. Su, S. Scott, B. Offerle, J. Randolph, and J. Ehman (2004). "Heat storage and energy balance fluxes for a temperate deciduous forest". *Agr. Forest Meteorol.*, 126.3, 185–201. DOI: 10.1016/j.agrformet.2004.07.003.

Padrón, R. S., L. Gudmundsson, D. Michel, and S. I. Seneviratne (2020). "Terrestrial water loss at night: global relevance from observations and climate models". *Hydrol. Earth Syst. Sc.*, 24.2, 793–807. DOI: 10.5194/hess-24-793-2020.

Padrón, R. S., L. Gudmundsson, P. Greve, and S. I. Seneviratne (2017). "Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis". *Water Resour. Res.*, 53.11, 9659–9678. DOI: 10.1002/2017WR021215.

Panosetti, D., L. Schlemmer, and C. Schär (2020). "Convergence behavior of idealized convection-resolving simulations of summertime deep moist convection over land". *Clim. Dynam.*, 55, 215–234. DOI: 10.1007/s00382-018-4229-9.

Peng, S.-S., S. Piao, Z. Zeng, P. Ciais, L. Zhou, L. Z. X. Li, R. B. Myneni, and H. Zeng (2014). "Afforestation in China cools local land surface temperature". *P. Natl. Acad. Sci. USA*, 111, 2915–2919. DOI: 10.1073/pnas.1315126111.

Perugini, L., L. Caporaso, S. Marconi, A. Cescatti, B. Quesada, N. de Noblet-Ducoudré, J. I. House, and A. Arneth (2017). "Biophysical effects on temperature and precipitation due to land cover change". *Environ. Res. Lett.*, 12.5, 053002. DOI: 10.1088/1748-9326/aa6b3f.

Pitman, A. J., N. de Noblet-Ducoudré, F. T. Cruz, E. L. Davin, G. B. Bonan, V. Brovkin, M. Claussen, C. Delire, L. Ganzeveld, V. Gayler, B. J. J. M. van den Hurk, P. J. Lawrence, M. K. van der Molen, C. Müller, C. H. Reick, S. I. Seneviratne, B. J. Strengers, and A. Voldoire (2009). "Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study". *Geophys. Res. Lett.*, 36.L14814. DOI: 10.1029/2009GL039076.

Pollock, M. D., G. O'Donnell, P. Quinn, M. Dutton, A. Black, M. E. Wilkinson, M. Colli, M. Stagnaro, L. G. Lanza, E. Lewis, C. G. Kilsby, and P. E. O'Connell (2018). "Quantifying and Mitigating Wind-Induced Undercatch in Rainfall Measurements". *Water Resour. Res.*, 54.6, 3863–3875. DOI: 10.1029/2017WR022421.

Pongratz, J., C. H. Reick, T. Raddatz, and M. Claussen (2009). "Effects of anthropogenic land cover change on the carbon cycle of the last millennium". *Global Biogeochem. Cy.*, 23.4. DOI: 10.1029/2009GB003488.

— (2010). "Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change". *Geophys.Res. Lett.*, 37.L08702. DOI: 10.1029/2010GL043010.

Pongratz, J., C. Reick, T. Raddatz, and M. Claussen (2008). "A reconstruction of global agricultural areas and land cover for the last millennium". *Global Biogeochem. Cy.*, 22.3. DOI: 10.1029/2007GB003153.

Pongratz, J., H. Dolman, A. Don, K.-H. Erb, R. Fuchs, M. Herold, C. Jones, T. Kuemmerle, S. Luyssaert, P. Meyfroidt, and K. Naudts (2018). "Models meet data: Challenges and opportunities in implementing land management in Earth system models". *Glob. Change Biol.*, 24.4, 1470–1487. DOI: 10.1111/gcb.13988.

Prein, A. F., W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli, N. P. M. van Lipzig, and R. Leung (2015). "A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges". *Rev. Geophys.*, 53.2, 323–361. DOI: 10.1002/2014RG000475.

Prein, A. F., A. Gobiet, H. Truhetz, et al. (2016). "Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits?" *Clim. Dynam.*, 46, 383–412. DOI: 10.1007/s00382-015-2589-y.

Priestly, C. H. B. and R. J. Taylor (1972). "On the assessment of surface heat flux and evaporation using large-scale parameters". *Mon. Weather Rev.*, 100, 81–92.

Ramankutty, N. and J. A. Foley (1999). "Estimating historical changes in global land cover: Croplands from 1700 to 1992". *Global Biogeochemical Cycles*, 13.4, 997–1027. DOI: 10.1029/1999GB900046.

Randerson, J. T., H. Liu, M. G. Flanner, S. D. Chambers, Y. Jin, P. G. Hess, G. Pfister, M. C. Mack, K. K. Treseder, L. R. Welp, F. S. Chapin, J. W. Harden, M. L. Goulden, E. Lyons, J. C. Neff, E. A. G. Schuur, and C. S. Zender (2006). "The Impact of Boreal Forest Fire on Climate Warming". *Science*, 314.5802, 1130–1132. DOI: 10.1126/science.1132075.

Rigden, A. J. and D. Li (2017). "Attribution of surface temperature anomalies induced by land use and land cover changes". *Geophys. Res. Lett.*, 44.13, 6814–6822. DOI: 10.1002/2017GL073811.

Rizwan, A. M., L. Y. Dennis, and C. Liu (2008). "A review on the generation, determination and mitigation of Urban Heat Island". *J. Environ. Sci.*, 20.1, 120–128. DOI: 10.1016/S1001-0742(08)60019-4.

Robertson, E. (2019). "The Local Biophysical Response to Land-Use Change in HadGEM2-ES". *J. Climate*, 32.22, 7611–7627. DOI: 10.1175/JCLI-D-18-0738.1.

Rüdisühli, S., M. Sprenger, D. Leutwyler, C. Schär, and H. Wernli (2020). "Attribution of precipitation to cyclones and fronts over Europe in a kilometer-scale regional climate simulation". *Weather and Climate Dynamics Discussions*, 2020, 1–38. DOI: 10.5194/wcd-2020-18.

Safranyik, L., A. Carroll, J. Regniere, D. Langor, W. Riel, T. Shore, B. Peter, B. Cooke, V. Nealis, and S. Taylor (2010). "Potential for Range Expansion of Mountain Pine Beetle into the Boreal Forest of North America". *Can. Entomol.*, 142, 415–442. DOI: 10.4039/n08-CPA01.

Saleska, S. R., K. Didan, A. R. Huete, and H. R. da Rocha (2007). "Amazon Forests Green-Up During 2005 Drought". *Science*, 318.5850, 612. DOI: 10.1126/science.1146663.

Samuelsson, P., C. G. Jones, U. Will ́En, A. Ullerstig, S. Gollvik, U. Hansson, E. Jansson, C. Kjellstro ̈M, G. Nikulin, and K. Wyser (2011). "The Rossby Centre Regional Climate model RCA3: model description and performance". *Tellus A*, 63.1, 4–23. DOI: 10.1111/j.1600-0870.2010.00478.x.

Sandel, B. and J. Svenning (2013). "Human impacts drive a global topographic signature in tree cover". *Nat. Commun.*, 4.2474. DOI: 10.1038/ncomms3474.

Schultz, N. M., P. J. Lawrence, and X. Lee (2017). "Global satellite data highlights the diurnal asymmetry of the surface temperature response to deforestation". *J. Geophys. Res.-Biogeo.*, 122.4, 903–917. DOI: 10.1002/2016JG003653.

Schultz, N. M., X. Lee, P. J. Lawrence, D. M. Lawrence, and L. Zhao (2016). "Assessing the use of subgrid land model output to study impacts of land cover change". *J. Geophys. Res.*, 121, 6133–6147. DOI: 10.1002/2016JD025094.

Schwaab, J., E. L. Davin, P. Bebi, A. Duguay-Tetzlaff, L. Waser, and R. Meier (2020). "Increasing the broad-leaved tree fraction in European forests mitigates hot temperature extremes". *Sci. Rep.*, 10.14153. DOI: 10.1038/s41598-020-71055-1.

Schwarz, C. F., E. C. Thor, and G. H. Elsner (1976). *Wildland planning glossary*.

Seneviratne, S. I., S. J. Phipps, A. J. Pitman, A. L. Hirsch, E. L. Davin, M. G. Donat, M. Hirschi, A. Lenton, M. Wilhelm, and B. Kravitz (2018). "Land radiative management as contributor to regional-scale climate adaptation and mitigation". *Nat. Geosci.*, 11, 88–96. DOI: 10.1038/s41561-017-0057-5.

Sheffield, J. and E. F. Wood (2008). "Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations." *Clim. Dynam.*, 31, 79–105. DOI: 10.1007/s00382-007-0340-z.

Siebert, S., V. Henrich, K. Frenken, and J. Burke (2013). *Global Map of Irrigation Areas version 5*. Rheinische Friedrich-Wilhelms-University, Bonn, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy.

Silberstein, R., A. Held, T. Hatton, N. Viney, and M. Sivapalan (2001). "Energy balance of a natural jarrah (Eucalyptus marginata) forest in Western Australia: measurements during the spring and summer". *Agr. Forest Meteorol.*, 109.2, 79–104. DOI: 10.1016/S0168-1923(01)00263-5.

Skinner, C. B., C. J. Poulson, and J. S. Mankin (2018). "Amplification of heat extremes by plant CO_2 physiological forcing". *Nat. Commun.*, 9.1094. DOI: 10.1038/s41467-018-03472-w.

Smith, A., N. Lott, and R. Vose (2011). "The Integrated Surface Database: Recent Developments and Partnerships". *B. Am. Meteorol. Soc.*, 92.6, 704–708. DOI: 10.1175/2011BAMS3015.1.

Sodemann, H., C. Schwierz, and H. Wernli (2008). "Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation influence". *J. Geophys. Res.-Atmos.*, 113.D3. DOI: 10.1029/2007JD008503.

Sonntag, S., J. Pongratz, C. H. Reick, and H. Schmidt (2016). "Reforestation in a high-CO2 world—Higher mitigation potential than expected, lower adaptation potential than hoped for". *Geophys. Res. Lett.*, 43, 6546–6553. DOI: 10.1002/2016GL068824.

Spracklen, D., S. Arnold, and C. Taylor (2012). "Observations of increased tropical rainfall preceded by air passage over forests". *Nature*, 489, 282–285. DOI: 10.1038/nature11390.

Sprenger, M. and H. Wernli (2015). "The LAGRANTO Lagrangian analysis tool - version 2.0". *Geosci. Model Dev.*, 8.8, 2569–2586. DOI: 10.5194/gmd-8-2569-2015.

Stoy, P., G. Katul, M. Siqueira, J.-Y. Juang, H. Mccarthy, A. Oishi, J. M. Uebelherr, H.-S. Kim, and R. Oren (2006). "Separating the Effects of Climate and Vegetation on Evapotranspiration along a Successional Chronosequence in the Southeastern US". *Glob. Change Biol.*, 12, 2115–2135. DOI: 10.1111/j.1365-2486.2006.01244.x.

Sud, Y. C., W. K.-M. Lau, G. K. Walker, J.-H. Kim, G. E. Liston, and P. J. Sellers (1996). "Biogeophysical Consequences of a Tropical Deforestation Scenario: A GCM Simulation Study". *J. Climate*, 9.12, 3225–3247. DOI: 10.1175/1520-0442(1996)009<3225:BCOATD>2.0.CO;2.

Swamy, S. L. and V. P. Tewari (2018). "Mitigation and Adaptation Strategies to Climate Change Through Agroforestry Practices in the Tropics". *In: Dagar J., Tewari V. (eds) Agroforestry*, DOI: 10.1007/978-981-10-7650-3_29.

Swenson, S. C., S. P. Burns, and D. M. Lawrence (2019). "The Impact of Biomass Heat Storage on the Canopy Energy Balance and Atmospheric Stability in the Community Land Model". *J. Adv.in Modeling Earth Sy.*, 11.1, 83–98. DOI: 10.1029/2018MS001476.

Tang, B., X. Zhao, and W. Zhao (2018). "Local Effects of Forests on Temperatures across Europe". *Remote Sens.*, 10.529. DOI: 10.3390/rs10040529.

Teuling, A. J., E. A. G. de Badts, F. A. Jansen, R. Fuchs, J. Buitink, A. J. Hoek van Dijke, and S. M. Sterling (2019). "Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe". *Hydrol. Earth Syst. Sc.*, 23.9, 3631–3652. DOI: 10.5194/hess-23-3631-2019.

Teuling, A. J., C. M. Taylor, J. Fokke Meirink, L. A. Melsen, D. G. Miralles, C. C. Van Heerwarden, R. Vautard, A. I. Stegehuis, G.-J. Nabuurs, and J. V.-G. de Arellano (2017). "Observational evidence for cloud cover enhancement over western European forests". *Nat. Commun.*, 8. DOI: 10.1038/ncomms14065.

Teuling, A. T., S. I. Seneviratne, R. Stöckli, M. Reichstein, E. Moors, P. Ciais, S. Luyssaert, B. van den Hurk, C. Ammann, C. Bernhofer, E. Dellwik, D. Gianelle, B. Gielen, T. Grünwald, K. Klumpp, L. Montagnani, C. Moureaux, M. Scottocornola, and G. Wohlfahrt (2010). "Contrasting response of European forest and grassland energy exchange ot heatwaves". *Nat. Geosci.*, 3, 722–727. DOI: 10.1038/ngeo950.

Teuling, A. J. (2018). "A Forest Evapotranspiration Paradox Investigated Using Lysimeter Data". *Vadose Zone J.*, 17.1, 170031. DOI: 10.2136/vzj2017.01.0031.

Thiery, W., E. L. Davin, D. M. Lawrence, A. L. Hirsch, M. Hauser, and S. I. Seneviratne (2017). "Present-day irrigation mitigates heat extremes". *J. Geophys. Res.*, 122.3, 1403–1422. DOI: 10.1002/2016JD025740.

Tölle, M. H., M. Breil, K. Radtke, and H.-J. Panitz (2018). "Sensitivity of European Temperature to Albedo Parameterization in the Regional Climate Model COSMO-CLM Linked to Extreme Land Use Changes". *Front. Environ. Sci.*, 6, 123. DOI: 10.3389/fenvs.2018.00123.

Tölle, M. H., O. Gutjahr, G. Busch, and J. C. Thiele (2014). "Increasing bioenergy production on arable land: Does the regional and local climate respond? Germany as a case study". *J. Geophys. Res.-Atmos.*, 119.6, 2711–2724. DOI: 10.1002/2013JD020877.

Trabucco, A., R. J. Zomer, D. A. Bossio, O. [Straaten], and L. V. Verchot (2008). "Climate change mitigation through afforestation/reforestation: A global analysis of hydrologic impacts with four case studies". *Agr. Ecosyst. Environ.*, 126.1. International Agricultural Research and Climate Change: A Focus on Tropical Systems, 81–97. DOI: 10.1016/j.agee.2008.01.015.

Trigo, I. F., L. F. Peres, C. C. DaCamara, and S. C. Freitas (2008). "Thermal Land Surface Emissivity Retrieved From SEVIRI/Meteosat". *IEEE T. Geosci. Remote*, 46.2, 307–315. DOI: 10.1109/TGRS.2007.905197.

Tyukavina, A., A. Baccini, M. C. Hansen, P. V. Potapov, S. V. Stehman, R. A. Houghton, A. M. Krylov, S. Turubanova, and S. J. Goetz (2015). "Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012". *Environ. Res. Lett.*, 10.7, 074002. DOI: 10.1088/1748-9326/10/7/074002.

Tzavali, A., J. P. Paravantis, G. Mihalakakou, A. Fotiadi, and E. Stigka (2015). "Urban heat island intensity: A literature review". *Fresenius Environ. Bull.*, 24.12b.

Van der Ent, R. J. and H. H. G. Savenije (2011). "Length and time scales of atmospheric moisture recycling". *Atmospheric Chemistry and Physics*, 11.5, 1853–1863. DOI: 10.5194/acp-11-1853-2011.

Van der Molen, M., A. Dolman, M. Waterloo, and L. Bruijnzeel (2006). "Climate is affected more by maritime than by continental land use change: A multiple scale analysis". *Global Planet. Change*, 54.1, 128–149. DOI: 10.1016/j.gloplacha.2006.05.005.

Van Heerwarden, C. and A. T. Teuling (2014). "Disentangling the response of forest and grassland energy exchange to heatwaves under idealized land-atmopshere coupling". *Biogeoscience*, 11.21. DOI: 10.5194/bg-11-6159-2014.

Vanden Broucke, S., S. Luyssaert, E. L. Davin, I. Janssens, and N. van Lipzig (2015). "New insights in the capability of climate models to simulate the impact of LUC based on temperature decomposition of paired site observations". *J. Geophys. Res.-Atmos.*, 120, 5417–5436. DOI: 10.1002/2015JD023095.

Vanden Broucke, S. and N. P. M. Van Lipzig (2017). "Do convection-permitting models improve the representation of the impact of LUC?" *Clim. Dynam.*, 49, 2749–2763. DOI: 10.1007/s00382-016-3489-5.

Vanderhoof, M., C. A. Williams, B. Ghimire, and J. Rogan (2013). "Impact of mountain pine beetle outbreaks on forest albedo and radiative forcing, as derived from Moderate Resolution Imaging Spectroradiometer, Rocky Mountains, USA". *J. Geophys. Res.-Biogeo.*, 118.4, 1461–1471. DOI: 10.1002/jgrg.20120.

Vanderhoof, M. K. and C. A. Williams (2015). "Persistence of MODIS evapotranspiration impacts from mountain pine beetle outbreaks in lodgepole pine forests, south-central Rocky Mountains". *Agr. Forest Meteorol.*, 200, 78–91. DOI: 10.1016/j.agrformet.2014.09.015.

Verseghy, D. L., N. A. McFarlane, and M. Lazare (1993). "Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs". *Int. J. Climatol.*, 13.4, 347–370. DOI: 10.1002/joc.3370130402.

Vivoy, N. (2009). *CRUNCEP dataset*.

Vogt, R., C. Bernhofer, L. W. Gay, L. Jaeger, and E. Parlow (1996). "The available energy over a Scots Pine plantation: What's up for partitioning?" *Theor. Appl. Climatol.*, 53.1, 23–31. DOI: 10.1007/BF00866408.

von Randow, C., A. O. Manzi, B. Kruijt, P. J. de Oliveira, F. B. Zanchi, R. L. Silva, M. G. Hodnett, J. H. C. Gash, J. A. Elbers, M. J. Waterloo, F. L. Cardoso, and P. Kabat (2004). "Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in South West Amazonia". *Theor. Appl. Climatol.*, 78, 5–26. DOI: 10.1007/s00704-004-0041-z.

Wang, Z., C. B. Schaaf, A. H. Strahler, M. J. Chopping, M. O. Román, Y. Shuai, C. E. Woodcock, D. Y. Hollinger, and D. R. Fitzjarrald (2014). "Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods". *Remote Sens. Environ.*, 140, 60–77. DOI: 10.1016/j.rse.2013.08.025.

Wan, Z. (2014). "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product". *Remote Sensing of Environment*, 140, 36–45. DOI: 10.1016/j.rse.2013.08.027.

Wan, Z. and G. Hulley (2015a). "MYD11C2 MODIS/Aqua Land Surface Temperature/Emissivity 8-Day L3 Global 0.05Deg CMG V006 [Data set]". DOI: 10.5067/MODIS/MYD11C2.006.

— (2015b). "MYD11C3 MODIS/Aqua Land Surface Temperature/Emissivity Monthly L3 Global 0.05Deg CMG V006 [Data set]". DOI: 10.5067/MODIS/MYD11C3.006.

Wang, K. and R. E. Dickinson (2012). "A Review of Global Terrrestrial Evapotranspiration: Observation, Modeling, Climatology, and Climatic Variablility". *Rev. Geophys.*, 50.RG2005. DOI: 10.1029/2011RG000373.

Wang, L., F. Tian, X. Wang, Y. Yang, and Z. Wei (2020). "Attribution of the land surface temperature response to land-use conversions from bare land". *Global Planet. Change*, 193, 103268. DOI: 10.1016/j.gloplacha.2020.103268.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp (1997). "The Resolution Dependence of Explicitly Modeled Convective Systems". *Mon. Weather Rev.*, 125.4, 527–548. DOI: 10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2.

Wernli, B. H. and H. C. Davies (1997). "A lagrangian-based analysis of extratropical cyclones. I: The method and some applications". *Q. J. Roy. Meteor. Soc.*, 123.538, 467–489. DOI: 10.1002/qj.49712353811.

Williams, C. A., M. Reichstein, N. Buchmann, D. Baldocchi, C. Beer, C. Schwalm, G. Wohlfahrt, N. Hasler, C. Bernhofer, T. Foken, D. Papale, S. Schymanski, and K. Schaefer (2012a). "Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers". *Water Resour. Res.*, 48.6. DOI: 10.1029/2011WR011586.

Williams, C. A., G. J. Collatz, J. Masek, and S. N. Goward (2012b). "Carbon consequences of forest disturbance and recovery across the conterminous United States". *Global Biogeochem. Cy.*, 26.1. DOI: 10.1029/2010GB003947.

Winckler, J., C. H. Reick, R. M. Bright, and J. Pongratz (2019a). "Importance of Surface Roughness for the Local Biogeophysical Effects of Deforestation". *J. Geophys. Res.-Atmos.*, 124.15, 8605–8618. DOI: 10.1029/2018JD030127.

Winckler, J., C. H. Reick, S. Luyssaert, A. Cescatti, P. C. Stoy, Q. Lejeune, T. Raddatz, A. Chlond, M. Heidkamp, and J. Pongratz (2019b). "Different response of surface temperature and air temperature to deforestation in climate models". *Earth Syst. Dynam.*, 10.3, 473–484. DOI: 10.5194/esd-10-473-2019.

Winckler, J., C. H. Reick, and J. Pongratz (2017). "Robust Identification of Local Biogeophysical Effects of Land-Cover Change in a Global Climate Model". *J. Climate*, 30.3, 1159–1176. DOI: 10.1175/JCLI-D-16-0067.1.

Winckler, J., Q. Lejeune, C. H. Reick, and J. Pongratz (2019c). "Nonlocal Effects Dominate the Global Mean Surface Temperature Response to the Biogeophysical Effects of Deforestation". *Geophys. Res. Lett.*, 46.2, 745–755. DOI: 10.1029/2018GL080211.

Wolf, S., W. Eugster, S. Majorek, and N. Buchmann (2011). "Afforestation of Tropical Pasture Only Marginally Affects Ecosystem-Scale Evapotranspiration". *Ecosystems*, 14.8, 1264–1275. DOI: 10.1007/s10021-011-9478-y.

Wood, S. N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". *J. R. Stat. Soc.*, 73.1, 3–36. DOI: 10.1111/j.1467-9868.2010.00749.x.

Wood, S. N. (2017). *Generalized additive models : an introduction with R*. Second edition. Chapman & Hall/CRC texts in statistical science. Boca Raton: CRC Press/Taylor & Francis Group.

Wood, S. N., Z. Li, G. Shaddick, and N. H. Augustin (2017). "Generalized Additive Models for Gigadata: Modeling the U.K. Black Smoke Network Daily Data". *J. Am. Stat. Assoc.*, 112.519, 1199–1210. DOI: 10.1080/01621459.2016.1195744.

World Urbanization Prospects (2018). *World Urbanization Prospects: The 2018 Revision*. Tech. rep. United Nations, Departement of Economic and Social Affairs.

Yosef, G., R. Walko, R. Avisar, F. Tatarinov, E. Rotenberg, and D. Yakir (2018). "Large-scale semi-arid afforestation can enhance precipitation and carbon sequestration potential". *Sci. Rep.*, 8.996. DOI: 10.1038/s41598-018-19265-6.

Zeng, X. (2001). "Global Vegetation Root Distribution for Land Modeling". *J. Hydrometeorol.*, 2, 525–530. DOI: 10.1175/1525-7541(2001)002<0525:GVRDFL>2.0.CO;2.

Zeng, X., M. Shaikh, Y. Dai, R. E. Dickinson, and R. Myneni (2002). "Coupling of the Common Land Model to the NCAR Community Climate Model". *J. Climate*, 15.14, 1832–1854. DOI: 10.1175/1520-0442(2002)015<1832:COTCLM>2.0.CO;2.

Zeppetello, L. R. V., L. A. Parsons, J. T. Spector, R. L. Naylor, D. S. Battisti, Y. J. Masuda, and N. H. Wolff (2020). "Large scale tropical deforestation drives extreme warming". *Environ. Res. Lett.*, 15.8, 084012. DOI: 10.1088/1748-9326/ab96d2.

Zhang, M., X. Lee, G. Yu, S. Han, H. Wang, J. Yan, Y. Zhang, Y. Li, T. Ohta, T. Hirano, J. Kim, N. Yoshifuji, and W. Wang (2014). "Response of surface air temperature to small-scale land clearing across latitudes". *Environ. Res. Lett.*, 9.3. DOI: 10.1088/1748-9326/9/3/034002.

Zhang, Y., A. Middel, and B. L. Turner (2019). "Evaluating the effect of 3D urban form on neighborhood land surface temperature using Google Street View and geographically weighted regression". *Landscape Ecol.*, 34, 681–697. DOI: 10.1007/s10980-019-00794-y.

Zhang, Y., C. Peng, W. Li, L. Tian, Q. Zhu, H. Chen, X. Fang, G. Zhang, G. Liu, X. Mu, Z. Li, S. Li, Y. Yang, J. Wang, and X. Xiao (2016). "Multiple afforestation programs accelerate the greenness in the 'Three North' region of China from 1982 to 2013". *Ecol. Indic.*, 61, 404–412. DOI: 10.1016/j.ecolind.2015.09.041.

Zhang, Y., A. T. Murray, and B. Turner (2017). "Optimizing green space locations to reduce daytime and nighttime urban heat island effects in Phoenix, Arizona". *Landscape Urban. Plan.*, 165, 162–171. DOI: 10.1016/j.landurbplan.2017.04.009.

Zomer Robert J., Neufeldt Henry, Xu Jianchu, Ahrends Antje, Bossio Deborah, Trabucco Antonio, N. van Meine, and Wang Mingcheng (2016). "Global Tree Cover and Biomass Carbon on Agricultural Land: The contribution of agroforestry to global and national carbon budgets". *Sci. Rep.-UK*, 6, 29987. DOI: 10.1038/srep29987;10.1038/srep29987.