Master Thesis

Local versus nonlocal effects of land cover changes: Comparing methods of signal separation for temperature responses

---

Rafael Stähli
17-107-509

Wyss Academy for Nature
Climate and Environmental Physics
Oeschger Center for Climate Change Research
University of Bern

Supervisor: Prof. Dr. Edouard Davin
Co-Supervisor: Dr. Quentin Lejeune (Climate Analytics)
Advisor: Dr. Shruti Nath (Climate Analytics, ETH Zurich)
Advisor: Steven De Hertog (Vrije Universiteit Brussel)

August 2023

Abstract

Land cover change (LCC) affects biogeochemical and biogeophysical processes, leading to changes in surface temperature. The effect of LCC has not only a local (i.e., at the immediate area of change) surface temperature effect but also nonlocal effects on surface temperature in neighboring or remote areas. To understand the local and nonlocal effects, we must separate the total effect into local and nonlocal effects. The two signal separation methods we investigated are the checkerboard method and the moving window (MW) method. Before the LAMACLIMA[1] project began, the small amount of dedicated checkerboard experiments limited comparisons between the checkerboard and MW methods. Our objective is to compare the results of the checkerboard and MW methods in the same dedicated LAMACLIMA experiments to find an agreement or middle ground between the two methodologies while highlighting the circumstances that indicate difficulties and disagreement. By highlighting the benefits and limitations of the two methods, our second aim is to make recommendations under which scenarios each method should be chosen. Focusing on local effects, we show that both methods generally agree well for deforestation. For afforestation, the local effects calculated by the two methods show more deviations. When looking at individual grid cells, some clear differences occur. We found that the checkerboard method presents spurious artefacts in areas with nearly absent LCC. In the MW method, the choice of the MW size notably influences local effects: larger MW sizes can include more diverse LCCs, increasing linear regression robustness, while reducing spatial variability. Our research takes a step toward a more comprehensive analysis of two signal separation methods while providing insights into the relative performance of the checkerboard and MW methods.

[1]LAMACLIMA - LAnd MAnagement for CLImate Mitigation and Adaptation. The LAMACLIMA project led by Climate Analytics aims to increase public and scientific understanding of the interactions between LCLM and climate to inform the development of effective land-based adaptation and mitigation strategies.

Contents

1 Introduction

The rise in temperature and hydrological extremes brought on by climate change is of great concern to society as it poses a risk to human health and ecosystem functioning (IPCC AR6 SPM, 2021). Therefore, developing strategies for global climate mitigation and local adaptation is of utmost importance, and land-based strategies are expected to play a key role in this (IPCC SRCCL, 2019). Besides their potential mitigation effect, land management practices such as forest restoration and management, no-tillage, cover crops, and crop albedo management have been shown to affect temperature locally in particular during heatwaves (Davin et al., 2014; Hirsch et al., 2017; Seneviratne et al., 2018; Hirsch et al., 2018; Schwaab et al., 2020). Through land cover change (LCC), e.g. the substitution of forests for croplands for agricultural purposes, humans have significantly altered the land's surface (Pongratz et al., 2008). Historical land cover changes has been shown to affect the past climate (Pitman et al., 2009; Boisier et al., 2012; De Noblet-Ducoudré et al., 2012), with regional climate effects being similar in magnitude to the concomitant rise of emissions throughout the period from 1850 - 2010 (De Noblet-Ducoudré et al., 2012).

This thesis focuses on the effects of tree cover change (TCC) as part of LCC. More specifically, we focus on deforestation and afforestation, including reforestation within afforestation. Afforestation and reforestation are commonly grouped under the term forestation. However, we use afforestation throughout this thesis to remain consistent with the terminology used by De Hertog et al. (2023). Since we use the same simulation setup as in (De Hertog et al., 2023).

Despite a reduction in the annual tree cover loss over recent years, the World Resources Institute indicates a substantial loss of 22.4 Mha of tree cover globally in 2022 (Figure 1). Furthermore, 459 Mha of tree cover have been lost since the beginning of the new millennium, showing the large scale of TCC in the past. Aside from significant tree cover losses, large tree cover increases are visible. With the Paris Agreement in 2016, 196 nations signed an agreement to undertake mitigation strategies to limit the mean global surface temperature increase, with respect to pre-industrial levels, to below 2 Kelvin. Regarding land management, afforestation has one of the largest mitigation potentials in terms of carbon dioxide removal, as graphically depicted in Figure 1. The mitigation potential has driven multiple organizations (e.g., Ecosia, One Tree Planted, 1t) to declare their participation in extensive tree-planting initiatives, committed to further amplifying their efforts in the forthcoming years. The large-scale deforestation and afforestation occurring underline the importance of comprehending the effects of TCC on the climate system.

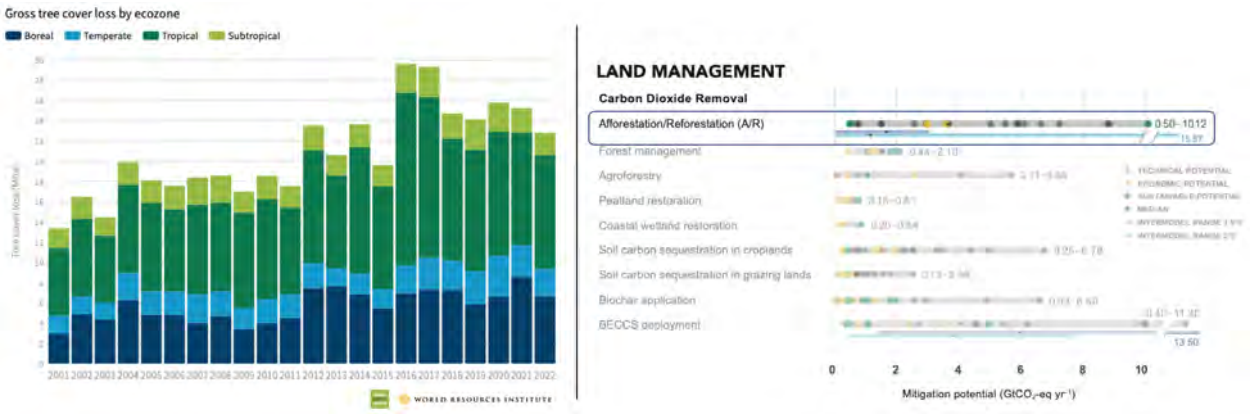

Figure 1: The left side shows the gross tree cover loss by ecozone (adapted Figure from World Resources Institute and the Global Forest Watch). The right side highlights the mitigation potential of afforestation/reforestation in 2020–2050 regarding carbon dioxide removal (Adapted Figure from IPCC SRCCL, 2019). For the right side only research that provides estimates of how much CO_2 equivalent per year could be reduced by 2050 is included. The mitigation potential is calculated by studies from 2010 - 2017 that represent technical, economic, or sustainable potential.

As indicated by the IPCC special report "Climate Change and Land" (2019), LCC not only affects biogeochemical (BGC) processes but also the energy and hydrological balance of the land surface, referred to as biogeophysical (BGP) effects (IPCC, 2019). In other words, the BGP effect reflect the impact of physical land surface changes, such as the changes in albedo, evapotranspiration, or surface roughness, while the BGC effect reflect changes in emissions of compounds, such as changes in the CO_2 emission (Pongratz et al., 2010; Davin and de Noblet-Ducoudré, 2010; Bright, 2015).

In this thesis, we focus on the BGP effects of TCC. More specifically, we analyze the BGP effect of TCC on surface temperature[2]. So, from now on, unless specified otherwise, when we refer to the effect of TCC, we refer to the BGP effect of TCC on surface temperature. The effects of TCC extend beyond local areas of change, affecting the surface temperature in neighboring and remote regions as well (Spracklen et al., 2018; Cohn et al., 2019; Davin et al., 2020). Therefore, when investigating the effect of TCC on surface temperature, it is essential to be aware of both the local and nonlocal effects.

The BGP effects of TCC have been analyzed through observational studies, particularly with paired in-situ and remote sensing measurements. These studies predominantly utilize a space-for-time strategy. Thereby, it compares two spatially distinct areas, in close proximity, assuming similar atmospheric forcing. By analyzing the same time period, the differences in surface temperature between these areas can be attributed to tree cover variations. Additionally, the space-for-time approach can be applied to areas wherein one area experiences TCC while the other remains unchanged (no-TCC) (Duveiller et al., 2018; Meier, 2021). Such observational studies encounter challenges. One challenge arises from the assumption of identical atmospheric forcing in both areas, which cannot be assured and introduces measurement noise. Furthermore, these observational studies inherently measure either local or nonlocal BGP effects but cannot simultaneously measure

[2]also called "skin" temperature in the IPCC Standard Output from Coupled Ocean-Atmosphere GCMs (Meehl et al., 2007)

the local and nonlocal effects. This inability highlights the need for good signal separation methods that can be applied to model simulations to disentangle the total effect into local and nonlocal effects. In this thesis, we will analyze and compare two existing signal separation methods applied to Earth System Model (ESM) simulation experiments.

The checkerboard method is one of the two signal separation methods we investigate in this thesis. The checkerboard method, originally described by Winckler et al. (2017), uses the design choice where TCC is altered only in every other grid cell[3] following a checkerboard pattern. This alternating TCC design choice must be implemented before running the simulation. Therefore implementing the checkerboard approach requires a dedicated ESM experiment, making it computationally expensive. In contrast, the moving window (MW) method is more flexible. The MW method is the second signal separation method investigated in this thesis. It can technically be applied to several existing LCC experiments since it does not depend on a regular spatial patterned LCC. Therefore, the MW method does not rely on dedicated ESM simulations, resulting in higher computational efficiency and flexibility (Lejeune et al., 2018; De Hertog et al., 2023).

Due to the lack of dedicated checkerboard experiments, there is a constrained capacity for comparing the checkerboard and the MW methods. The LAMACLIMA project created dedicated checkerboard experiments that offer the opportunity to compare the checkerboard to the MW method in-depth. Thus, our objective is to compare the results of the checkerboard and MW methods in the same dedicated LAMACLIMA experiments to find an agreement or middle ground between the two methodologies while highlighting the circumstances that indicate difficulties and disagreement.

The following steps are taken in order to accomplish the objective of this thesis. First, we compare the results of the checkerboard and the MW methods in the LAMACLIMA experiments for two scenarios: deforestation and afforestation relative to the 2015 "present-day" control levels. This comparison will be made through a spatial comparison, the analysis of 2-D histograms showing the Pearson correlation coefficient of both methods and a focus on the local effects sign agreement and disagreement of both methods. Secondly, we will investigate possible reasons explaining the comparison differences. Thereby we will investigate, among others, possible spurious artefacts from the checkerboard method and the influence and the added values of different MW sizes in the MW approach.

[3]Grid cells are equally-sized rectangular (2D) or cubic (3D) spatial locations that collectively form a spatial coverage of a large area.

2 Methods & Data

2.1 Earth System Models

Earth System Models (ESMs) consist of multiple submodels that describe, among others, the atmosphere, ocean, land, runoff, and sea ice and how their processes react upon change. To do so, the submodels calculate interactive processes and feedback within three-dimensional grid cells with a certain resolution, thereby modeling the Earth's climate (Collins et al., 2011). Linking these submodels together gives a fully-coupled ESM, which enables a comprehensive understanding of global climate effects based on anthropogenic forcings. Figure 2 presents the submodels employed in the ESMs used in this thesis. We are especially interested in understanding the climate effects based on changes in the land surface submodels. More precisely in changes in land cover, specifically tree cover. Within the land surface submodels, each grid cell consists of smaller tiles representing distinct vegetation types. Depending on the model, the vegetation types may differ. These vegetation types are typically categorized into two main groups: Plant Functional Types (PFT) and Crop Functional Types (CFT). Each of these groups further incorporates subgroups representing, e.g., different PFTs such as deciduous and extra-tropical evergreen trees.

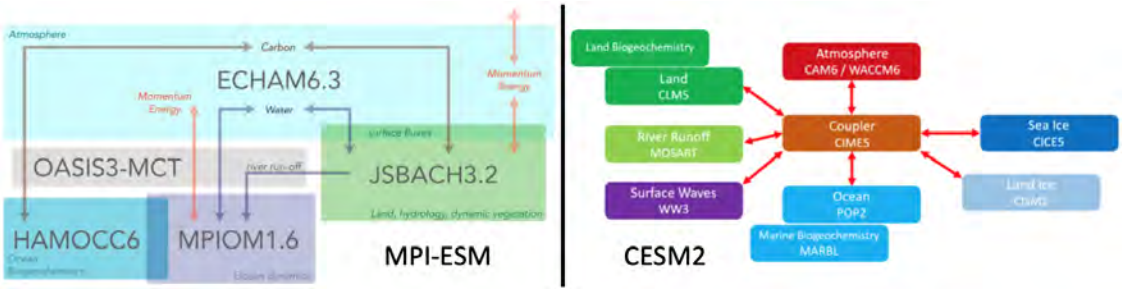

Figure 2: Visualization of the submodels within the fully-coupled MPI-ESM on the left and the submodels within the fully-coupled CESM2 on the right. Adapted Figure from Mauritsen et al. (2019); Danabasoglu et al. (2020).

The data used in this study are ESM simulations specifically designed to isolate the effects of TCC and separate their local and nonlocal components using the checkerboard method. The simulation experiments employed in this thesis have been generated through the LAMACLIMA project (De Hertog et al., 2023) and utilize the following two ESMs:

MPI-ESM: Stands for Max Planck Institute Earth System Model.
This fully-coupled ESM, illustrated in the left part of Figure 2, makes use of the land component JSBACH3.2 and the atmospheric component ECHAM6.3, which are connected to the ocean dynamic and biogeochemistry models (MPIOMI1.6 and HAMOCC6) via OASIS3-MCT. The land cover in MPI-ESM is represented through a total of 12 different PFTs, including 2 CFTs and 4 tree PFTs (extra-tropical evergreen and deciduous trees as well as tropical broadleaf evergreen trees) (Mauritsen et al., 2019).

CESM2: Stands for Community Earth System Model 2.
This open source, fully-coupled model illustrated on the right side of Figure 2, described by Dan-

abasoglu et al. (2020), includes submodels for the ocean, atmosphere, land ice, sea ice, river runoff, and land. The CESM2 includes 8 CFTs, which can grow on rainfed or irrigated patches, and 14 natural PFTs (Danabasoglu et al., 2020).

It is important to recognize that CESM2 and MPI-ESM have differing interpretations of identical input land cover maps, given that each model incorporates distinct PFTs. Consequently, we anticipate slight disparities in the surface temperature effects of these two models due to the dissimilarities in the TCC maps. Davin et al. (2020) describe how to construct idealized land cover maps.
In this research, if not mentioned differently, we use the native grid cell resolution of the CESM2 and the MPI-ESM (lat $\times$ lon; CESM2: $0.90^\circ \times 1.25^\circ$ and MPI-ESM: $1.88^\circ \times 1.88^\circ$). MPI-ESM has a lower resolution than CESM2, resulting in spatially larger grid cells and, thus, fewer data points. It is important to consider this grid cell resolution difference and its potential influence when comparing the TCC effects of both models. We expect the pattern of the TCC effect from MPI-ESM to have less spatial variability than from CESM2 due to spatially larger grid cells. Figure 3 visualizes the MW for 5x5, 7x7, 9x9 and 11x11 grid cells. We added the 11x11 MW to visually show why we believe that we lose the "locality" aspect if the MW is spatially so large. Further, we did not include the 3x3 MW since we assumed it would not have enough data points for a statistically robust response. We analyze the 5x5, 7x7, and 9x9 MW sizes based on these assumptions.

Figure 3: Visualization of the spatial dimension of different MW sizes. The central point corresponds to the initial MW centered grid cell. The first window visible is the 5x5 MW, then 7x7, 9x9 and 11x11.

2.2 Signal separation of local and nonlocal effects

2.2.1 Definition of local and nonlocal effects

As indicated in Section 2.1, ESMs consist of three-dimensional grid cells, modeling the Earth's climate. For grid cells on land, we distinguish between grid cells that undergo TCC and grid cells that do not undergo TCC. We will refer to them as TCC grid cells and no-TCC grid cells. A total or "net" effect of TCC, consisting of local and nonlocal effects, is visualized in Figure 4. We define local effects as those effects that exclusively result from changes in the surface properties of the grid cell "A" itself and which occur on "A". The nonlocal effect is the emerging effect received on "A" due to TCC of nearby or remote grid cell "D". A grid cell that does not undergo TCC can have a nonlocal effect even though its surface properties have not changed. Grid cells "B" and "C"

experience no local TCC. Thus, as a consequence, the nonlocal effect is equal to the total BGP effect for no-TCC grid cells.

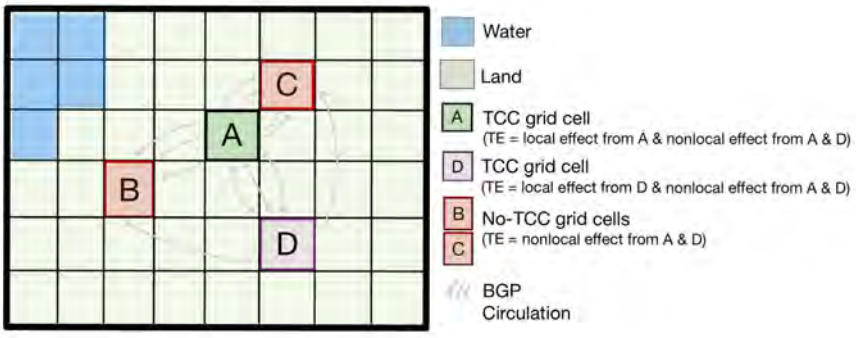

Figure 4: Schematic to visualize how TCC generates local and nonlocal effects. TE stands for total effect, and BGP for biogeophysical. No-TCC grid cells stand for grid cells that do not undergo TCC. The gray arrows depict the BGP circulation change brought on by TCC, which impacts remote grid cells.

2.2.2 Signal separation methods

The checkerboard method is a systematic method created by Winckler et al. (2017) that requires dedicated ESM simulations. In the checkerboard method, only every second grid cell undergoes TCC, resulting in a checkerboard pattern, visible in Figure 5b). Grid cells without TCC represent a nonlocal effect, capturing climatic changes due to neighboring TCC. Since there is no local effect in grid cells with no-TCC, their nonlocal effect equals the total effect. The nonlocal effects of these grid cells are illustrated in Figure 5b). The grid cells that undergo TCC are masked-out in gray, resulting in a checkerboard pattern of nonlocal effects. To differentiate the total effect at TCC grid cells into local and nonlocal effects, the nonlocal effects from no-TCC grid cells in Figure 5b) are linearly interpolated onto the TCC grid cells. Therewith, the nonlocal effect is known for every grid cell, illustrated in Figure 5c). By subtracting the calculated nonlocal effect from the total effect at the TCC grid cells, the local effect, visible in 5d), can be generated. Then the local effects are interpolated to the grid cells without TCC, receiving a global map of local effects presented in Figure 5e). In order to receive the total effect if every grid cell would undergo TCC, the Sub-figures 5c) and 5e) have to be added up. The MW method will be applied to these total effects (based on Figure 5c & 5e). The initial total effect map from the simulation experiment, visible in Figure 5a), consists of grid cells without TCC (nonlocal effect in Figure 5b) and the grid cells with TCC (combining the nonlocal effect in Figure 5c and the local effect in Figure 5d).

Based on Winckler et al. (2017), the checkerboard pattern where every second grid cell undergoes TCC was used, as it is the most robust. An advantage of the checkerboard method is that it provides a clear understanding of the origin of nonlocal effects. It enables comprehensive isolation of all components (i.e., local, nonlocal, and total effects)(Meier, 2021). However, the requirement of a simulation setup where only every second grid cell undergoes TCC poses a constraint on the potential application of the checkerboard method on other existing simulation experiments.

As the checkerboard method depends on interpolation, the results can differ based on the TCC pattern applied. A linear interpolation on a setup where every second grid cell undergoes TCC has been suggested and applied (Winckler et al., 2017; De Hertog et al., 2023). Winckler et al. (2017) calculated an interpolation bias of approximately 0.25 K, which agrees with our findings (Appendix A). Thus, this bias can be considered as negligible.

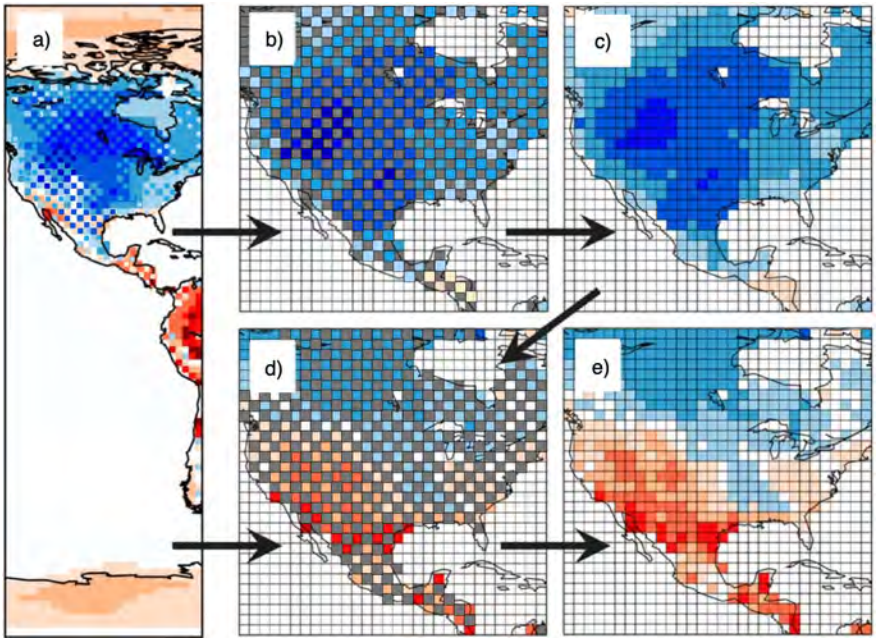

Figure 5: Schematic of the signal separation of the checkerboard method. (a) The signal from the dedicated checkerboard simulation. Showing a simulation total effect based on a checkerboard patterned TCC map. (b) The nonlocal effects at no-TCC grid cells. (c) The nonlocal effects of (b) interpolated to the TCC grid cells. (d) The local effect at TCC grid cell. Calculated by subtracting (c) from (a). (e) The local effects of (d) interpolated to the TCC grid cells. The gray grid cells are grid cells where no information is needed. This is an adapted Figure from Winckler et al. (2017).

The moving window (MW) method used in this research follows the methodology outlined by Lejeune et al. (2018). To extract the TCC effects, a factorial approach has been applied. The factorial approach compares the surface temperature of TCC grid cells versus grid cells without TCC. However, such a factorial approach can not be applied to all-forcing experiments, i.e., experiments where every grid cell undergoes TCC. Therefore, Kumar et al. (2013) developed a MW methodology that differentiates between low TCC and high TCC and can calculate the TCC effect of all-forcings experiments. To be able to apply the Kumar et al. (2013) methodology, the grid cells within the MW are assumed to be affected similarly by external forcings. This assumption is also essential within the MW method we apply in this thesis. We used a MW method that was extended by Lejeune et al. (2017) and outlined by Lejeune et al. (2018). Both Kumar et al. (2013) and Lejeune et al. (2018) demonstrated that the TCC effect can be captured, to some extent, by this MW method. To disentangle the local and nonlocal effects in one specific grid cell, we calculate a linear regression between the simulated total surface temperature changes and the imposed TCC within a MW, i.e., a quadratic window consisting of several grid cells centered around grid cell i,j:

$$\Delta T_{i,j} = \beta * TCC_{i,j} + c \tag{1}$$

The regression coefficient β represents the slope of the linear regression, capturing the relationship between the change in tree cover ($TCC_{i,j}$) and the resulting total effect ($\Delta T_{i,j}$) for grid cell i,j. The intercept c has no direct relationship to the $TCC_{i,j}$ considered and corresponds, thus, to the nonlocal effect. The MW method is illustrated in Figure 6.

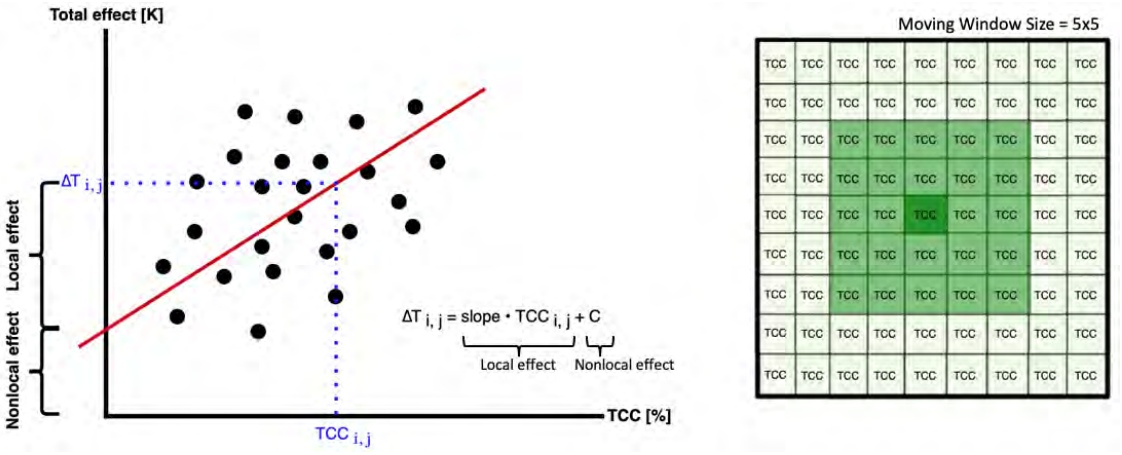

Figure 6: Schematic illustrating the signal separation of the MW method. A 5x5 MW is illustrated on the right side, which is used to calculate the signal separation of the MW centered grid cell. On the left side, the linear regression, which is calculated for each MW, is presented. The black points represent signals for each grid cell within the MW. The red line shows the slope, and the blue dotted lines help to see the TCC and its corresponding total effect of the MW centered grid cell i,j.

It is important to note that the choice of the MW size influences the calculated effects, as the linear regression is performed using the grid cells within the MW. We analyze three different MW sizes to explore the impact of different MW sizes. We focus on the 5x5, 7x7, and 9x9 MW as we assume they have a good balance between capturing the influence of the neighboring grid cells (without a too large spatial extent) and ensuring enough grid cells for a robust linear regression calculation. These assumptions, as well as possible guidelines when selecting the MW sizes, will be analyzed in the course of this thesis. Note that our analysis, if nothing else is mentioned, is conducted at the native resolution of the ESMs, meaning that the spatial extent of a 5x5 MW varies across different ESMs.

2.3 Simulation setup

In this thesis the LAMACLIMA simulations described by De Hertog et al. (2023) are used. Thereby the focus is on the BGP effects of TCC. To achieve idealized TCCs in comparison to the known tree cover data of the year 2015 (control), the following three sensitivity experiments have been analyzed on the ESM-specific native resolution:

- Control (CTL): Land cover maps and other external forcings are kept constant at 2015 levels.
- Forest (FRST): Complete expansion of tree cover to a fully afforested world in relation to 2015 levels while holding all external forcings constant. Thereby, non-forest PFTs are removed,

and forest PFTs increased to 100 %, while bare soil is preserved in order to not plant forests in unrealistic areas, e.g., desert. The latitudinal mean forest distribution is taken for grid cells where no forest PFT fraction existed in the control year.

- Crop (CROP): Complete expansion of crop cover in relation to 2015 levels while holding all external forcings constant. Thereby natural PFTs are removed, and cropland CFTs increased to 100 %, while bare soil is preserved in order to not plant crops in unrealistic areas, e.g., deserts.

The difference between FRST and CTL simulations will be used to identify the afforestation effect. The difference between CROP and CTL simulations can technically not directly be interpreted as a deforestation effect, since all PFTs, including non-forest PFTs such as shrubs and grass, have been replaced by crops. Therefore, CROP - CTL is more precisely referred to as LCC towards complete crop expansion. However, for simplicity, we continue to refer to the LCC crop expansion experiment used for the checkerboard method as deforestation. For the MW method, we only use the deforestation TCC map. This creates a slight difference between the LCC (respectively TCC) maps used by the two methods (Appendix B) and consequently their interpretation. The reasoning behind this distinction of considering only tree-PFTs for the MW method is due to the heterogeneity of the LCC. For example, there could be grid cells where mostly grasses or shrubs are removed, while other grid cells mainly remove tree PFTs. This heterogeneity of the LCC would make the analysis more difficult, as different vegetation types have different local effects. Thus, within the MW size, respectively for the linear regression, it would become more complex to calculate the relationship (slope) between the LCC and the total effect. The TCC maps used are visible in Figure 7, while the LCC maps used for the checkerboard can be seen in Appendix B.

Further, it is important to note that we use a simulation setup that was created to optimize the checkerboard method. The FRST and CROP simulation setup is implemented in a checkerboard pattern since we know from Section 2.2.2 that a dedicated simulation setup is needed to apply the checkerboard method. This means that during the 160-year simulation period investigated, only every second grid cell undergoes TCC.

In the LAMACLIMA simulations, the same present-day (2015) CO_2 level and other climate forcings remain constant during the entire time frame. The CMIP6 historical simulations from 2015 serve as the initial conditions. These initial conditions are then applied to both ESMs to run model simulations over 160 years to average out internal variability. To ensure that stratospheric aerosol concentrations in the LAMACLIMA simulations are representative, the first 10 years (from 2015 - 2025) function as BGP spin-up based on the Scenario Model Intercomparison Project (ScenarioMIP) data. For the remaining 150 years of simulations, the stratospheric aerosol concentrations were used constantly at the spin-up level. For CESM2, the atmospheric aerosol concentration of 2025 was directly used from the simulation start because of technological limitations. Furthermore, for CESM2, the solar forcing has been kept constant at the average rate for the full simulation period. For MPI-ESM, the solar forcing has been kept at the natural oscillation. The reasoning behind these choices is to find the future BGP effects of a TCC in the 2015 “present-day” setup,

where every other factor is held constant and equal to the 2015 levels (De Hertog et al., 2023).

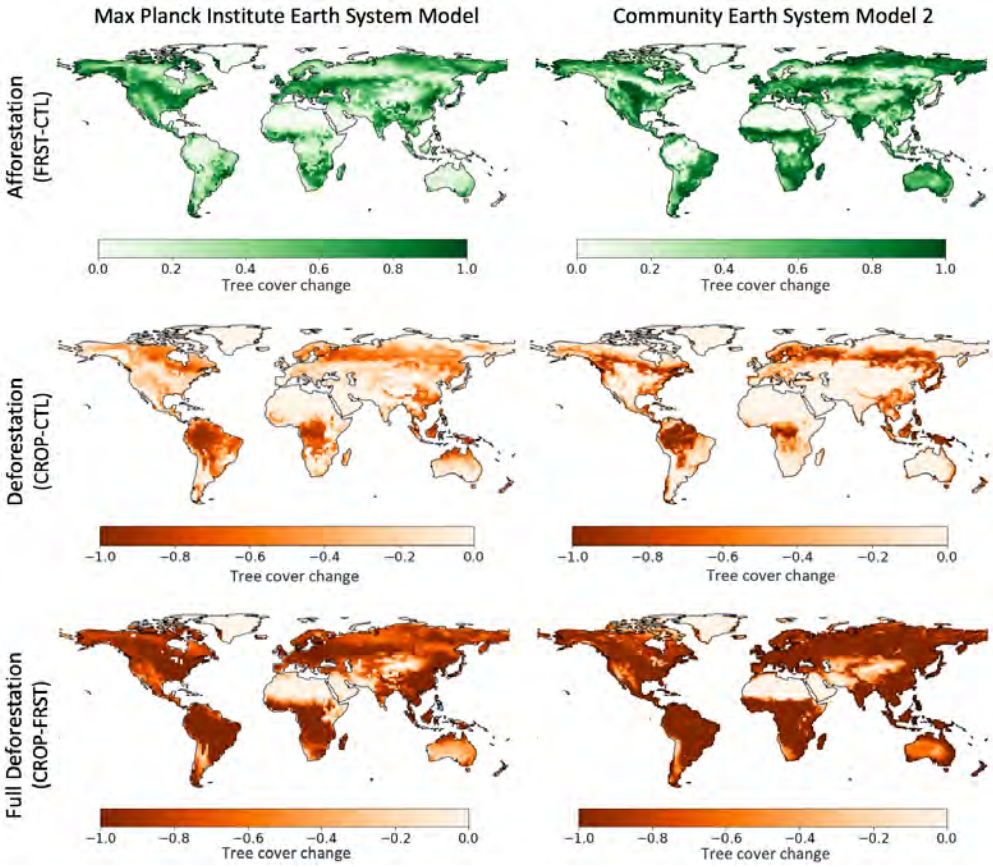

Figure 7: Overview of the tree cover change maps used in the MW method for each scenario investigated. The tree cover changes are relative to the present-day 2015 levels. Note that CESM2 has a higher resolution than MPI-ESM.

2.4 Analysis setup

2.4.1 Seasonal analysis

In our analysis, we focus on seasonal effects. Instead of considering the annual mean, we focus on the boreal summer mean (June, July, August) and the boreal winter mean (December, January, February). Considering the seasonal mean instead of the annual mean enables us to explain better the TCC effects in terms of the different BGP drivers since BGP drivers can vary between seasons (Bonan, 2008; Boisier et al., 2012; Davin et al., 2020).

2.4.2 Surface temperature vs. 2-meter air surface temperature

The MW and the checkerboard method could be used to investigate the signal separation of several variables. Be it surface temperature, 2-meter air temperature, or precipitation. Our study will only focus on the effect on surface temperature. Winckler et al. (2019b) demonstrated that for MPI-ESM, the deforestation nonlocal effect of surface temperature and 2-meter air temperature are rather similar. However, in the case of the local effect, they demonstrated that for MPI-ESM, the annual mean surface temperature response is nearly twice as strong compared to that of the

2-meter air temperature, except for northern high- to mid-latitudes. The different temperature definitions can explain the difference between the effects of the two temperature variables. For MPI-ESM, for example, the 2-meter air temperature is defined by a linear regression between the surface temperature and the lowest atmospheric layer. In this case, the 2-meter air temperature is also influenced by atmospheric changes, in contrast to surface temperature, which is mainly influenced by changes in surface variables such as albedo (Winckler et al., 2017). Furthermore, Winckler et al. (2019b) highlighted that different models define 2-meter air temperature differently, indicating that the comparison of surface temperature is more consistent, also in comparison with satellite observations. Knowing these temperature differences, we have to keep in mind that our results will be surface temperature specific and consequently show mostly higher magnitudes compared to local effects on 2-meter air temperature. For more details on when to use surface temperature or 2-meter air temperature, see Winckler et al. (2019b).

3 Results

3.1 Comparison of the results of the checkerboard and moving window methods in the LAMACLIMA experiments

In this section, we compare the local effects calculated with the two methods (checkerboard and MW) for the deforestation and afforestation scenarios. We calculate the local effect using a 5x5 MW size for the MW method unless specified otherwise.

3.1.1 Deforestation (CROP - CTL)

Investigating Figure 8, we see areas where both methods show similar TCC effects, while in other areas, the methods show opposite TCC effects. During the winter months in northern latitudes, both methods show a strong cooling effect as a result of deforestation. The snow-masking effect enhances the albedo effect, which can explain the cooling effect observed during winter (Qu and Hall, 2007; Abe et al., 2017). For summer, the methods show less agreement on the TCC effect in boreal and temperate regions. Another example of an area where both methods show similar TCC effects is in the tropics for CESM2. Both methods show a clear warming effect for CESM2 in the rainforest regions. For MPI-ESM, the TCC effect of both methods in the tropics is less clear. While the checkerboard method shows a slight warming effect, the MW method shows cooling effects over certain tropical regions. According to Chen and Dirmeyer (2016, 2020), a decrease in evapotranspiration can explain the warming in the tropics resulting from deforestation. Additionally, the change in surface roughness mainly drives the warming over boreal and temperate regions.

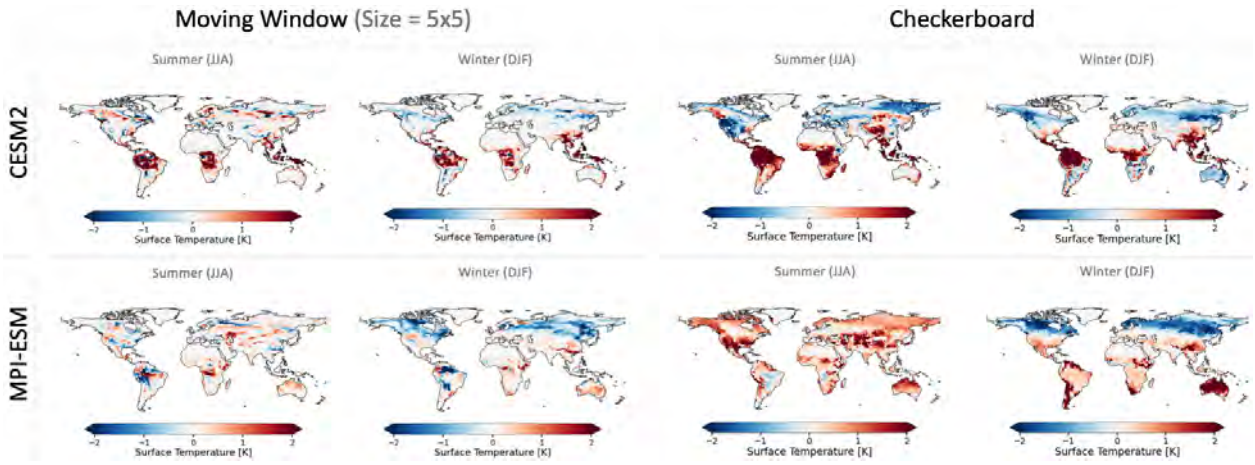

Figure 8: Spatial visualization of the local effect for the deforestation scenario. The columns represent the local effect for deforestation of a specific signal separation method. We differentiate between summer (representing the mean of June, July, August) and winter (representing the mean of December, January, February). A 5x5 MW size has been used.

Overall, using both methods, we can calculate the local effects of deforestation. Whereby the agreement on TCC impacts are higher in some areas than others. Another way of comparing the two methods is with a 2-D histogram. Figure 9 makes it easier to compare individual grid cells and enables us to calculate the Pearson correlation coefficient. By investigating the Pearson correlation

coefficients, we can verify the findings from the spatial world map. Generally, we see a relatively high correlation between the methods (checkerboard and MW), with 0.69 and 0.77 for CESM2. Thus, for deforestation, CESM2, both methods show similar TCC local effects. For MPI-ESM, the correlation is slightly lower, with 0.51 for the winter mean. We see no correlation between the methods regarding the summer mean of MPI-ESM. This zero correlation links to the global warming in MPI-ESM summer, which looks quite different from the TCC effect patterns from CESM2 and, most importantly, highlights that the checkerboard method identifies effects of TCC (especially warming in summer for MPI-ESM) that the MW method does not show.

Comparing both model outputs, it becomes evident that CESM2 shows larger TCC effects than MPI-ESM. The higher magnitudes of the local effect for CESM2 agree with the annual mean findings by De Hertog et al. (2023). Further, we recognize seasonal differences if we compare the patterns between the seasons. We see that data points show opposite TCC effects between the methods. To further investigate these disagreements on the sign of the TCC effects, we investigate Figure 10, showing the number of grid cells where both methods agree or disagree on the sign of the local effect.

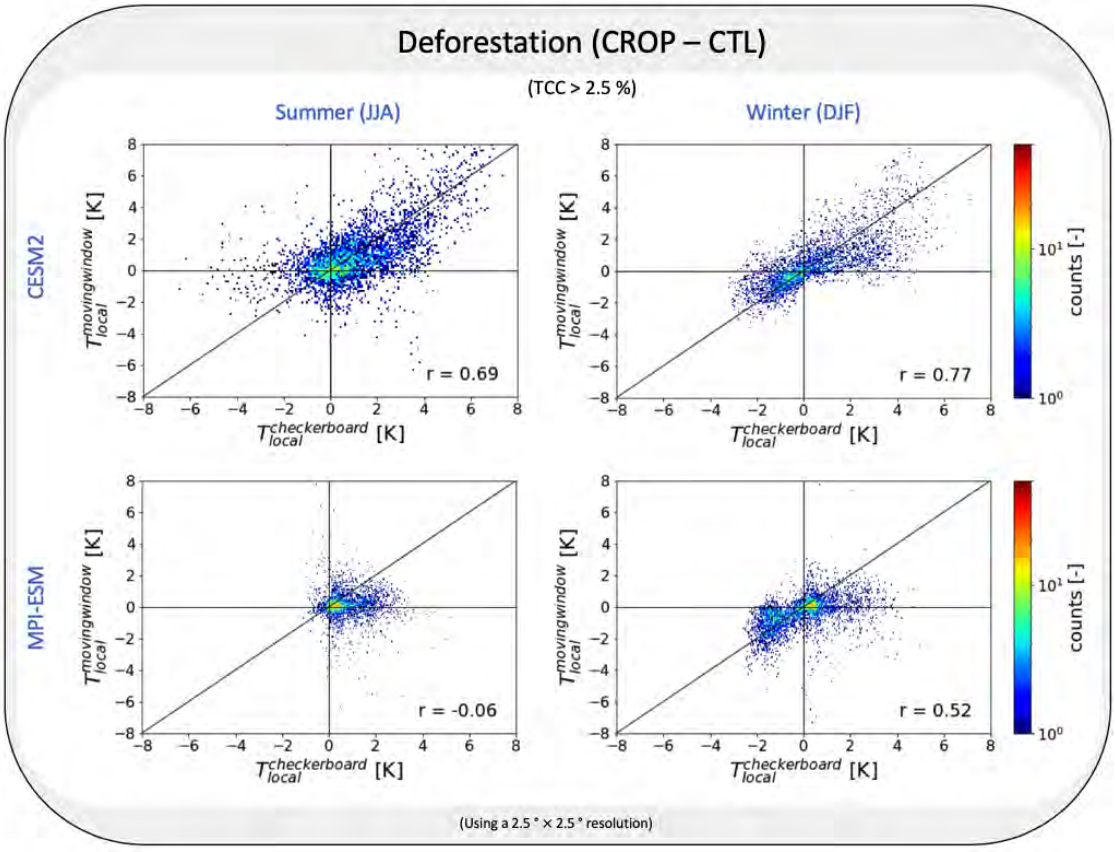

Figure 9: Illustrates 2-D histograms of the deforestation local effects from both methods for both models and seasons. The x-axis corresponds to the local effect of the checkerboard method, while the y-axis represents the local effect of the MW method. A 2.5° × 2.5 ° model grid cell resolution has been used here. The r-values represent the corresponding Pearson correlation coefficient. A 5x5 MW size has been used.

Figure 10 presents on the left side the spatial distribution of the grid cells where both methods agree (light-grey) or disagree (dark-grey) on the sign of the TCC effect for deforestation. Additionally, on the right side, a bar plot visualizes the cumulative count of grid cells that agree or disagree on the sign. Since we want to focus on the local effects based on TCC, we exclude the grid cells with a TCC lower than 2.5 % as these are areas with mostly bare soil, i.e., not representing a TCC-based local effect. We can observe that the agreement or disagreement grid cells are widely spread without a clear pattern for both ESMs, on the left side of Figure 10. The right side of Figure 10 shows relatively good agreement on the sign of the local effects calculated between the two methods, with an agreement between 63 % and 78.5 % for both seasons and models. However, this also highlights that there are many areas where the methods show the opposite TCC effect. Around 21.5 % up to 37 % of the total amount of grid cells show disagreement on the TCC effect sign between the two methods, which is a non-negligible amount. Looking at the bar plot on the right side in Figure 10, we can verify our findings of Figure 8 and Figure 9 of higher agreement between the methods for winter mean compared to summer mean.

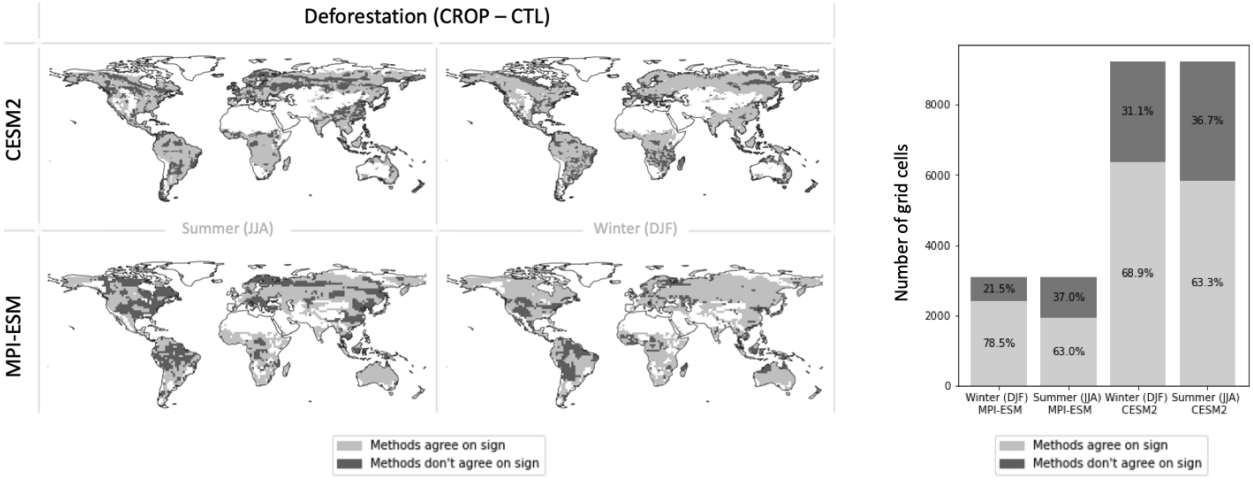

Figure 10: Illustrates the disagreement and agreement on the sign of deforestation local effects from both methods for both seasons and models. On the left side, the disagreement and the agreement on the sign of the local effect are shown spatially, whereby white grid cells correspond to masked-out grid cells with a TCC < 2.5 % or grid cells with the exact same local effect. The number of disagreement/agreement grid cells has been added up on the right side in a bar plot. The percentage values correspond to the disagreement/agreement relative to the total number of grid cells within the model. A 5x5 MW size has been used.

3.1.2 Afforestation (FRST - CTL)

From a visual inspection, the local effects of both methods visualized in Figure 11 agree pretty well. For example, in winter in the northern latitudes, we observe a strong agreement on the warming effect, which finds an explanation in the reverse of the snow-masking albedo effect mentioned for deforestation. However, we see clear and substantial differences between the two methods when focusing on certain areas. For the area around Botswana or Australia in CESM2, we see that the MW method presents very prominent warming effects, while the checkerboard method shows a very prominent cooling effect. Thus, in these areas, they not only show different TCC effects, they also have substantial magnitudes.

Further, looking at CESM2 in summer, for Europe, the Middle East, or North Siberia, we notice much stronger magnitudes for the local effect calculated by the MW method than for the checkerboard. MPI-ESM shows such differences too. However, it is more difficult to compare for MPI-ESM since the local effects calculated by the checkerboard method for summer MPI-ESM show nearly global cooling.

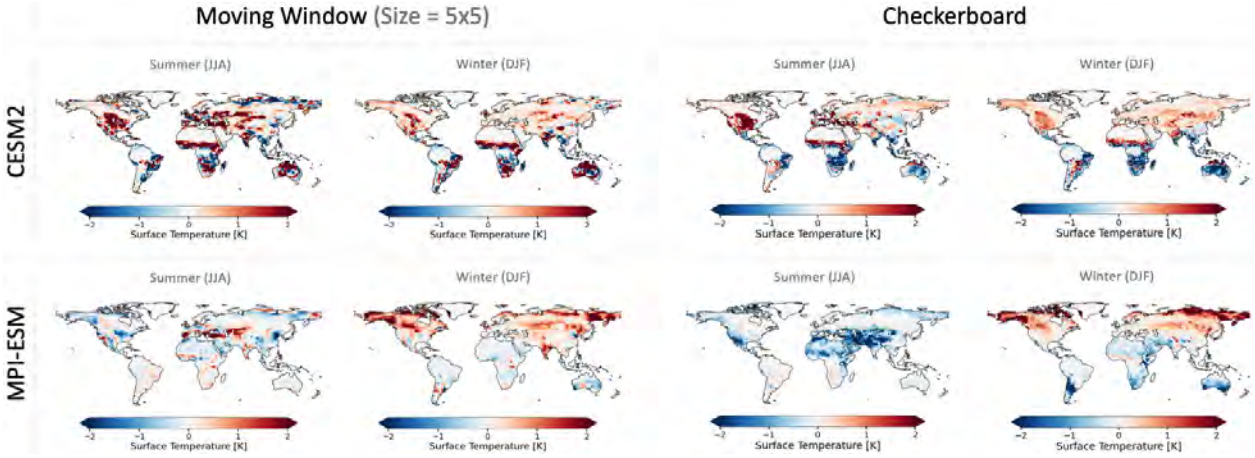

Figure 11: Spatial visualization of the local effect for the afforestation scenario. The columns represent the local effect for afforestation of a specific signal separation method. We differentiate between summer (representing the mean of June, July, August) and winter (representing the mean of December, January, February). A 5x5 MW size has been used.

Therefore, when investigating the 2-D histograms and the correlation between the local effects of both methods in Figure 12, it is not surprising that we see low correlations. For CESM2, we notice a correlation of 0.24 (in summer) and 0.29 (in winter). Thus, for afforestation the methods show a lot lower correlation than for deforestation. However, the correlation between both methods has remained very similar for MPI-ESM. For winter MPI-ESM, the correlation only decreased by 0.06 and slightly improved by 0.07 for winter. As noticed in the comparison within Figure 12 for CESM2 summer, we see that most data points lie above the diagonal line, highlighting that the MW method overestimates the local effects compared to the checkerboard method. Thus, it shows strong methodological differences. Similar to deforestation, we also see afforestation grid cells where both methods do not agree on the sign of the local effect. We investigate the agreement and disagreement on the sign of the local effects in Figure 13.

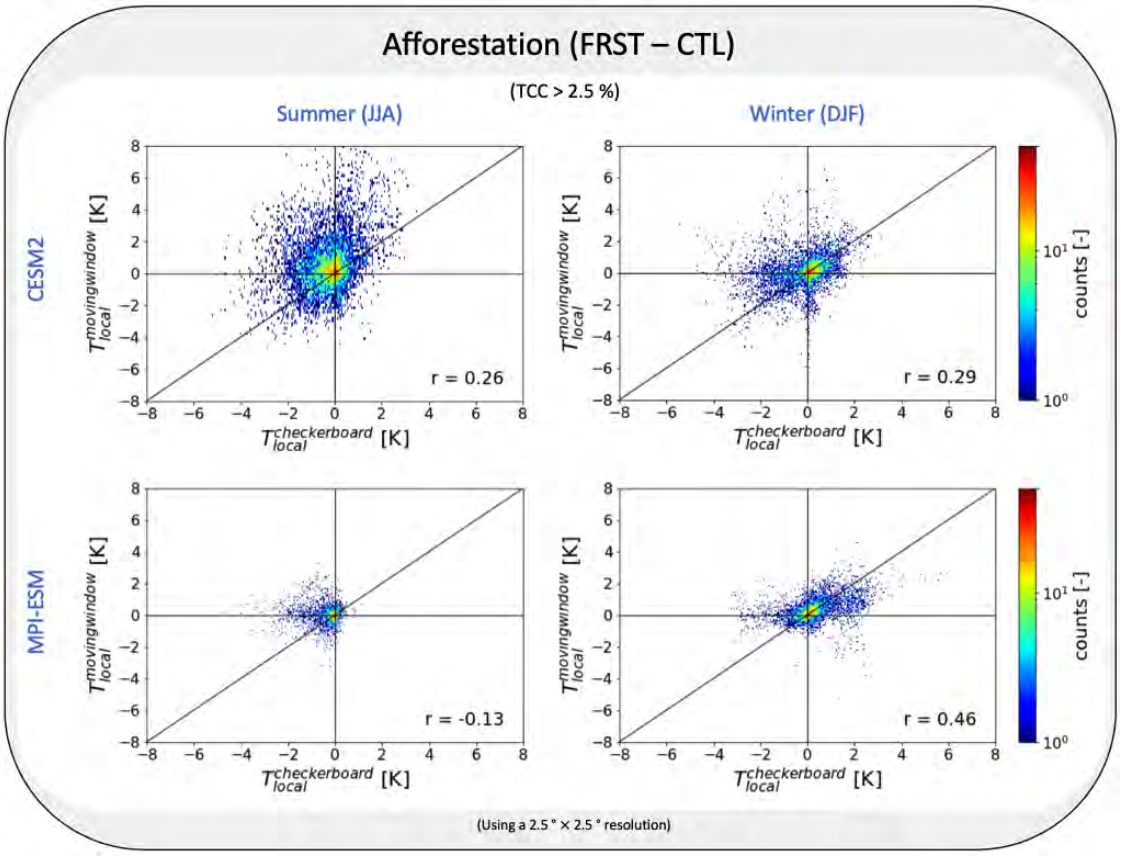

Figure 12: Illustrates 2-D histograms of the afforestation local effects from both methods for both models and seasons. The x-axis corresponds to the local effect of the checkerboard method, while the y-axis represents the local effect of the MW method. A 2.5° × 2.5 ° model grid cell resolution has been used here. The r-values represent the corresponding Pearson correlation coefficient. A 5x5 MW size has been used.

Even though we have a lower correlation for CESM2 for afforestation compared to deforestation, the agreement (percentage-wise) of the two methods on the local effects sign is very similar. It ranges from 60.2 % to 74.5 %. Thus, also for afforestation, there are a non-negligible amount of grid cells where both methods do not agree on the local effects sign. Further, as for deforestation, we see that also for afforestation, the agreement in winter is higher compared to the summer mean. Looking at the spatial distribution of the grid cells where both methods disagree on the sign, we do not see a clear pattern.

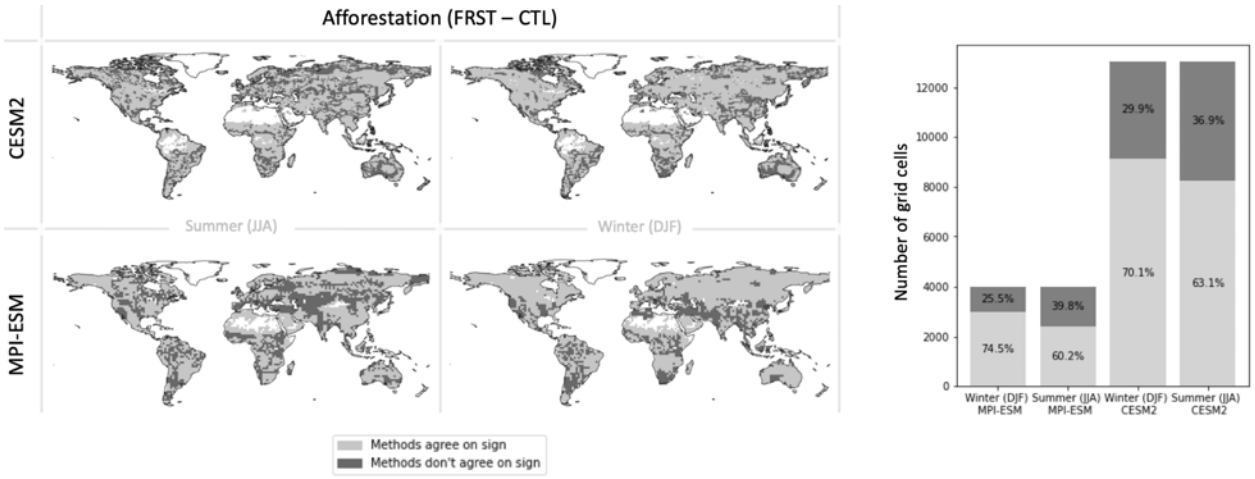

Figure 13: Illustrates the disagreement and agreement on the sign of afforestation local effects from both methods for both seasons and models. On the left side, the disagreement and the agreement on the sign of the local effect are shown spatially, whereby white grid cells correspond to masked-out grid cells with a TCC < 2.5 % or grid cells with the exact same local effect. The number of disagreement/agreement grid cells has been added up on the right side in a bar plot. The percentage values correspond to the disagreement/agreement relative to the total number of grid cells within the model. A 5x5 MW size has been used.

3.2 Investigation of possible reasons explaining differences

There are multiple reasons which potentially explain the differences between both methods. Be it the calculation bias within the method, the different assumptions needed to apply the methods, the different definitions of the local and nonlocal effect of the method, or the simulation experiments the methods are applied on. In the first section, we will investigate what happens when both methods are applied on very low TCC grid cells. Secondly, we will investigate the added values of choosing a larger MW size and lastly, we will have a closer look at the impacts of different MW sizes on the local effect.

3.2.1 Possible artefacts in areas with low TCC

In Section 2.2.1, we defined the local effect to be exclusively from changes in surface properties of the grid cell itself. Consequently, we expect a very low effect on surface temperature for grid cells with very low TCC.

Per definiton, the MW method calculates the local effect as the slope (corresponding to the relationship between the TCC and the resulting total effect) times the TCC of the MW-centered grid cell. Thus, if the TCC of the MW-centered grid cell is nearly zero, the slope gets multiplied by a near-zero value, resulting in low surface temperature changes. On the other hand, the checkerboard method has no constraint for very low TCC in its calculation. Its local effect is determined by subtracting the nonlocal effect from the total effect. In conclusion, the two methods might produce different results because they rely on slightly different definitions.

In Figure 14, for deforestation, the grid cells with less than 2.5 % TCC (very low TCC) have been isolated (masked out in gray) from the other grid cells. As expected, we see very low (or even

absent) local effects for the MW method. On the other hand, we see strong local effects for the checkerboard method. These strong local effects make limited sense, showing a strong local TCC effect based on nearly absent TCC. We assume these effects could be spurious artefacts from the checkerboard method.

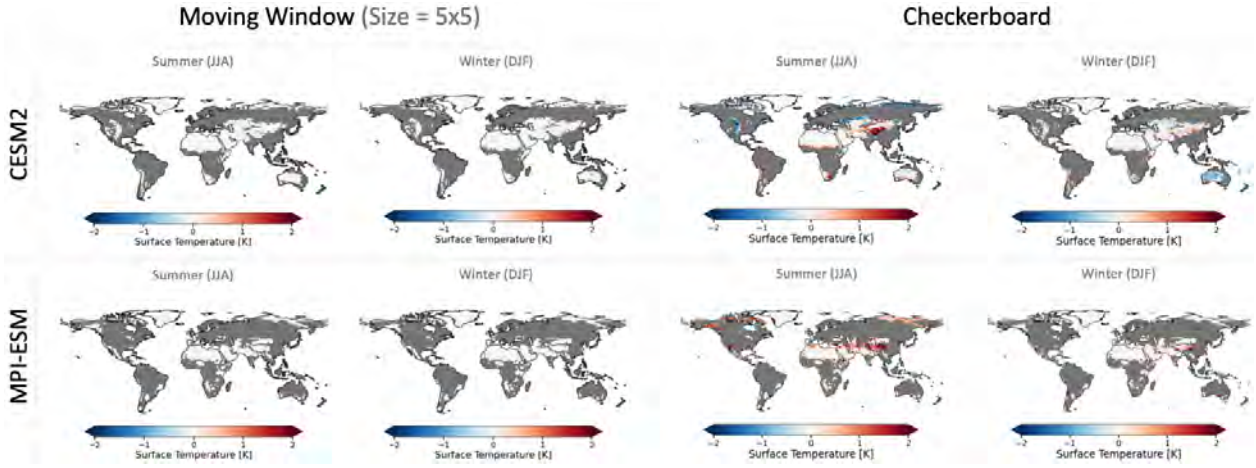

Figure 14: Spatial visualization of the local effects of grid cells with TCC < 2.5 % for the deforestation scenario. The grid cells in gray are masked-out grid cells with TCC larger than 2.5 %. The columns represent the local effect for deforestation of a specific signal separation method. We differentiate between summer mean and winter mean. A 5x5 MW size has been used.

3.2.2 Added values of choosing a larger MW size

In the comparison of both methods in Section 3.1, we have been comparing a 5x5 MW method to the checkerboard method. However, as mentioned in Section 2.1, the MW size can be increased. We will now investigate the differences between the 5x5, 7x7, and 9x9 MW size local effects. Note that for larger MW sizes, the linear regression includes a larger sample size to separate the local and nonlocal effects. Figure 15 compares the disagreement on the sign between both methods of deforestation for all three MW sizes. The differences between the local effects of different MW sizes for afforestation can be seen in Appendix C. Analyzing Figure 15, we see higher agreement on the sign between the methods for larger MW sizes. Thus, although the improvements are small, the 9x9 MW size shows the highest agreement between the MW and checkerboard methods. The most significant improvement from a 5x5 to a 9x9 MW size in terms of agreement increase is 4 % for winter CESM2.

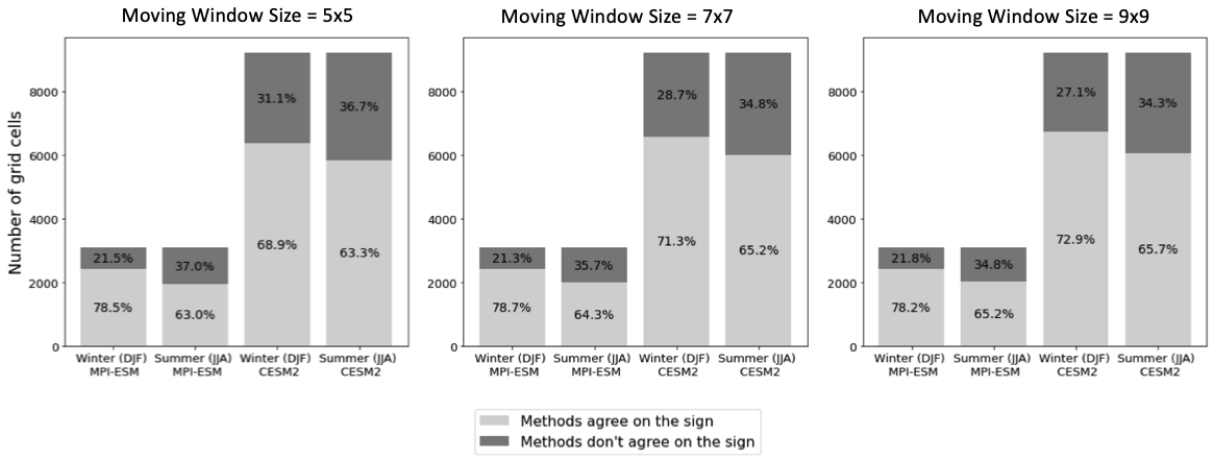

Figure 15: Only considering grid cells with TCC > 2.5 %, for deforestation, the bar plots show the disagreement and agreement on the sign of local effects from the checkerboard and MW methods for both seasons and models. The disagreement and the agreement grid cells have been added up. The percentage values correspond to the disagreement/agreement relative to the total number of grid cells within the model. Each bar plot uses a different MW size for the MW method: starting from the left with a 5x5, 7x7 to a 9x9.

Knowing that both methods have some methodological uncertainties, a disagreement on the sign could still be seen as a similar local effect, as long as the magnitude difference is minimal, respectively, within the uncertainty bar. Therefore we calculate the magnitude difference of the grid cells with opposite local effects. The magnitude differences are distinguished into four intervals. The number of grid cells within the intervals relative to the total number of grid cells with opposite local effects is visible in Figure 16. Comparing the orange bars of all three MW sizes, shown in Figure 16, we note that for larger MW sizes, the number of sign disagreeing grid cells with a magnitude difference of more than 2.5 K between the methods decreases. This highlights, that not only the disagreement on the sign between both methods slightly decreases with a larger MW size, but also the extreme magnitude differences.

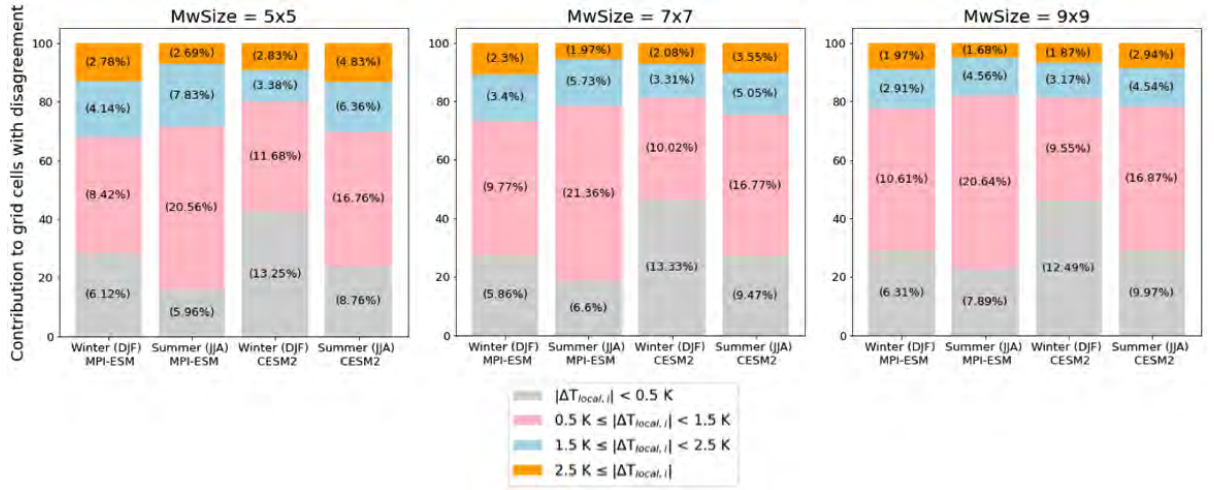

Figure 16: Only considering grid cells with TCC > 2.5 %, for deforestation, the bar plots show the magnitude differences between the checkerboard and MW methods for grid cells which disagree on the sign of local effects from both methods for both seasons and models. $|\Delta T_{\text{local, i}}|$ represents the magnitude difference of the local effect for grid cell i. The percentage values visible in the bar plots correspond to the number of grid cells within the magnitude interval relative to the total number of grid cells within the model. Each bar plot uses a different MW size for the MW method: starting from the left with a 5x5, 7x7 to a 9x9.

Increasing the MW size also improves the signal-to-noise ratio in the global median (Appendix D). The signal corresponds to the local effect, while the noise corresponds to the standard error of the corresponding slope. The decrease in standard error of the corresponding slope for larger MW sizes explains the increase in signal-to-noise ratio (Appendix E). A possible hypothesis could be that the larger the MW size, the more data points can be used for the linear regression calculation, thereby reducing the slope's standard error.

3.2.3 Impact of moving window size on local effect and signal separation

In this section, we do not examine the differences of the MW sizes relative to the checkerboard results, as investigated in Section 3.2.2, but we examine the local effect difference directly between the MW sizes. We calculate for each grid cell the standard deviation of the local effect across all three MW sizes. Figure 17 displays the standard deviations (on the y-axis) with the corresponding TCC values (on the x-axis) of all grid cells. Focusing on full deforestation (CROP - FRST) in Figure 17, we note huge standard deviations for TCC values higher than approximately 90 %. Such grid cells with very large TCCs show standard deviations across MW sizes that sometimes reach more than 3K. We conclude, that the choice of the MW size can significantly impact the local effect's magnitude and, thus, the signal separation.

For deforestation and afforestation (in Figure 17), such extreme standard deviations between the local effects of the three MW sizes are more rare compared to the full deforestation (CROP - FRST) scenario. Deforestation seems to show the least amount of strong standard deviations of the local effects between the MW sizes. This could be linked to the good agreement with the checkerboard method visible in Section 3.1.1. Further, it seems as if for deforestation and, more prominently, for afforestation, there is a linear increase in the standard deviation for increasing TCC under 80 %. A possible explanation could be the increase in TCC itself. Within the MW local effect calculation, the slope gets multiplied by a linearly increasing TCC if we move along the x-axis to higher TCC. This multiplication possibly explains the linear increase in standard deviation with higher TCC and why standard deviations for very low TCC are close to zero. Knowing the local effect calculation formula, it also becomes clear that the difference between the local effects of the three MW sizes comes from a difference in the slope.

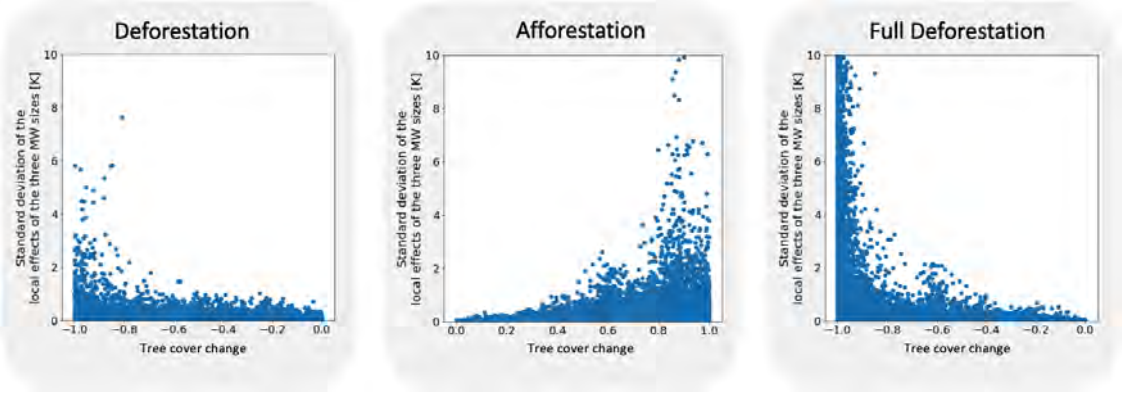

Figure 17: The scatter plots illustrate the standard deviations of the local effects of the MW sizes for each grid cell in relation to the tree cover change of each grid cell for all three scenarios. From the left, deforestation, afforestation, and full deforestation are shown for the annual mean on CESM2.

Examining the spatial distributions of grid cells with huge standard deviations of the MW local effect, visible in the (bottom row of Figure 18), we can identify the areas where the MW size should be carefully chosen. Comparing in Figure 18 the standard deviation maps (bottom row) to the corresponding TCC maps (top row), we note that large differences occur mostly in areas with large homogeneous TCCs. These large homogeneous TCCs can be seen most prominently for full deforestation in South America, Sub-Saharan region, or in Europe. For deforestation, high standard deviations are visible in the rainforest, where multiple grid cells close to each other show very large deforestation rates. For afforestation high standard deviations are visible in the same areas, such as Australia, the Sub-Saharan region, or Botswana, where the MW showed a higher warming effect for CESM2 than the checkerboard in Section 3.1.2.

It seems that if there is a lack of heterogeneity (i.e. different representations) of TCC in the MW, the MW method cannot calculate a reliable linear regression for certain grid cells and therefore the goodness of fit is poor.

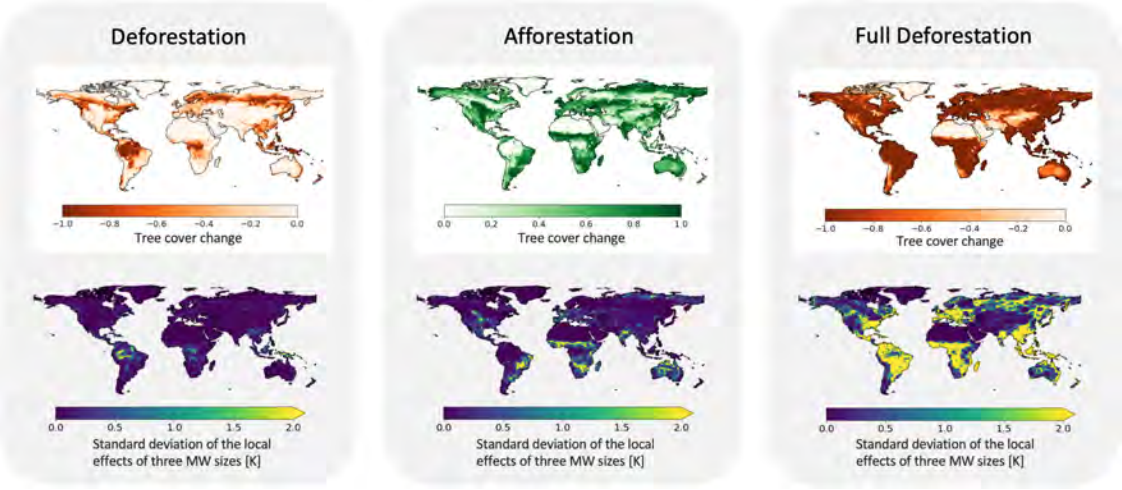

Figure 18: In the bottom row, the spatial distribution of the standard deviation of the local effects of the MW sizes can be seen for all three scenarios. From the left, deforestation, afforestation, and full deforestation are shown for the annual mean on CESM2. The tree cover change for each grid cell, used for the MW method, is shown in the top row for all three scenarios.

4 Discussion

In this study, we discovered that the two signal separation methods, the checkerboard and MW methods, show more similarity during winter than in summer. Thus, they agree more on showing the same TCC effect sign and presenting higher correlation coefficients during winter than summer. By investigating models of the Land-Use and Climate, Identification of Robust Impacts (LUCID) Lejeune et al. (2017) showed that out of the six different LUCID models studied, changes in albedo are strongest in the winter mean and changes in the latent heat are strongest in the summer mean. This seasonal variation agrees with our finding of good alignment between the signal separation methods regarding the snow-masking albedo effect in northern latitudes in winter. However, the concordance is less pronounced for summer latent heat change in tropical regions (Appendix H). By analyzing the reconstruction method (which is similar to the MW method), Lejeune et al. (2017) showed good agreement with the factorial method regarding albedo change and latent heat change. In our results, the agreement between the MW and the checkerboard method on latent heat change (summer mean) seems less good than for albedo (winter mean).

By comparing the factorial method to the reconstruction method, Lejeune et al. (2017) further found that the reconstruction method underestimates the TCC impacts compared to the factorial method. This underestimation is not visible if we look at the comparison between the MW method and the checkerboard method in this thesis. Best visible for CESM2 in summer (Figure 12), we see that the MW overestimates the TCC impact compared to the checkerboard method. The observed discrepancy in the MW method's tendencies relative to the checkerboard effect could stem from its limited capacity to robustly calculate the local effect for too similar TCCs within the MW, leading to a higher slope calculation.

Analyzing our model differences, we recognize, in accordance with De Hertog et al. (2023), strong disparity in the magnitude of the TCC local effects from the models. CESM2 and MPI-ESM have different submodels and thus different TCC maps, however, different partitioning of the available energy could also play a role. De Noblet-Ducoudré et al. (2012) and Lejeune et al. (2017) indicated that models partition the available energy differently between latent heat and sensible heat. The disparity between the TCC effects of different models highlights that precaution should be taken when comparing results from different models, as the resulting local effects are inherently model dependent.

Checkerboard method

In our findings, the local effects calculated by the checkerboard method for the summer mean of MPI-ESM showed a nearly global warming for deforestation and a nearly global cooling for afforestation. As this global warming and cooling are not visible for the local effect calculated by the MW method, this nearly zero correlation cannot be linked to the model's partitioning of the sensible and latent heat. Therefore, the difference must come within the underlying mechanism of the signal separation methods. Possible explanations could be that other (not tree-related) LCCs might be at the source of this discrepancy, leading to a miscalculation of the local effect from TCC

and, consequently, a difference in signal separation output.

According to the definition of the local effect, surface temperature change should solely result from local TCC. However, in some regions the calculations with the checkerboard method indicate strong local effects even with nearly absent TCC, contradicting the local effect's definition and suggesting potential spurious artefacts. Considering that the checkerboard method directly computes the nonlocal effect and derives the local effect by subtracting the nonlocal effect from the total effect, it is possible that some factors overlooked in the nonlocal effect calculation might have been mistakenly incorporated into the local effect. Thus, a deeper investigation of the nonlocal effects might be helpful in understanding these spurious artefacts.

Given the spurious artefacts identified in the local effects from the checkerboard and the predominantly warming (for deforestation) and cooling effect (for afforestation) for MPI-ESM summer, it is evident that obtaining a "clean" local effect from the checkerboard method might be more challenging than initially expected. We expected a clean and robust local effect since the dedicated ESM simulation setup was created to optimize the checkerboard signal separation. Further, De Hertog et al. (2023) have demonstrated that the annual mean local effect of the checkerboard method agrees with observational studies. However, seeing these predominantly strong responses for summer MPI-ESM in multiple scenarios and the spurious artefacts for very low TCC raises questions about the checkerboard's robustness.

MW method

Regarding the MW method, we have seen that if the MW contains mostly grid cells of very similar TCC, the standard deviation of the local effect calculated by different MW sizes becomes very large. This suggests that the MW method is unsuitable for robustly calculating the local effect if different TCCs are not represented well enough within the MW. This finding supports the requirement made by Lejeune et al. (2017), who required a minimum of three grid cells with low TCC and a minimum of three grid cells with high TCC for their reconstruction method to calculate the local effect. Further, if the ratio of high and low TCC grid cells was not around 0.5, Lejeune et al. (2017) increased the MW size, which shows the importance of a good representation of different TCCs. However, we found that even by increasing the MW size to 9x9, the ratio of around 0.5 between high and low TCC is often not given (Appendix F). This makes it difficult to find a high-to-low ratio-based rule from which the "best" corresponding MW size for each grid cell can be chosen. When the high-to-low ratio is not around 0.5, the local effect calculated by the MW method needs to be considered cautiously. Calculating the local effect across multiple MW sizes can be beneficial to grasp the local effect for such grid cells. It would offer insight into the standard deviation among the three MW size local effects, providing a sense of result stability. Nevertheless, when the conditions of a good representation of different TCCs within the MW were given, we were able to calculate a local effect similar to the one of the checkerboard method with the MW method. Examples of areas with a good representation of different TCCs are Alaska and South Siberia. For further research, it would be interesting to calculate the Pearson correlation coefficients for certain

areas to validate the importance of representing different TCCs within the MW size.

We used linear regression to calculate the relationship between the TCC and the BGP effect using the MW method. Thereby assuming, as in Lejeune et al. (2018), that the effect of TCC on surface temperature is linear. De Noblet-Ducoudré et al. (2012) also concluded that the amount of deforestation is almost linear to the effect of Land Use and Land Cover Change (LULCC). In Section 3.2.3, we observed a rather exponential behavior for the standard deviation of the local effects of the three MW sizes. This raises the question if the assumption of a linear relationship between the TCC and the BGP effect is the right choice. A possible alternative would be to use a polynomial regression, allowing for larger BGP effects for larger TCCs. Although a possible downside of this option might be the underestimation of the effect of low TCCs. Further, the MW approach is limited by the low amount of data points within the MW to fit on, making it more challenging to fit the relationship between TCC and BGP effects correctly.

The signal separation of the MW method includes the assumption of a homogeneous background climate. Chen and Dirmeyer (2020) highlight that this space-for-time assumption, where paired sites or grid cells within the MW have the same background climate, may not be applicable at broader scales due to significant differences in incoming radiation and cloud cover linked with deforestation. They further highlight differences between observations and model simulations in terms of atmospheric feedback for deforestation. Having seen in section 2.1 that the MW sizes (grid cell resolution) can be spatially large, we question how well the assumption of a homogeneous background climate holds and, because of that, how precisely the local effects can be calculated. As this assumption is key to a successful MW signal separation of the TCC effect, it must be carefully investigated whether the assumption holds or not.

Lastly, it has to be taken into consideration that Malyshev et al. (2015) and Winckler et al. (2019a) showed that the nonlocal effects dominate the global mean surface temperature response for BGP effects of deforestation. Thus, it is important to keep in mind that in this thesis, we only focused on the local effect, which can be overwhelmed by the atmospheric feedback (nonlocal effects) (Chen and Dirmeyer, 2020). To explore the magnitude of nonlocal effects in the context of deforestation and afforestation, the calculation of these effects can be found in Appendix G. Initially, the aim of this study was also to increase the understanding of the still rather uncertain behavior of the nonlocal effect. However, we investigated and compared only the local effect since the local effect can be easier compared to observations and cross-checked with results from other studies (Winckler et al., 2017; Lejeune et al., 2017, 2018; Duveiller et al., 2018; Chen and Dirmeyer, 2020; De Hertog et al., 2023). Technically, the nonlocal effect can be seen as the difference between the total and the local effect, and therefore, if we increase the confidence level of the local effect, we theoretically also increase the confidence level of the nonlocal effect. Nevertheless, future studies could and should further investigate the nonlocal effect.

Generally, as simplified methods, the signal separations underlie some assumptions and are subject

to limitations and uncertainties. For example, we assume that all external forcings are not affecting the total effect (showing a homogeneous background climate) and that the total effect on surface temperature is only due to local and nonlocal TCC. However, from other studies (Lejeune et al., 2018; Thiery et al., 2020), we know that geographical variables, such as elevation, affect the surface temperature. In other words, non-TCC components might be involved, which, if not considered, can wrongly attribute TCC to surface temperature change. Another example is the choice of a linear interpolation in the checkerboard method or the assumption in the MW method that the relationship between TCC and surface temperature change is linear.

All in all, investigating the assumptions and understanding their impact on the results are immensely important for trustworthy signal separation. We gain confidence in the methods when we gain confidence in the underlying assumptions used within the methods. Therefore, for future studies, the focus should not only lie in comparing the nonlocal and local effects to observations but also in investigating and understanding the method (and the assumptions) itself.

5 Conclusion

This study compared the TCC effects, calculated by the checkerboard and the MW method. This comparison was applied to the same dedicated LAMACLIMA experiments (using CESM2 and MPI-ESM) to find an agreement or middle ground between the two methodologies. This further enabled us to highlight circumstances where the methodologies indicate difficulties and disagreement. We compared the results of the signal separation methods (checkerboard and MW) for two seasonal means (summer and winter) and three scenarios (afforestation, deforestation, and full deforestation). The signal separation methods were selected as they can signal separate the total effect into the local and the nonlocal effect.

We discovered that the local effects calculated using the checkerboard and MW methods generally match well for deforestation. However, there is an exception during summer for the MPI-ESM model. For afforestation the correlation between the methods is lower compared to deforestation. Additionally, we discovered that there are seasonal variations, with higher correlation between signal separation methods during winter compared to summer.

Table 1 highlights the main findings of the two signal separation methods.

Methods	Findings
Moving Window	• Larger moving window size can improve slope calculation but with risk of smoothing out spatial variations • Importance of the representation of different TCCs within the moving window
Checkerboard	• Possible spurious artefacts for low LCC

Table 1: Summary of the findings for both signal separation methods.

Regarding the MW method, we found that the representation of different TCCs within the MW is key for a trustworthy calculation of the local effect. Based on this finding, a diverse representation of different TCCs within the MW is a crucial requirement to apply the MW method in future studies. Secondly, by increasing the MW size, we found that the likelihood of more diverse TCC within the MW increases and that it reduces the standard error of the slope. However, increasing the MW size comes with the risk of smoothing out spatial variations.

Regarding the checkerboard method, we encountered difficulties calculating the local effects in areas with nearly absent TCC. As we have not found clear explanations for these strong local effects in areas with nearly absent TCC, we assume that the visible effects could be spurious artefacts.

Based on these main findings, we can make a recommendation under which scenarios the specific signal separations should be used. These recommendations are highlighted in Table 2.

Methods	Recommendation
Moving Window	• For low LCC areas • Limited computational capacity (first estimate)
Checkerboard	• For homogeneous LCC areas • Good understanding of the nonlocal effect

Table 2: Recommendations under which scenarios the signal separation methods should be chosen.

The MW method should be chosen when the focus lies on areas with low TCC. The reasoning behind this recommendation comes from the spurious artefacts found for the checkerboard method and the more trustworthy calculation for low TCC with the MW method. Secondly, the MW method should be chosen when there is only limited computational capacity. The MW method can directly be applied on data sets containing irregular information on TCC and its total effect, while for the checkerboard method, a computationally more expensive dedicated ESM experiment is needed. Thus, a third benefit of the MW method is that it can be used to get a quick first estimate of the local and nonlocal effects.

The checkerboard method should be preferred when investigating areas with homogeneous TCC, as the MW cannot calculate a robust linear regression in these areas. Further, the checkerboard method should be chosen, due to its setup, if the interest lies in a good understanding of the nonlocal effect.

In conclusion, this study represents a detailed comparison between the MW and checkerboard methods on the same data set, providing a foundational understanding of their comparability. We showed that both methods agree relatively well on the signal separation under certain conditions. Therefore, the choice of which method to choose for future projects should be based on the research question and the available computational capacity. Further research can build upon our findings and provide deeper insights into how these methods perform across diverse data sets. For example, it would be helpful to understand why the methods agree best in winter and especially in CESM2 and less in summer and MPI-ESM. Such investigations would contribute valuable information to guide the selection of the most suitable method for future projects.

Looking ahead, our findings open the door to promising future research and exploration areas. Be it a more in-depth comparison of the nonlocal effects, analyzing the underlying assumptions, or, more generally, applying the signal separation methods on other simulation experiments.

6 Appendix

A) Visualization of the interpolation bias for both models

In Section 2.2.2 we saw that the checkerboard method linearly interpolates every second grid cell to receive a signal for the full world map (i.e. every grid cell). Here, we investigate how large this interpolation bias can be. Therefore, we started with the original total surface temperature effect of CESM2 for full deforestation (CROP - FRST), whereby we only used every second grid cell. This gave us the world map in the top left of Figure 19. As in the checkerboard method, we linearly interpolated to have a world map with signals in every grid cell (top right of Figure 19). Subtracting from the original total effect map (bottom left of Figure 19) the interpolated world map (top right of Figure 19), we can see the interpolation bias in every second grid cell (bottom right of Figure 19). On the right-hand side, we calculated the latitudinal mean of the interpolation bias, showing that interpolation bias lies within $\pm$ 0.25 K.

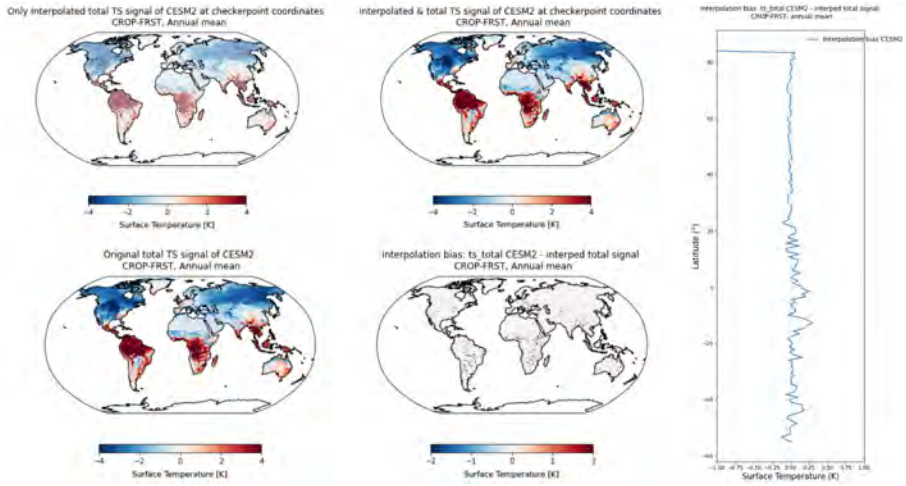

Figure 19: Visualization of the interpolation bias embedded within the checkerboard method. In this Figure, the interpolation bias is calculated for CESM2, annual mean, CROP - FRST, total TCC effect on surface temperature.

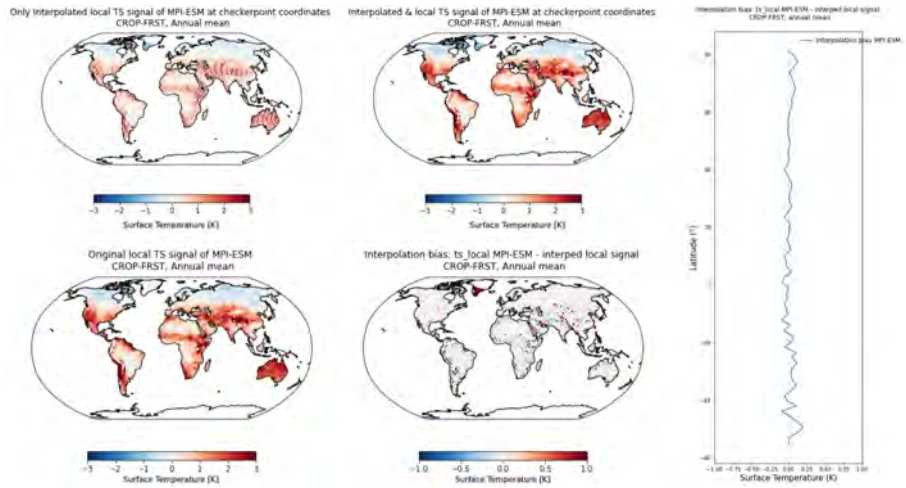

Figure 20: Visualization of the interpolation bias embedded within the checkerboard method. In this Figure, the interpolation bias is calculated for MPI-ESM, annual mean, CROP - FRST, total TCC effect on surface temperature.

B) TCC maps used for the MW method and LCC maps used for the checkerboard method. Adapted Figures from De Hertog et al. (2023).

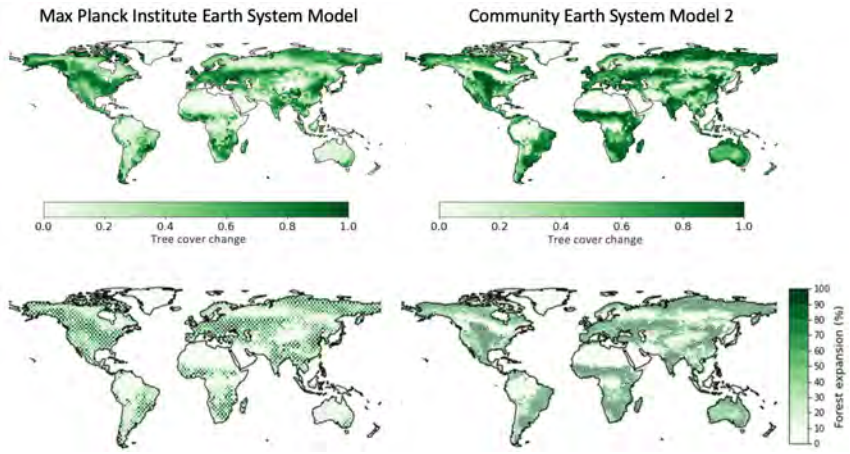

Figure 21: In the top row the two TCC maps used within the MW method are visible for both methods. In the bottom row the two LCC maps used within the checkerboard method are visible for both methods. For FRST-CTL.

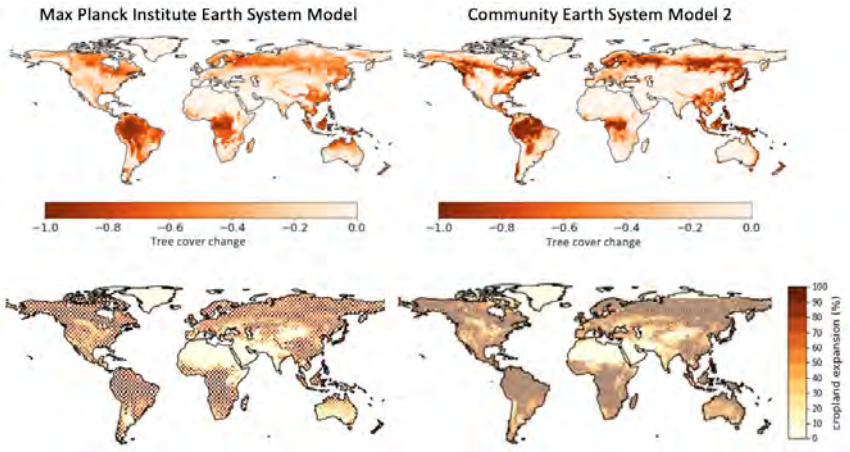

Figure 22: In the top row the two TCC maps used within the MW method are visible for both methods. In the bottom row the two LCC maps used within the checkerboard method are visible for both methods. For CROP-CTL.

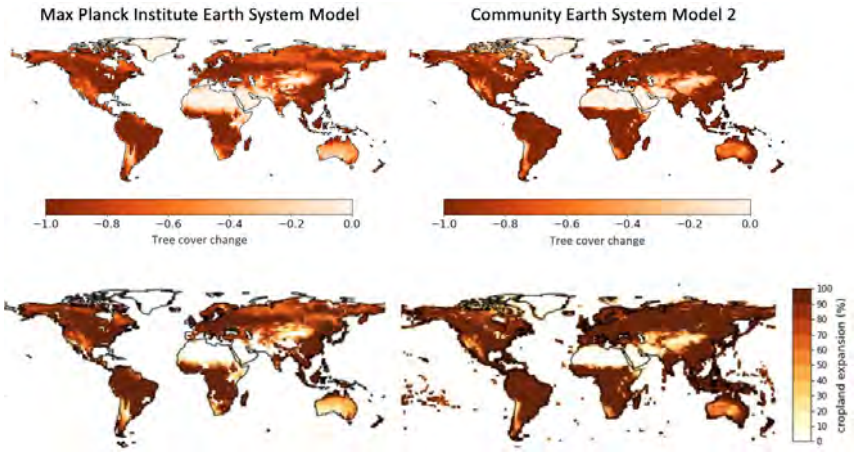

Figure 23: In the top row the two TCC maps used within the MW method are visible for both methods. In the bottom row the two LCC maps used within the checkerboard method are visible for both methods. For CROP-FRST.

C) Barplots for FRST - CTL focusing on the grid cells where both methods show opposite local effects

In Section 3.2.2 we analyzed the benefits of choosing a larger window size for the deforestation scenario. Here the same has been done but for the afforestation and full deforestation scenario

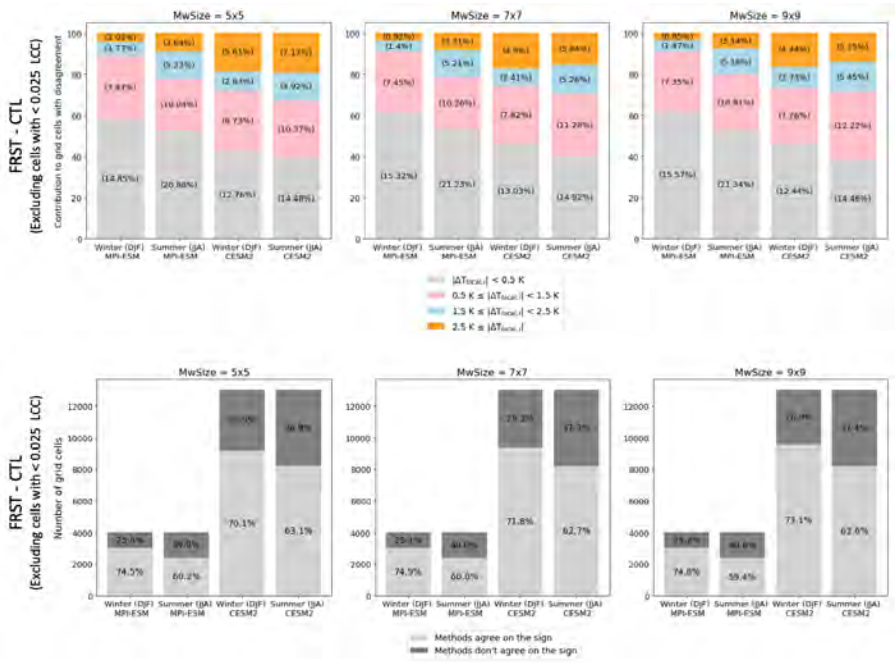

Figure 24: For FRST - CTL. In the bottom row the relative disagreement between the methods can be shown for all three MW sizes. Top row shows the magnitude difference intervals of the grid cells with opposite local effect.

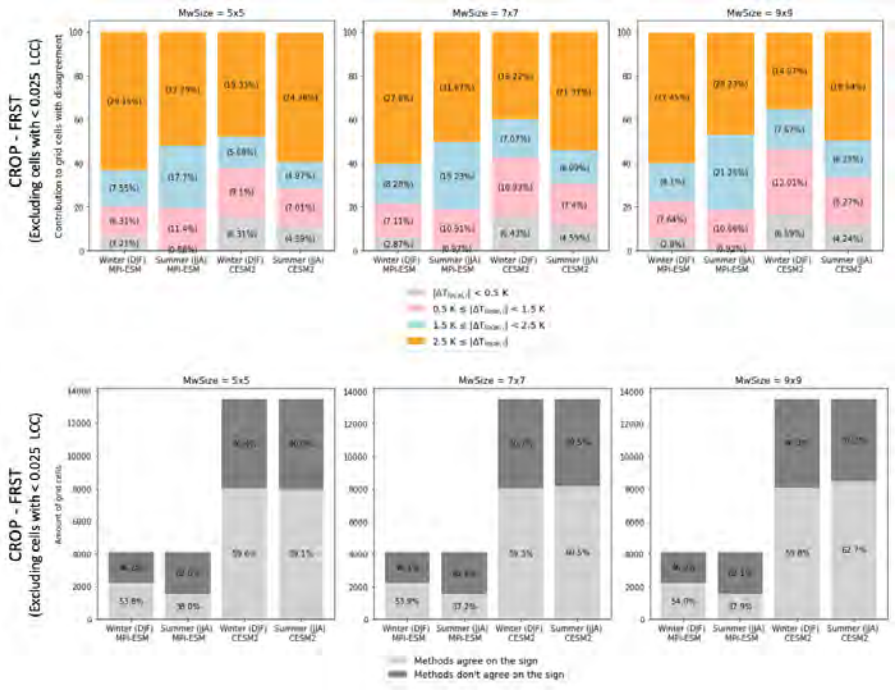

Figure 25: For CROP - FRST. In the bottom row the relative disagreement between the methods can be shown for all three MW sizes. Top row shows the magnitude difference intervals of the grid cells with opposite local effect.

D) Signal-to-noise ratio calculated for both models

Visualization of the in Section 3.2.2 mentioned increase in signal-to-noise ratio for higher MW sizes.

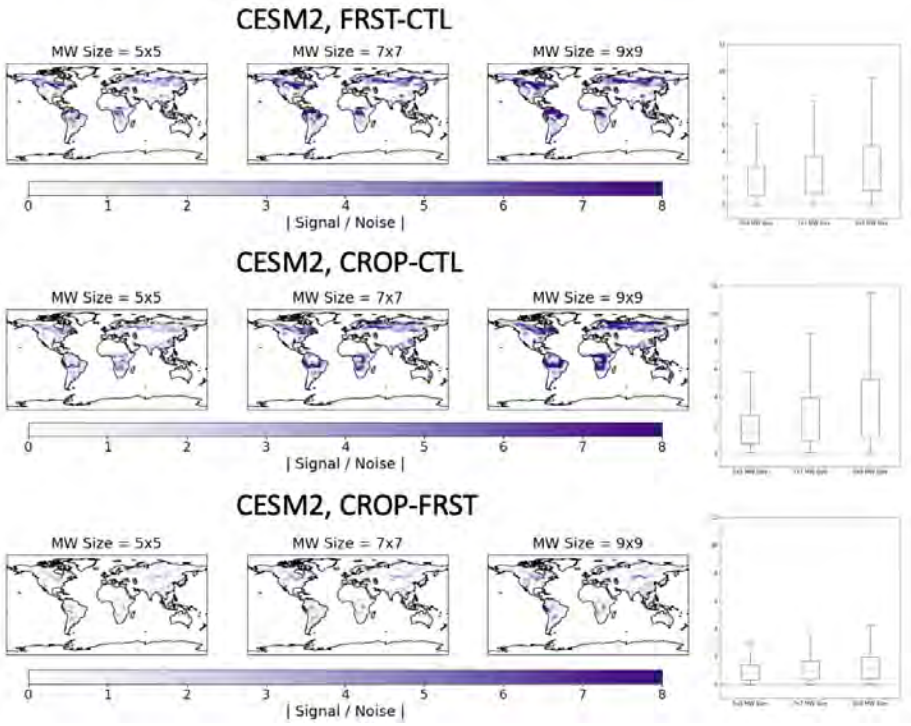

Figure 26: Calculation of the signal-to-noise ratio for CESM2. Whereby the Signal corresponds to the slope and the noise to the standard error of the slope. Here the absolute value of the signal-to-noise value is shown.

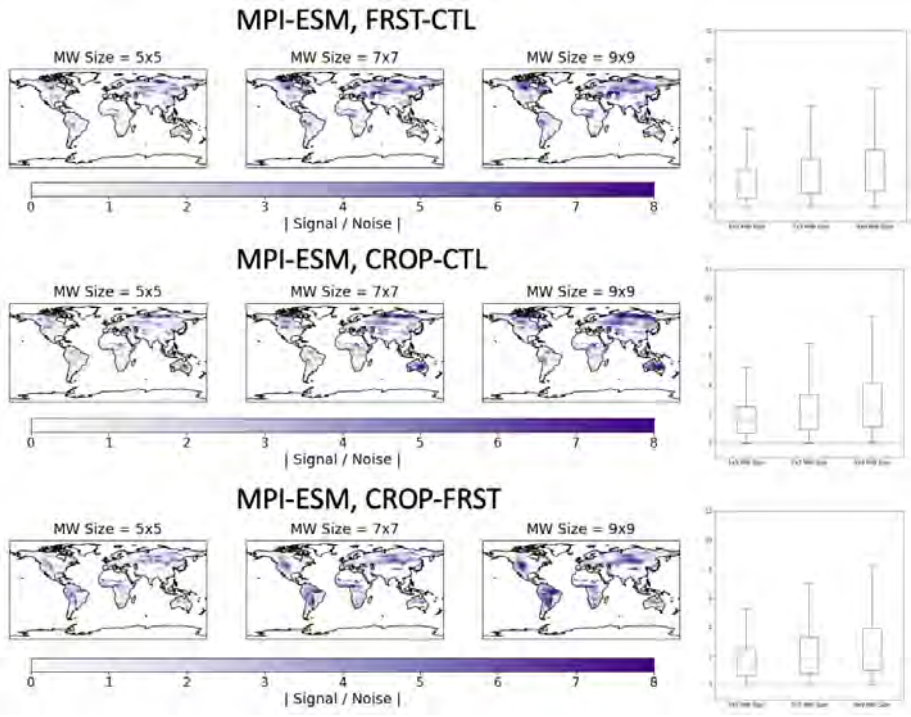

Figure 27: Calculation of the signal-to-noise ratio for MPI-ESM. Whereby the Signal corresponds to the slope and the noise to the standard error of the slope. Here the absolute value of the signal-to-noise value is shown.

E) Standard error of the corresponding slope (noise) calculated for both models

Visualization of the in Section 3.2.2 mentioned decrease in standard error of the corresponding slope for higher MW sizes.

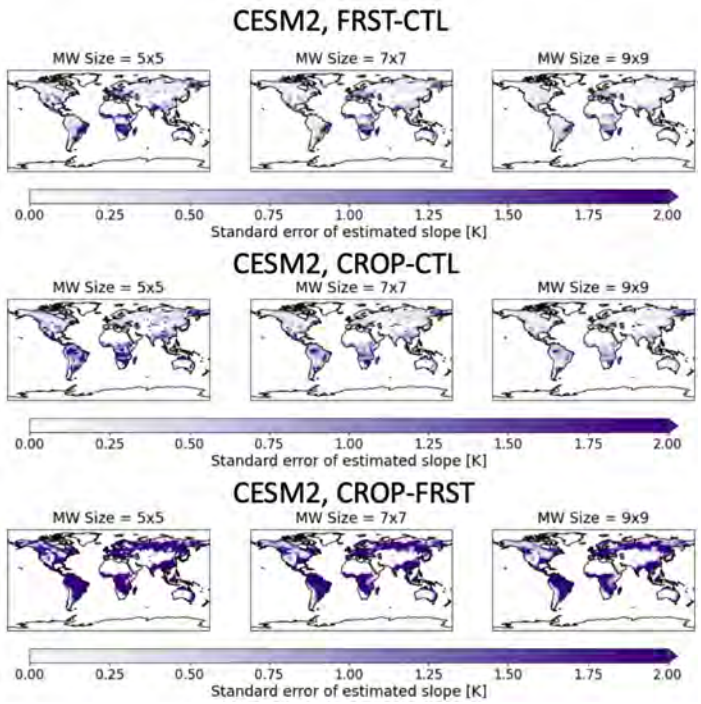

Figure 28: Calculation of the signal-to-noise ratio for CESM2. Whereby the Signal corresponds to the slope and the noise to the standard error of the slope. Here the absolute value of the signal-to-noise value is shown.

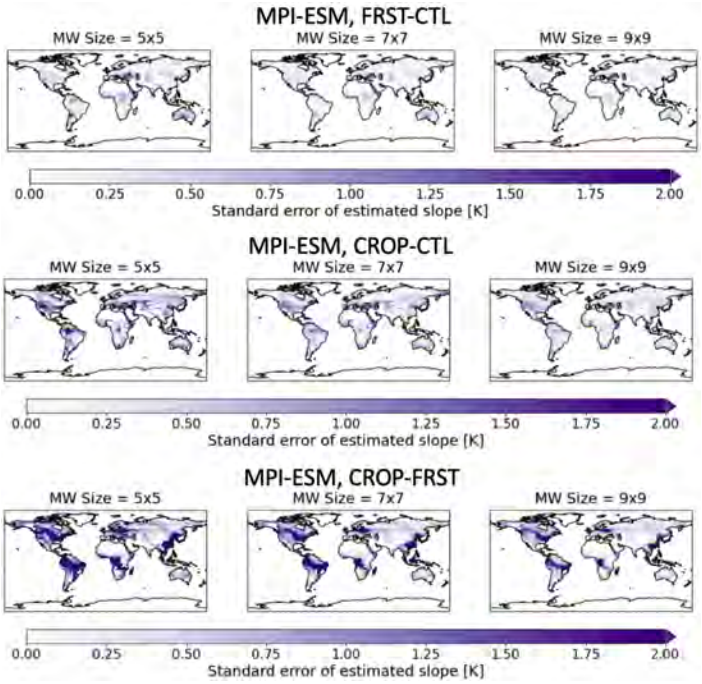

Figure 29: Calculation of the signal-to-noise ratio for MPI-ESM. Whereby the Signal corresponds to the slope and the noise to the standard error of the slope. Here the absolute value of the signal-to-noise value is shown.

F) Analysis of the TCC ratio within the MW

By analyzing the different results from Figure 30 and 31 we can see, that also for the 9x9 MW size the TCC ratio is not around 0.5. This, is something that should be kept in mind, when choosing a MW size based on the TCC ratio within the MW. For future investigation it would be interesting to see if the grid cells where the TCC ratio of around 0.5 can been achieved with a maximum of a 9x9 MW size, show higher agreement with the checkerboard effects than the remaining grid cells. If yes, this could potentially be a way to find out where the MW can be applied best.

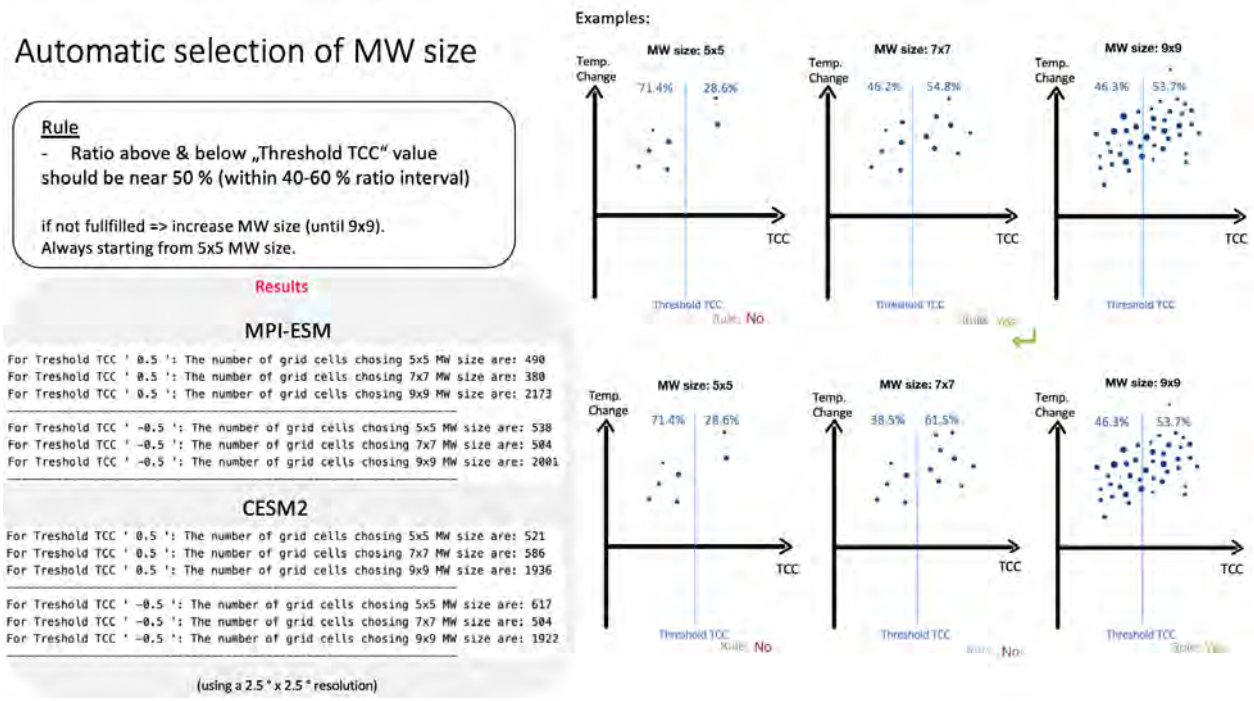

Figure 30: Based on a rule, which is explained in the top left, we choose the smallest grid cell which still fulfills the rule. This should help insure locality and the diverse representation of TCC within the MW. On the right two examples are illustrated, how the rule works. In the gray box in the bottom left, the results are shown. If the Threshold TCC is negative, the automatic selection has been applied on the deforestation scenario. The positive Threshold TCC show the results for afforestation. This calculation has been applied on a 2.5 ° x 2.5 ° resolution for both models.

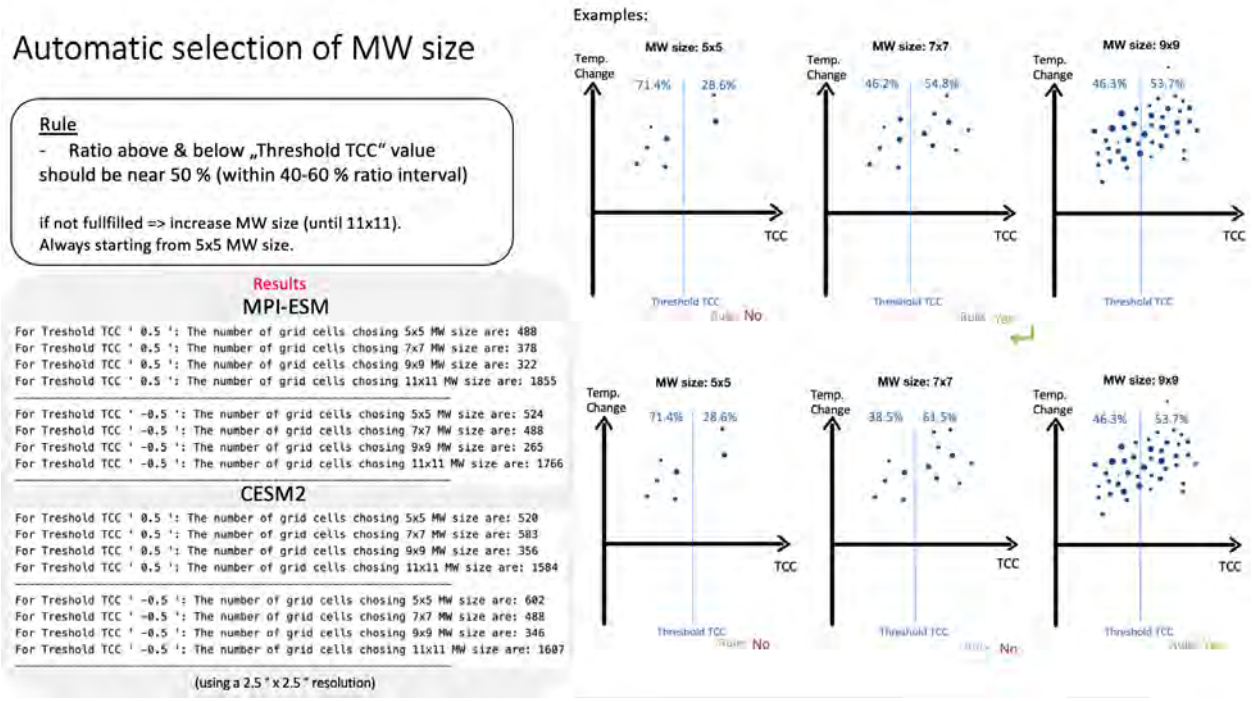

Figure 31: Based on a rule, which is explained in the top left, we choose the smallest grid cell which still fulfills the rule. This should help insure locality and the diverse representation of TCC within the MW. On the right two examples are illustrated, how the rule works. In the gray box in the bottom left, the results are shown. If the Threshold TCC is negative, the automatic selection has been applied on the deforestation scenario. The positive Threshold TCC show the results for afforestation. This calculation has been applied on a 2.5 ° x 2.5 ° resolution for both models.

G) Nonlocal effects on surface temperature for afforestation and deforestation

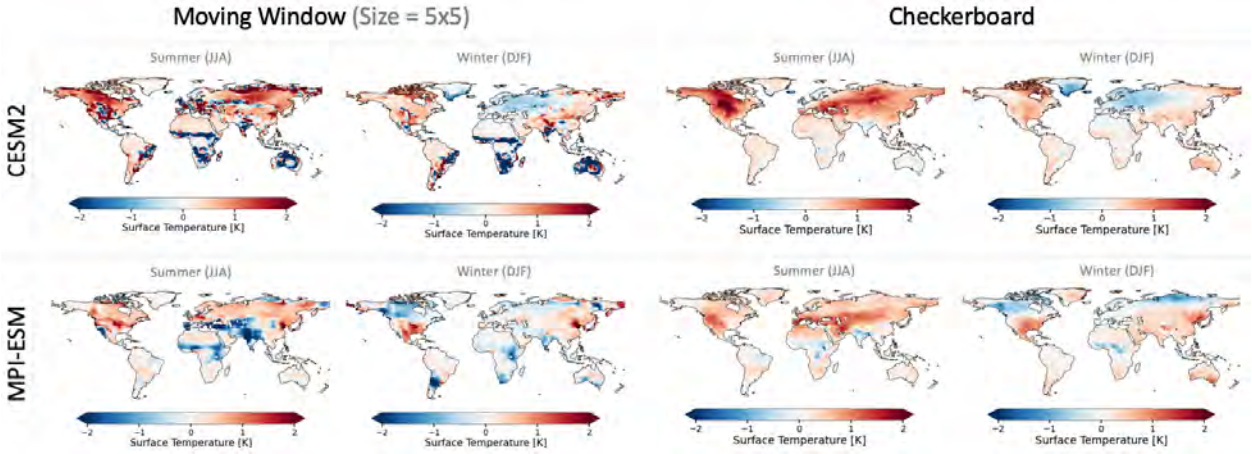

Figure 32: In this Figure, the nonlocal effects on surface temperature are shown for afforestation (FRST-CTL), for both methods, models, and seasonal means

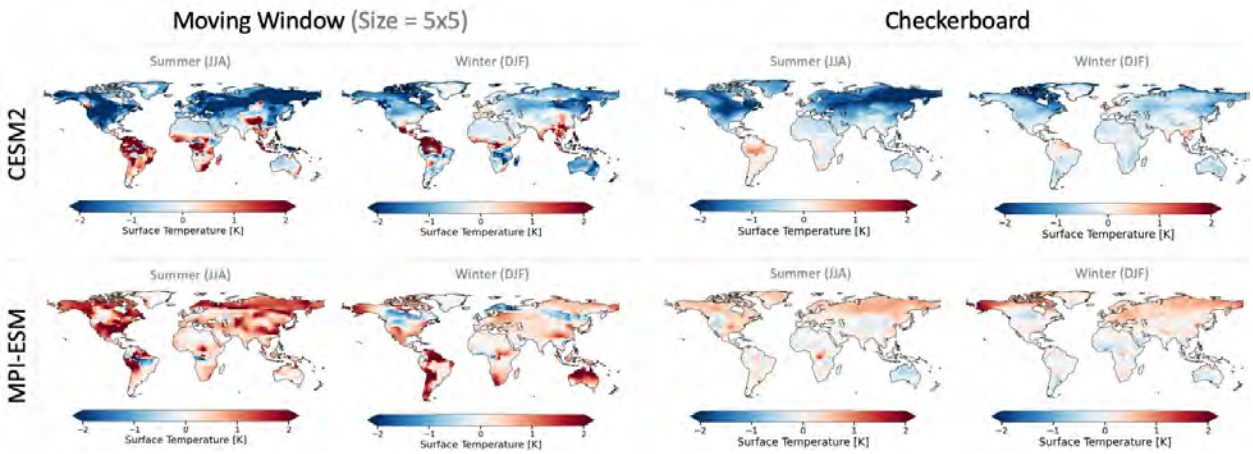

Figure 33: In this Figure, the nonlocal effects on surface temperature are shown for deforestation (CROP-CTL), for both methods, models, and seasonal means

H) Comparison of the MW size local effects to the checkerboard effect for latitude bands

Here, the local effects on surface temperature have been visualized in boxplots within the latitude band. On the left side of the Figures the deforestation effects (corresponding TCC visualized on x-axis) are visualized. On the right side of the Figures the afforestation effects (corresponding TCC visualized on x-axis) are visualized. These plots highlight that for a latitude band average, the 9x9 MW size does not always show the most similar effect than the checkerboard method. Further we see increased method differences (dotted lines) for higher TCC changes. These were first insights of the standard deviations between the local effects of the three MW sizes and their possible underlying reasons. Note that for MPI-ESM summer, the checkerboard method shows a global warming for deforestation and a global cooling for afforestation, this seems to be different to what we expect due to other studies. Thus, MPI-ESM summer should be considered carefully.

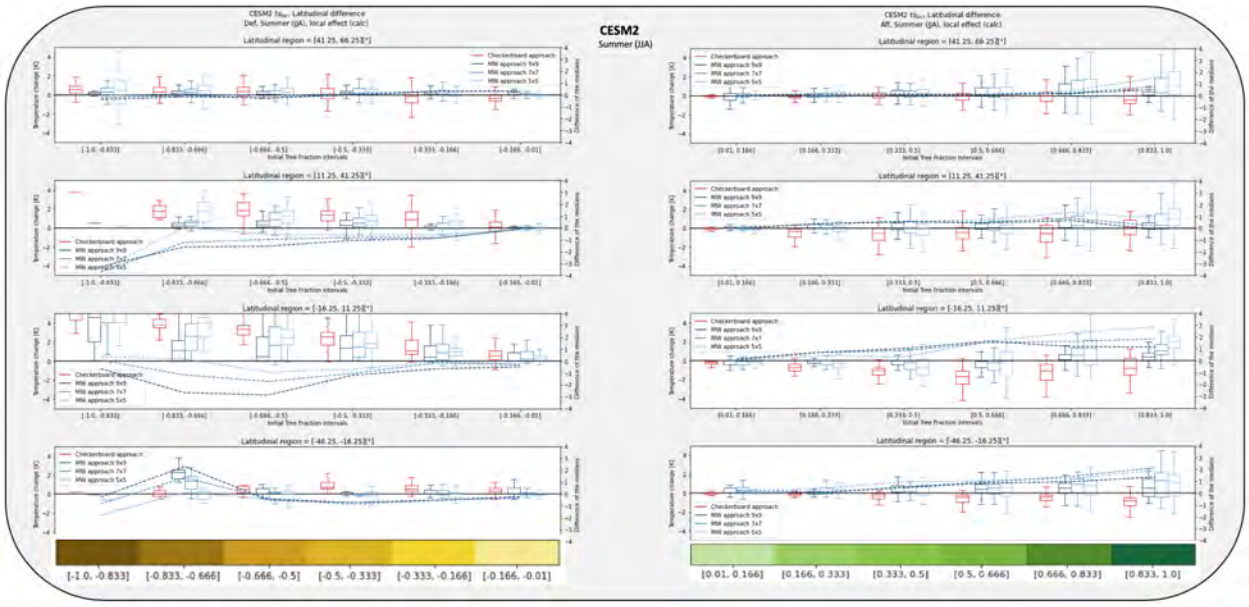

Figure 34: On the right side, we see the local effect of deforestation for four latitude bands. On the left side, we see the local effect of afforestation for four latitude bands. The different colored boxplots show the locla effects of the different MW sizes and the checkerboard method. The dotted lines highlight the difference of the MW size to the checkerboard response. The calculations are made for CESM2, summer.

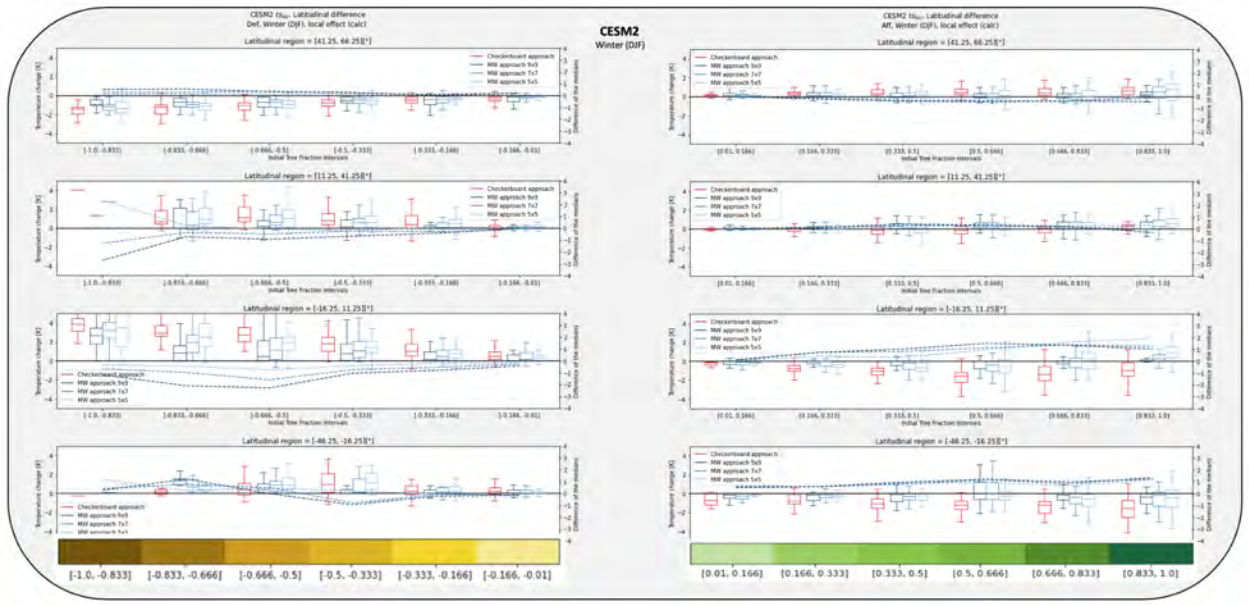

Figure 35: On the right side, we see the local effect of deforestation for four latitude bands. On the left side, we see the local effect of afforestation for four latitude bands. The different colored boxplots show the locla effects of the different MW sizes and the checkerboard method. The dotted lines highlight the difference of the MW size to the checkerboard response. The calculations are made for CESM2, winter.

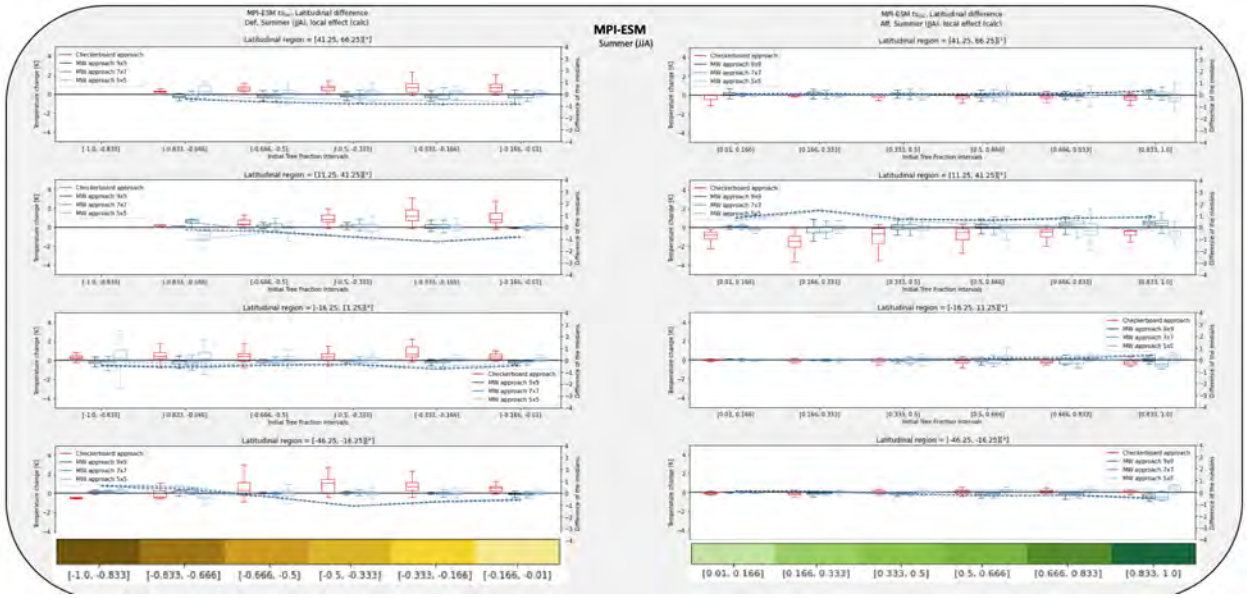

Figure 36: On the right side, we see the local effect of deforestation for four latitude bands. On the left side, we see the local effect of afforestation for four latitude bands. The different colored boxplots show the locla effects of the different MW sizes and the checkerboard method. The dotted lines highlight the difference of the MW size to the checkerboard response. The calculations are made for MPI-ESM, summer.

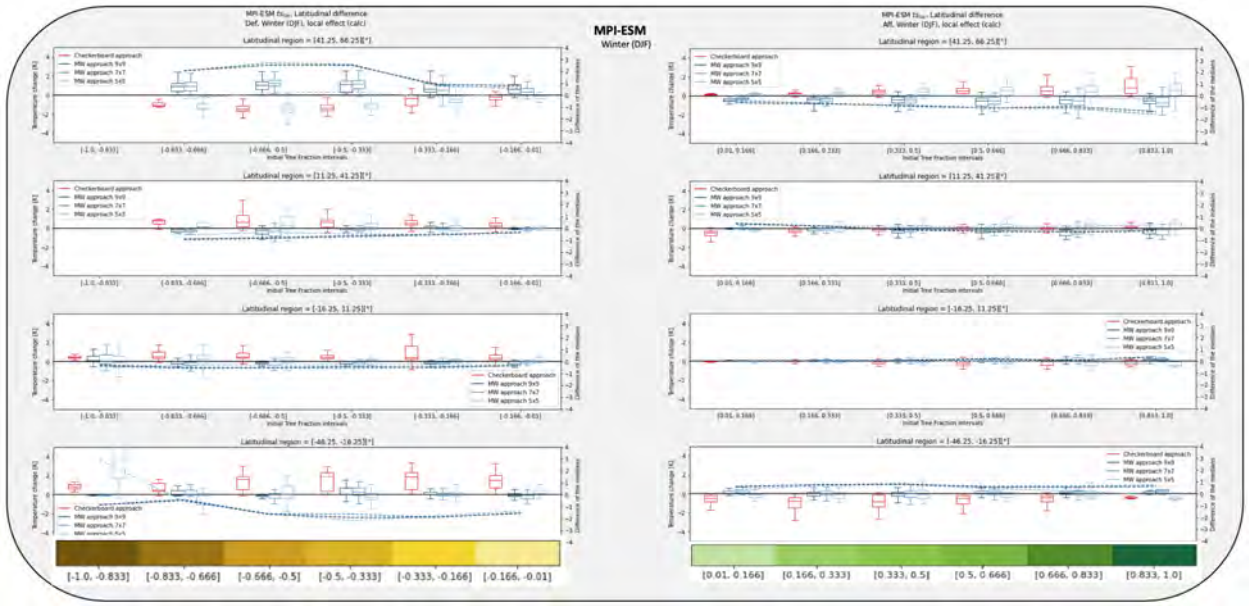

Figure 37: On the right side, we see the local effect of deforestation for four latitude bands. On the left side, we see the local effect of afforestation for four latitude bands. The different colored boxplots show the locla effects of the different MW sizes and the checkerboard method. The dotted lines highlight the difference of the MW size to the checkerboard response. The calculations are made for MPI-ESM, winter.

I) Spatial visualization of the local effects of grid cells with TCC < 2.5 % for the afforestation scenario, showing possible spurious artefacts of the checkerboard method.

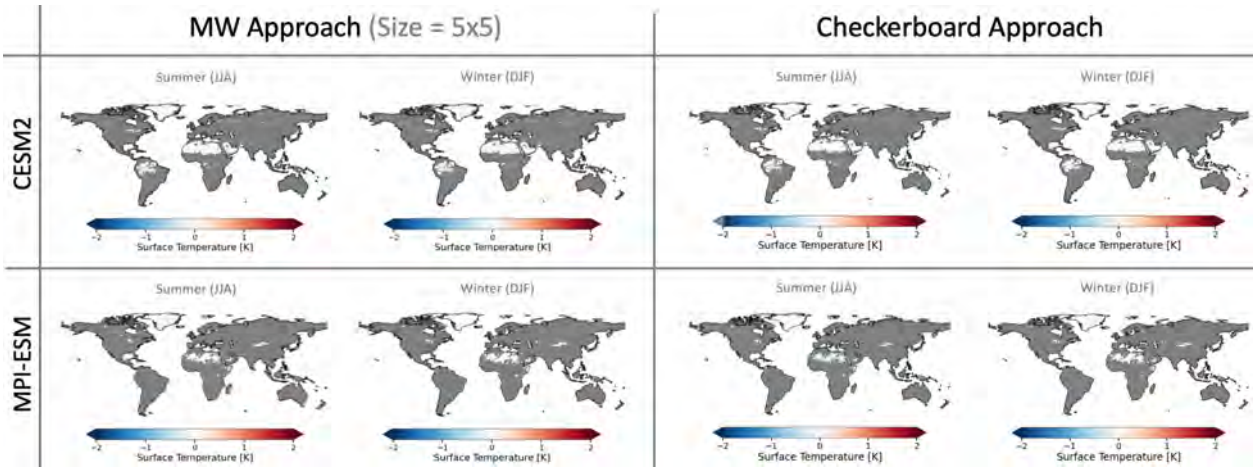

Figure 38: Spatial visualization of the local effects of grid cells with TCC < 2.5 % for the afforestation scenario. The grid cells in gray are masked-out grid cells with TCC larger than 2.5 %. The columns represent the local effect for deforestation of a specific signal separation method. We differentiate between summer mean and winter mean. A 5x5 MW size has been used.

Figure reference

Figure 1 (right side): World Resources Institute and Global Forest Watch; 2023; Gross tree cover loss by ecozone; Viewed at 16.08.2023; <https://research.wri.org/gfr/forest-extent-indicators/forest-loss>

Data and code availability

We created a function which allows to apply the MW method on data with following input: total effect data, variable change (e.g., TCC) data. This function, along with the other scripts created for this thesis, is accessible on the Wyss Academy for Nature GitHub page. For the data please contact Rafael Stähli, Steven De Hertog, Shruti Nath, Édouard Davin or Quentin Lejeune.

Acknowledgement

I want to say a big thank you to some really important people who helped me a lot with my master's thesis. First, I want to thank Édouard. You made me feel welcome in the research team and provided me with a good balance between freedom and guidance. I'm really thankful for the support, input, ideas and the openness. Then, there's Quentin. I want to thank you for helping me shape the thesis. You sometimes challenged me with good questions and suggestions, but they always ended up significantly improving the quality of my work. So, thank you for that. Lastly, I want to give a really big thank you to my advisors Shruti and Steven. I can't thank you enough for investigating an hour (or more) every week for a year to guide me through this thesis. These meetings were not only super helpful, but they also kept me motivated and were a space to enjoy the work.

Overall, I'm grateful because this year was more than just work. Working with all of you has been really enjoyable, and I appreciate it a whole lot.

7 References

Abe M. , Takata K. , Kawamiya M. and Watanabe S. (2017). Vegetation masking effect on future warming and snow albedo feedback in a boreal forest region of northern eurasia according to miroc-esm. *Journal of Geophysical Research: Atmospheres*, 122 (17), 9245–9261.

Boisier J. , de Noblet-Ducoudré N. , Pitman A. , Cruz F. , Delire C. , Van den Hurk B. , Van der Molen M. , Müller C. and Voldoire A. (2012). Attributing the impacts of land-cover changes in temperate regions on surface temperature and heat fluxes to specific causes: Results from the first lucid set of simulations. *Journal of Geophysical Research: Atmospheres*, 117 (D12).

Bonan G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *science*, 320 (5882), 1444–1449.

Bright R. M. (2015). Metrics for biogeophysical climate forcings from land use and land cover changes and their inclusion in life cycle assessment: a critical review. *Environmental Science & Technology*, 49 (6), 3291–3303.

Chen L. and Dirmeyer P. A. (2016). Adapting observationally based metrics of biogeophysical feedbacks from land cover/land use change to climate modeling. *Environmental Research Letters*, 11 (3), 034 002.

Chen L. and Dirmeyer P. A. (2020). Reconciling the disagreement between observed and simulated temperature responses to deforestation. *Nature communications*, 11 (1), 1–10.

Cohn A. S. , Bhattarai N. , Campolo J. , Crompton O. , Dralle D. , Duncan J. and Thompson S. (2019). Forest loss in brazil increases maximum temperatures within 50 km. *Environmental Research Letters*, 14 (8), 084 047.

Collins W. , Bellouin N. , Doutriaux-Boucher M. , Gedney N. , Halloran P. , Hinton T. , Hughes J. , Jones C. , Joshi M. , Liddicoat S. and others (2011). Development and evaluation of an earth-system model–hadgem2. *Geoscientific Model Development*, 4 (4), 1051–1075.

Danabasoglu G. , Lamarque J.-F. , Bacmeister J. , Bailey D. , DuVivier A. , Edwards J. , Emmons L. , Fasullo J. , Garcia R. , Gettelman A. and others (2020). The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, 12 (2), e2019MS001 916.

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

Davin E. L. , Seneviratne S. I. , Ciais P. , Olioso A. and Wang T. (2014). Preferential cooling of hot extremes from cropland albedo management. *Proceedings of the National Academy of Sciences*, 111 (27), 9757–9761.

Davin E. L. , Rechid D. , Breil M. , Cardoso R. M. , Coppola E. , Hoffmann P. , Jach L. L. , Katragkou E. , de Noblet-Ducoudré N. , Radtke K. and others (2020). Biogeophysical impacts of forestation in europe: first results from the lucas (land use and climate across scales) regional climate model intercomparison. *Earth System Dynamics*, 11 (1), 183–200.

De Hertog S. J. , Havermann F. , Vanderkelen I. , Guo S. , Luo F. , Manola I. , Coumou D. , Davin E. L. , Duveiller G. , Lejeune Q. and others (2023). The biogeophysical effects of idealized land cover and land management changes in earth system models. *Earth System Dynamics*, 14 (3), 629–667.

De Noblet-Ducoudré N. , Boisier J.-P. , Pitman A. , Bonan G. , Brovkin V. , Cruz F. , Delire C. , Gayler V. , Van den Hurk B. , Lawrence P. and others (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. *Journal of Climate*, 25 (9), 3261–3281.

Duveiller G. , Hooker J. and Cescatti A. (2018). The mark of vegetation change on earth's surface energy balance. *Nature communications*, 9 (1), 1–12.

Hirsch A. L. , Prestele R. , Davin E. L. , Seneviratne S. I. , Thiery W. and Verburg P. H. (2018). Modelled biophysical impacts of conservation agriculture on local climates. *Global change biology*, 24 (10), 4758–4774.

Hirsch A. L. , Wilhelm M. , Davin E. L. , Thiery W. and Seneviratne S. I. (2017). Can climate-effective land management reduce regional warming? *Journal of Geophysical Research: Atmospheres*, 122 (4), 2269–2288.

IPCC AR6 SPM V. , Masson-Delmotte, Zhai P. , Pirani A. , Connors S. , Péan C. , Berger S. , Caud N. , Chen Y. , Goldfarb L. , Gomis M. and others (2021). Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. *Climate Change 2021: The Physical Science Basis*.

IPCC SRCCL P. R. , 2019: Shukla, Skeg J. , Buendia E. C. , Masson-Delmotte V. , Pörtner H.-O. , Roberts D. , Zhai P. , Slade R. , Connors S. , Van Diemen S. and others (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. *In press*.

Kumar S. , Dirmeyer P. A. , Merwade V. , DelSole T. , Adams J. M. and Niyogi D. (2013). Land use/cover change impacts in cmip5 climate simulations: A new methodology and 21st century challenges. *Journal of Geophysical Research: Atmospheres*, 118 (12), 6337–6353.

Lejeune Q. , Davin E. L. , Gudmundsson L. , Winckler J. and Seneviratne S. I. (2018). Historical deforestation locally increased the intensity of hot days in northern mid-latitudes. *Nature climate change*, 8 (5), 386–390.

Lejeune Q. , Seneviratne S. I. and Davin E. L. (2017). Historical land-cover change impacts on climate: Comparative assessment of lucid and cmip5 multimodel experiments. *Journal of Climate*, 30 (4), 1439–1459.

Malyshev S. , Shevliakova E. , Stouffer R. J. and Pacala S. W. (2015). Contrasting local versus regional effects of land-use-change-induced heterogeneity on historical climate: analysis with the gfdl earth system model. *Journal of Climate*, 28 (13), 5448–5469.

Mauritsen T. , Bader J. , Becker T. , Behrens J. , Bittner M. , Brokopf R. , Brovkin V. , Claussen M. , Crueger T. , Esch M. and others (2019). Developments in the mpi-m earth system model version 1.2 (mpi-esm1. 2) and its response to increasing co2. *Journal of Advances in Modeling Earth Systems*, 11 (4), 998–1038.

Meehl G. , Covey C. , Latif M. , McAvaney B. , Mitchell J. and Stouffer R. (2007). Ipcc standard output from coupled ocean-atmosphere gcms. *Lawrence Livermore National Laboratory Program for Climate Model Diagnosis and Intercomparison*.

Meier R. (2021). On the biogeophysical consequences of forestation. *Doctoral thesis*, URL https://doi.org/10.3929/ethz-b-000484696.

Pitman A. J. , de Noblet-Ducoudré N. , Cruz F. , Davin E. L. , Bonan G. , Brovkin V. , Claussen M. , Delire C. , Ganzeveld L. , Gayler V. and others (2009). Uncertainties in climate responses to past land cover change: First results from the lucid intercomparison study. *Geophysical Research Letters*, 36 (14).

Pongratz J. , Reick C. , Raddatz T. and Claussen M. (2008). A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, 22 (3).

Pongratz J. , Reick C. , Raddatz T. and Claussen M. (2010). Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change. *Geophysical Research Letters*, 37 (8).

Qu X. and Hall A. (2007). What controls the strength of snow-albedo feedback? *Journal of Climate*, 20 (15), 3971–3981.

Schwaab J. , Davin E. L. , Bebi P. , Duguay-Tetzlaff A. , Waser L. T. , Haeni M. and Meier R. (2020). Increasing the broad-leaved tree fraction in european forests mitigates hot temperature extremes. *Scientific reports*, 10 (1), 1–9.

Seneviratne S. I. , Wartenburger R. , Guillod B. P. , Hirsch A. L. , Vogel M. M. , Brovkin V. , van Vuuren D. P. , Schaller N. , Boysen L. , Calvin K. V. and others (2018). Climate extremes, land–climate feedbacks and land-use forcing at 1.5 c. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376 (2119), 20160 450.

Spracklen D. , Baker J. , Garcia-Carreras L. , Marsham J. and others (2018). The effects of tropical vegetation on rainfall. *Annual Review of Environment and Resources*, 43 (1), 193–218.

Thiery W. , Visser A. J. , Fischer E. M. , Hauser M. , Hirsch A. L. , Lawrence D. M. , Lejeune Q. , Davin E. L. and Seneviratne S. I. (2020). Warming of hot extremes alleviated by expanding irrigation. *Nature communications*, 11 (1), 290.

Winckler J. , Lejeune Q. , Reick C. H. and Pongratz J. (2019a). Nonlocal effects dominate the global mean surface temperature response to the biogeophysical effects of deforestation. *Geophysical Research Letters*, 46 (2), 745–755.

Winckler J. , Reick C. H. and Pongratz J. (2017). Robust identification of local biogeophysical effects of land-cover change in a global climate model. *Journal of Climate*, 30 (3), 1159–1176.

Winckler J. , Reick C. H. , Luyssaert S. , Cescatti A. , Stoy P. C. , Lejeune Q. , Raddatz T. , Chlond A. , Heidkamp M. and Pongratz J. (2019b). Different response of surface temperature and air temperature to deforestation in climate models. *Earth System Dynamics*, 10 (3), 473–484.

Declaration of consent

on the basis of Article 30 of the RSL Phil.-nat. 18

Name/First Name: Stähli Rafael

Registration Number: 17-107-509

Study program: Master of Science in Climate Sciences

Bachelor ☐ Master ☑ Dissertation ☐

Title of the thesis:

Local versus nonlocal effects of land cover changes:
Comparing methods of signal separation for temperature responses

Supervisor:

Prof. Dr. Édouard Davin

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of 5 September, 1996 is authorized to revoke the title awarded on the basis of this thesis.
For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with future theses submitted by others.

Place/Date Biel, 17.08.2023

Signature