Otryakhin et al. applied downscaling to the meteorological input data for LPJ-GUESS and assessed the impact of using high- versus low-resolution climate data on model outputs. They statistically evaluated the differences introduced by the orographic downscaling by comparing mountainous and relatively flat regions, demonstrating that differences in model outputs due to climate data resolutions are more pronounced in mountainous areas. The statistically robust approach presented in the manuscript provides valuable insights not only for researchers applying downscaling techniques, but also for those using coarse-resolution climate data. For example, it provides information for evaluating whether statistical errors arising from the coarseness of climate data fall within the range of uncertainties caused by other factors, such as model parameterization or variability in input data.

I recommend this manuscript for publication in GMD; however, I request a major revision due to several concerns. While the writing is mostly logical and clear, there are sections where the lack of detail makes it difficult to fully understand the methodological flow from approach to results. Although the statistical procedures are described in detail, the downscaling method, the core aspect of the study, is not explained clearly. Individual comments are provided below. I believe that the revised manuscript will be suitable for publication in GMD.

We thank the reviewer for the positive assessment of our work and valuable suggestions, which we think have contributed to the increase of quality and clarity of our manuscript. We address those comments below and use green text for our replies. For quotations of the text we use *Italic* font, while the newly introduced amendments are in ***bold Italic***.

Major comments:

L4-9

The abstract seems too simple. It should elaborate more on the unique aspects of this study, the insights gained, and its advantages. Specifically, the differences in climate variables caused by elevation gradients and their effects on the model should be clearly described. The introduction is similar in this regard. It would benefit from more detailed information that is linked to the experimental design. Since the analysis to investigate the effects of elevation differences is well-executed, it would be better to explicitly explain how high-resolution climate data influences dynamic vegetation models.

We augmented the abstract by adding insights into the downscaling method used and the key results of the study:

Using the CHELSA algorithm, we create an elevation-informed high-resolution climate dataset for a domain encompassing the European Union. ***Distinctive features of this algorithm include orographic nature of formation of precipitation, a negative derivative of temperatures with respect to elevation, and also, detailed consideration of shadowing and exposure of the terrain to the Sun in computations of solar radiation. We design a custom experiment protocol and use it to perform LPJ-GUESS simulations on both resolutions.*** *Comparative analysis reveals significant systematic discrepancies between the two resolutions.* ***In mountainous areas, all of the considered output variables show statistically significant differences. In particular, carbon pools are smaller on the high resolution, with the total carbon pool being 37-39% smaller. Furthermore, we***

quantify the extent to which the under-representation of orographic climate variation affects regional predictions across the European Union. This is expressed as a difference in the total value, which ranges from -3.8% for the net ecosystem productivity to 2.9% for the litter and soil C pools. These values are found to be comparable to differences caused by miss-representation of water bodies and shorelines on the low resolution.

In the introduction, we added an overview on modern downscaling methods and put our study into a broader context of research of vegetation response to high-resolution climate forcings as follows:

Downscaling methods can be applied to overcome the mismatch between coarse global climate projections, and the fine-resolution needs of impact models (Karger et al., 2023). At present, terrain-informed downscaling could be executed by either regional climate models for dynamical downscaling, or by topogaphic downscaling methods. Algorithms of the first class are very precise as they directly model physical and chemical processes in the atmosphere. This comes with the disadvantage of being computationally slow, which makes their application on large scales challenging (Giorgi et al., 2009; Sørland et al., 2021; Schär et al., 2020). Topogaphic downscaling uses mechanistic relationships to turn low-resolution climatologies into high-resolution ones based on knowledge of terrain. These relationships are quite simple and do not capture atmospheric effects unrelated to topography, so this class of algorithms fails to represent some small-scale effects, such as convective precipitation (Karger et al., 2021). Also, topographic downscaling is characterized by less computational complexity than that of dynamical downscaling. The two best performing and widely known topogaphic methods are CHELSA (Karger et al., 2017, 2021, 2023) and PRISM (Daly et al., 1994, 1997). For this study we choose CHELSA to perform downscaling for two reasons. First, we need a computationally fast algorithm as we examine a region covering the whole of Europe. Second, out of the two best performing topogaphic downscaling methods, CHELSA provides the easiest way to interpret the results from the point of view of atmospheric physics.

Regarding explanation of how high-resolution climate data influences dynamic vegetation models, establishing the exact mechanisms of how high-resolution climate changes the vegetation dynamics was not one of the objectives of our study. In this paper, we rigorously prove that the vegetation dynamic does change when resolution increases, and we discuss what processes may be involved in this, but we do not analyze which processes play significant roles in that and which ones do not. Thus, we do not mention these processes in the introduction, since it is not a major part of our study.

Section 2.1

It would be better to include a justification for the selection of CHELSA. Clarifying the differences from dynamic downscaling methods would help make the objectives of this study clearer.

Please, see our comment above.

The version of CHELSA used in the study should be specified.

The CHELSA version used is V2.1 (we have added this information at the beginning of the revised Sect. 2.1).

L16

The authors mention local extreme weather events, but is the downscaling approach used in this study capable of reproducing such events? For instance, how accurately can CHELSA represent localized extreme precipitation caused by topographic effects, and what specific types of events can it capture?

Yes, downscaled data represents extreme events better than low-resolution ones. Consider a 50-by-50 km gridcell with a narrow tall mountain chain. Due to the mechanism of orographic precipitation, wind pushes moisture from a large area towards the top of this chain, so that a large portion of water precipitates in a small area. On the low resolution, precipitation per square meter may be just slightly above average, but on the high resolution it may be extreme because of the size of the low-resolution gridcell. Another effect happens to the temperature. Since it is averaged on the low resolution, we cannot observe late spring frosts in high altitudes, which will be present on the high resolution. We decided not to go into details of this topic since local weather extremes have not been studied in the CHELSA setting. We leave it for future studies.

Precipitation in methods

What is the spatial resolution of the satellite data? In the manuscript, some information such as climate variables is summarized in tables. It may be helpful to include this information in a table as well. Overall, the description of the downscaling methods is ambiguous. In particular, for precipitation and shortwave radiation, additional details are needed to ensure reproducibility. It is necessary to include a clear explanation of how low-resolution data are distributed across the high-resolution grid cells (e.g., Eq. 24 in Karger et al., 2023).

We checked once again the main reference for our version CHELSA V2.1 and also confirmed with the CHELSA team, that satellite data is not used in this version. The associated text was removed. We largely added details for the CHELSA method to improve the general understanding of it. See the updated Sections 2.1.2 Precipitation and 2.1.3 Surface downwelling shortwave radiation (RSDS) in the appendix at the end of this document. We note that, the CHELSA algorithm is reproducible in any case, as we provided the reference to the main article on the algorithm, and also included a link to the exact software implementation. In this work, we strive to give the working understanding of the downscaling method, rather than a thorough recipe for replicating it. An interested reader can follow the references, read file Readme, download the source code and data, and study the fine details of the algorithm.

L68-76

It is difficult to understand from the presented equations how the downscaling from low to high resolution is actually performed.

We added a more detailed description of the method in the revised manuscript. See the updated Section 2.1.3 Surface downwelling shortwave radiation (RSDS) in the appendix at the end of this document. We refer to the original article on CHELSA V2.1 (Karger et.al, 2023) for more information.

L88-89

Is Equation (6) essential? The statistical testing is described in detail, whereas the downscaling method lacks sufficient explanation, leading to an imbalance in the presentation.

We believe that equation (6) is required to unambiguously define variables in equation (7). We enlarged the description of the downscaling methods, please see our responses above.

L101

It is unclear whether the “50-100 observations” refer to the number of grid cells at the downscaled or raw resolution. This should be stated more explicitly. Also, is this number limited by computational constraints? In Fig. 8, for instance, a simulation is performed at the European scale, so a more detailed explanation would be helpful.

We made clarifications in Sect. 2.2 explaining that “50-100 observations” refers to the regional averages of values computed over the period 1850 – 2014. We made a note that this limitation arises from the computational constrains. Whenever we do simulations at the European scale, we make it only 1 time on each resolution.

To clarify this point, the text was modified as follows:

In the context of studies of large regions over the historical period 1850–2014, LPJ-GUESS simulations are computationally demanding especially on the high resolution. Because of this, *generating samples that contain more than 50–100 observations* ***of averages in the Alpine region is*** *a challenging task* ***on both 0.5◦ and 0.083(3)◦ resolutions.***

L118-119

Since the manuscript includes fire on–off experiments, it should include a more detailed explanation of the fire-related processes to enhance clarity and reproducibility.

We expanded on the description of the fire model within the context of the LPJ-GUESS model description. The fire model is composed of two submodels: the SIMFIRE model to estimate burned area annually, and the BLAZE model to simulate wildfire ignition stochastically and calculate CO2 and N fluxes. The text now reads:

(L118) Wildfires are ***simulated explicitly with the SIMFIRE-BLAZE submodel (Knorr et al., 2014, 2016; Rabin et al., 2017). The potential burned area for each gridcell is calculated annually as a function of land cover type, meteorological information, and the fraction of absorbed photosynthetically-active radiation (FAPAR) as a proxy for vegetation cover. This is then used to model ignition stochastically, and calculate combustion rates and associated carbon and nitrogen fluxes. A comprehensive description of the fire submodel is available in Molinari et al. (2021).***

Fig. 2

It might be helpful to provide more information, such as what i represents and the sample size.

We changed the caption as follows:

Figure 2. Scheme of computations in the ensemble experiment. Here, X is the average of values at the end of the computation period 1850–2014 in the region, lr and hr are the indicators of the low and high resolution correspondingly, i = {1, .., 50} is the experiment id, μ's are the sample mean estimates.

4.2 Results

It would be helpful to illustrate the characteristics of both the high-resolution and low-resolution climate data, for example using maps. This would make it easier to understand how downscaling affects climate variables, especially in regions with significant elevation differences.

Overall, the results are presented primarily as statistical information, but it would be helpful to also show the spatial differences visually using graphs or maps.

That is a very interesting topic. As a matter of fact, it is very hard to fully represent climate as a map because daily data is very dynamic--- there are lots of differences from day to day. Our historical dataset include about 60 000 days on the European scale, so visualization of every day is impossible. Averaging over time periods would smooth out this variation, so climates on both resolutions would look quite alike. An even stronger smoothing effect would happen after spatial averaging, so that analysis of time series is not possible this way. At the same time, this daily dynamics is what makes a difference in vegetation simulations. Therefore, we have a classical big data problem: we know the mechanics behind the dataset, but cannot inspect the data thoroughly. Partially, this problem was investigated in the CHELSA papers we cite (Karger et.al 2017, 2021, 2023). These works explain the mechanisms behind the differences of climatologies on both resolutions and provide daily maps as examples. Of course, they also do not provide an exhaustive list of differences since they proceed from known causes to evidence, and cannot catch differences of unknown causes. We do not do the same investigation here, because we would like to avoid repetition.

In any case, the results are presented as maps in Sect. 5 with the images provided in the Supplementary material. Sect. 4 has a different purpose and is primarily focused on the rigorous proof that there are significant differences between resolutions.

The statistical explanation of the errors arising from differences in resolution was very clear. Has the study examined whether using downscaled climate data improves the agreement between model simulations and observed fluxes?

If so, a brief description of this result would help strengthen the justification for using downscaled climate data in the modeling framework.

In this study, we have not investigated whether using downscaled climate data improves the agreement between model simulations and observed fluxes. Although this is a very important task, our study concerns itself with evaluating the differences between modeled outputs on high and low resolution. We note that the downscaled climate is closer to climate observations, so if the model output on low resolution was closer to observed fluxes, that would suggest that the model needs recalibration or revision. In any case, we leave this for future research.

Table 3

Aren’t the units of fluxes kgC m^{-2} yr^{-1}? Isn’t stored carbon expressed on an area basis?

The units on this table refer to regional aggregates and averages of the variables, which is the focus of our study. Hence, the units are not on a per-area basis, even if the raw model output is. We now clarify this in the table's caption:

"List of ecosystem variables modeled by LPJ-GUESS that were included in the experiment. These include carbon fluxes (...), carbon pools (...), water cycle variables (...), and vegetation structural variables (...). ***The units refer to regional aggregates (for all variables except FPC and LAI) and regional averages (for FPC and LAI) of the selected variables."***

Are the characteristic outputs of a DGVM, such as vegetation transitions, not evaluated in this study?

In this study we focused on evaluating the likely magnitude of the impact of resolution on aggregated diagnostics. The spatial PFT distribution was consistent between the two simulations, but a full evaluation of species distribution, including a comparison with observations and with results of previous versions of the model, will be the object of future work.

Fig3:

Roff showed remarkable difference between experiments in Fig. 3(b). Roff exhibited a notable difference between experiments in Fig. 3(b). Could you clarify the cause of this discrepancy?

The discrepancy is only in relative terms. The Roff values in the different experiments are actually very similar in absolute terms in the control region, ranging from 42.5 to 49 mm/y (see tables Therefore, a small difference between experiments of a few mm per year amounts to a large difference in relative terms. In the study region, the differences are much larger (~40mm/year), but smaller in relative terms because Roff in those regions is much larger. We draw attention to the small absolute Roff difference between the experiments by including the following text (L216):

“The differences between ensemble means in the study and control regions, δS and δC, are now both negative (Fig. 3). ***Runoff shows the largest relative discrepancy with respect to the previous experiment, but the difference in absolute terms is very small.*** *This sign switch...”*

L234

The discussion on the contribution of fire appears somewhat abrupt. Could you clarify why fire is considered to have a significant impact? Additionally, if fire events are infrequent, wouldn't ensemble averaging tend to smooth out their influence?

Fire is a rare but destructive event, so ensemble averaging does not necessarily have to smooth out its influence. Because of this, we wanted to check whether it played an important role.

We added the following paragraph in section Discussion on the contribution of fire to LPJ-GUESS results:

“The effect of fire on simulation results was found to be somewhat important, but not as strong as those of non-conservative properties of CHELSA and differences in climate due to orography.

The effect includes 2 parts. First, since ignition is stochastic, the presence of fire module is supposed to increase variation of the simulation results. Comparison of the standard deviations in Tables 6 and 7 shows that this effect does not play a significant role. Second, fire is a rare but destructive event which introduces changes in the potential vegetation structure. This could be one of the reasons why we see more variables with statistically indistinguishable muC_hr and muC_lr in the uCH/NoFire experiment than in the uCH/Fire one. In the study region on the high resolution, ignition is expected to occur more in valleys, which are warmer and drier than mountain tops, thus the effect of reduced vegetation in mountainous areas should be decreased in the uCH/NoFire experiment. However, from Fig. 3 we see that the influence of fire on vegetation in the study region is negligible compared to the influence of orography-induced climate difference."

Is geographical bias a particularly important and non-negligible source of uncertainty for the processes simulated by LPJ-GUESS?

Our results in Table 8 show that geographical bias is 3%-4% on the European scale, which is comparable to the climate-response bias (0.6%-3.8%). Together, these 2 sources constitute the total bias of up to 7%, which by far exceeds the standard deviation of the sample for almost all output variables. Therefore, this total bias might be a significant confounding factor in future studies involving statistical tests on samples of DGVM outputs. Also, please see our response under the question about the carbon budget calculations.

L265:

How were delta(cli) and (geo) calculated?

Please, see the appendix attached to this file for the details of this calculation. We will also include it as an appendix in the revised version of the manuscript.

In carbon budget calculations, the proportion of land cover within each grid cell is usually taken into account, so the error in the climate response would appear to be the more important factor.

We thank the reviewer for this observation. Scaling by land cover fraction is a very good approximation in gridcells that have both a fraction of water and small altitude variability. This is not always the case (as, eg., in the northern parts of the coast of Norway). We agree, however, that rescaling would work for most shoreline gridcells. A criterion for wether to downscale a particular gricell based only on elevation variability (independently of whether the gridcell in question contains water) would completely address this problem. We have expanded our discussion by including the following paragraph, where we mention this issue in the context of a broader discussion of the impact of resolution on model outputs:

(L270) Earlier work by Müller and Lucht (2007) showed little impact on model results when running the LPJ DGVM between 10◦ and 0.5◦, at 0.5◦ intervals, suggesting that a resolution of 0.5◦ is still too coarse to account for relevant effects of spatial heterogeneity. Our study suggests that the impacts of resolution on the modeled output, linked to the influence of orography on the input climate, become noticeable at higher resolutions. The relative importance of these effects depends

strongly on the focus region. Europe-wide simulations show an impact of resolution on aggregated ecosystem pools and fluxes of ∼ 3%, likely smaller than the uncertainty derived from the spread in climate forcings by different GCMs (see, e.g., Schaphoff et al., 2006; Morales et al., 2007; Schurgers et al., 2018). By contrast, these differences increase up to ∼ 46% in an Alpine region. ***Additional bias may result from poor representation of shorelines and small inland water bodies, but this effect could be mitigated by scaling the model output by the land-cover fraction in the affected gridcells. In areas of low variability in surface elevation, the difference between LPJ-GUESS outputs at different resolutions is much smaller and may be safely ignored in calculations involving regional averages of ecosystem variables. For this type of studies, one could optimize the resource requirements of the simulations by using a coarser resolution in areas with low elevation variability.***

We now also mention this point in the summary:

(L323) We studied systematic differences between high-resolution LPJ-GUESS simulations, forced with the new dataset, and low-resolution simulations. We found that low-resolution simulations are systematically biased. Two main sources of bias were identified: (a) bias associated to the non-linear response of the model to orographical climate variability, and (b) bias associated to the poor representation of coastlines and inland water bodies on a coarse grid. ***While the latter may be mitigated by rescaling the output by the land cover fraction in the affected gridcells, reducing the climate-response bias requires a finer grid resolution.***

L297 “correlations”

While I can infer the intended meaning, it would be better to explain it in more concrete terms.

In the real world, climate variables are correlated with each other. For example, at points where light is obstructed, the temperature is lower than that at neighboring points with no obstruction. Analogously, a spot with significant amount of precipitation would be colder and darker than the same spot with no precipitation.

CHELSA processes all climate variables independently of each other, possible correlations between variables that might exist in the physical world are not factored in by the algorithm. These correlations, however, might be built-in in more complex algorithms, and will likely be captured by dynamical downscaling, because it simulates the full physics of the system. To clarify this point, we modified the text as follows:

(L296) *In the context of climate change mitigation, correlations between different climate variables might influence relevant modeled variables (Zscheischler et al., 2019).* ***To give an example of mechanisms responsible for these correlations, we notice that at points where light is obstructed, the temperature is lower than at neighboring points with no obstruction. Analogously, a spot with a significant amount of precipitation would be colder and darker than the same spot without precipitation.*** *Such correlations are not built into univariate methods like CHELSA but can be captured by dynamical or multivariate downscaling methods.*

L278-290

The discussion lacks sufficient consideration of the model processes. While nonlinear responses are mentioned, it remains unclear how the model processes and the downscaled climate inputs interact

and what specifically leads to the nonlinear responses. Is the influence of climate variables other than temperature not addressed in the discussion?

We agree with the reviewer that the discussion between lines 278-290 focus almost exclusively in the impact of temperature differences on productivity, although the redistribution of precipitation in the high-resolution grid is also mentioned. We suggest adding the following text to highlight the influence of radiation and precipitation on the modeled processes.

[L289] "The interplay between these factors will depend on the specific region being simulated, which emphasizes the complexity of the model's response to orographical and climate drivers. ***There are many other modeled processes that respond non-linearly to climate forcings. Leaf-level photosynthesis shows a saturating (as opposed to linear) response to absorbed photosynthetically-active radiation when not limited by RuBisCo production (see Haxeltine and Prentice, 1996, for a discussion of the scaling of leaf-level photosynthesis to canopy-level productivity). Soil water transport follows a power law of available water content, which in turn depends on the amount of rainfall (see Gerten et al. 2004). The amount of radiation reaching the forest floor, which determines potential establishment of new saplings, obeys an exponential law that depends on the forest canopy's LAI (Monsi and Saeki, 1953, 2005). The decay rate of C in the different soil carbon pools is a non-linear function of soil temperature (driven by air temperature in the model) and soil water content (which depends non-linearly on precipitation rate, as mentioned above; see description of the carbon cycle submodel in Smith et al., 2014).***

L300-315

The proposed testing protocol in this section lacks specificity and its necessity is questionable. The statistical tests already presented in methods are sufficient to serve as reference information for other future studies. If a new approach is to be proposed, it would be better presented in text rather than as equations.

We agree with this point of view, and we have significantly simplified the end of the section by removing the proposed testing protocol and mathematical notation, while leaving only short textual description of the proposed studies. The text was modified as follows:

These methods are, however, generally more complex, and might require intensive use of computational resources. Therefore, it might be of interest to find systematic differences between simulations forced by the different methods. ***This could be done with the help of the methodology presented in Sect. 2.2 and 4. A similar setup could also be employed to investigate systematic differences originating from alternative model configurations. For example, one could assess whether the modeled impacts of two different forest managing strategies on regional carbon sinks are significantly different from each other.***

Minor comments:

L45 "(3)"

That is likely a typographical error.

Not at all. It is a common notation for a periodic decimal. E.g., 1/3=0.3(3).

References:

Gerten, Dieter, Sibyll Schaphoff, Uwe Haberlandt, Wolfgang Lucht, and Stephen Sitch. “Terrestrial Vegetation and Water Balance—Hydrological Evaluation of a Dynamic Global Vegetation Model.” Journal of Hydrology 286, no. 1 (2004): 249–70. https://doi.org/10.1016/j.jhydrol.2003.09.029.

Haxeltine, A., and I. C. Prentice. “A General Model for the Light-Use Efficiency of Primary Production.” Functional Ecology 10, no. 5 (1996): 551–61. https://doi.org/10.2307/2390165.

Karger, Dirk Nikolaus, Olaf Conrad, Jürgen Böhner, et al. “Climatologies at High Resolution for the Earth’s Land Surface Areas.” Scientific Data 4, no. 1 (2017): 170122. https://doi.org/10.1038/sdata.2017.122.

Karger, Dirk Nikolaus, Adam M. Wilson, Colin Mahony, Niklaus E. Zimmermann, and Walter Jetz. “Global Daily 1 Km Land Surface Precipitation Based on Cloud Cover-Informed Downscaling.” Scientific Data 8, no. 1 (2021): 307. https://doi.org/10.1038/s41597-021-01084-6.

Karger, Dirk Nikolaus, Stefan Lange, Chantal Hari, et al. “CHELSA-W5E5: Daily 1 Km Meteorological Forcing Data for Climate Impact Studies.” Earth System Science Data 15, no. 6 (2023): 2445–64. https://doi.org/10.5194/essd-15-2445-2023.

Knorr, W., L. Jiang, and A. Arneth. “Climate, CO2 and Human Population Impacts on Global Wildfire Emissions.” Biogeosciences 13, no. 1 (2016): 267–82. https://doi.org/10.5194/bg-13-267-2016.

Knorr, W., T. Kaminski, A. Arneth, and U. Weber. “Impact of Human Population Density on Fire Frequency at the Global Scale.” Biogeosciences 11, no. 4 (2014): 1085–102. https://doi.org/10.5194/bg-11-1085-2014.

Molinari, Chiara, Stijn Hantson, and Lars Peter Nieradzik. “Fire Dynamics in Boreal Forests Over the 20th Century: A Data-Model Comparison.” Frontiers in Ecology and Evolution 9 (September 2021). https://doi.org/10.3389/fevo.2021.728958.

Monsi, Masami, and Toshiro Saeki. “On the Factor Light in Plant Communities and Its Importance for Matter Production.” Japanese Journal of Botany 14, no. 1 (1953): 22--52.

Monsi, Masami, and Toshiro Saeki. “On the Factor Light in Plant Communities and Its Importance for Matter Production.” Annals of Botany 95, no. 3 (2005): 549–67. https://doi.org/10.1093/aob/mci052.

Morales, Pablo, Thomas Hickler, David P. Rowell, Benjamin Smith, and Martin T. Sykes. “Changes in European Ecosystem Productivity and Carbon Balance Driven by Regional Climate Model Output.” Global Change Biology 13, no. 1 (2007): 108–22. https://doi.org/10.1111/j.1365-2486.2006.01289.x.

Müller, Christoph, and Wolfgang Lucht. “Robustness of Terrestrial Carbon and Water Cycle Simulations against Variations in Spatial Resolution.” Journal of Geophysical Research: Atmospheres 112, no. D6 (2007). https://doi.org/10.1029/2006JD007875.

Rabin, Sam S., Joe R. Melton, Gitta Lasslop, et al. “The Fire Modeling Intercomparison Project (FireMIP), Phase 1: Experimental and Analytical Protocols with Detailed Model Descriptions.” Geoscientific Model Development 10, no. 3 (2017): 1175–97. https://doi.org/10.5194/gmd-10-1175-2017.

Schaphoff, Sibyll, Wolfgang Lucht, Dieter Gerten, Stephen Sitch, Wolfgang Cramer, and I. Colin Prentice. “Terrestrial Biosphere Carbon Storage under Alternative Climate Projections.” Climatic Change 74, no. 1 (2006): 97–122. https://doi.org/10.1007/s10584-005-9002-5.

Schurgers, Guy, Anders Ahlström, Almut Arneth, Thomas A. M. Pugh, and Benjamin Smith. “Climate Sensitivity Controls Uncertainty in Future Terrestrial Carbon Sink.” Geophysical Research Letters 45, no. 9 (2018): 4329–36. https://doi.org/10.1029/2018GL077528.

Smith, B., D. Wårlind, A. Arneth, et al. “Implications of Incorporating N Cycling and N Limitations on Primary Production in an Individual-Based Dynamic Vegetation Model.” Biogeosciences 11, no. 7 (2014): 2027–54. https://doi.org/10.5194/bg-11-2027-2014.

Zscheischler, Jakob, Erich M. Fischer, and Stefan Lange. “The Effect of Univariate Bias Adjustment on Multivariate Hazard Estimates.” Earth System Dynamics 10, no. 1 (2019): 31–43. https://doi.org/10.5194/esd-10-31-2019.

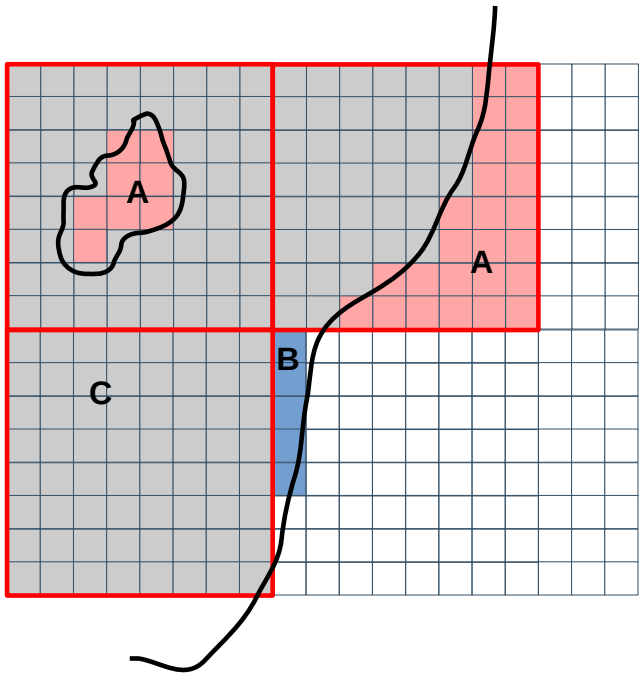

Figure 1: High- and low- resolution gridcells overlayed on the high resolution grid. The low-resolution gridcells are outlined in red. The thick black line represents the shoreline. *Gray:* areas present in the low-res simulation but not in the hi-res simulation. *Blue:* Areas present only in the hi-res simulation. *Red:* Areas present only in the low-res simulation.

1 Bias decomposition

Let X be a modeled variable, S_X the aggregated value of X over the simulated domain, and μ_X the domain-average. In order to calculate the climate-response and shoreline-representation components of the bias, we consider the following quantities, defined in the high resolution grid:

1. X_{ij}^{HR}: Value of the high-resolution output at grid point (i, j).
2. X_{ij}^{LR}: Value of the low-resolution output at grid point (i, j). We note that this value will be the same for all (i, j) within the same low-resolution gridcell (see Fig. 1.
3. A_{ij}: Surface area of the gridcell at gridpoint (i, j)
4. $M_{ij}^{\mathrm{LR,HR}}$: *Overlap mask*. It takes the value 1 at land points where low-resolution values and high-resolution values overlap (gray cells in Fig. 1), and 0 everywhere else.
5. $M_{ij}^{\overline{\mathrm{LR}},\mathrm{HR}}$: Only *high-resolution mask*. It takes the value 1 at land points present in the high-resolution simulation, but not present in the low resolution one (blue cells in Fig. 1) and 0 everywhere else.

6. $M_{ij}^{\mathrm{LR},\overline{\mathrm{HR}}}$: Only *high-resolution mask*. It takes the value 1 at land points present in the low-resolution simulation, but not present in the high resolution one (red cells in Fig. 1) and 0 everywhere else.

1.1 Regionally aggregated quantities

For regionally aggregated variables, such as the carbon fluxes and pools, the bias between high- and low- resolution outputs is:

$$\begin{aligned}\delta &= S_X^{\mathrm{LR}} - S_X^{\mathrm{HR}} \\ &= \sum_{i,j} X_{ij}^{\mathrm{LR}} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\mathrm{LR},\overline{\mathrm{HR}}}) \\ &\quad - \sum_{i,j} X_{ij}^{\mathrm{HR}} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\overline{\mathrm{LR}},\mathrm{HR}}),\end{aligned} \tag{1}$$

where the indices (i, j) cover the whole domain. In this equation, the first sum represents the regional sum of the low resolution values, and the second term is the regional sum of the high-resolution values. Rearranging terms yields:

$$\begin{aligned}\delta = &\underbrace{\sum_{i,j} (X_{ij}^{\mathrm{LR}} - X_{ij}^{\mathrm{HR}}) A_{ij} M_{ij}^{\mathrm{LR,HR}}}_{\delta_{\mathrm{cli}}} \\ &+ \underbrace{\sum_{i,j} A_{ij} (X_{ij}^{\mathrm{LR}} M_{ij}^{\mathrm{LR},\overline{\mathrm{HR}}} - X_{ij}^{\mathrm{HR}} M_{ij}^{\overline{\mathrm{LR}},\mathrm{HR}})}_{\delta_{\mathrm{sho}}} .\end{aligned} \tag{2}$$

The first term of the above equation, labeled as δ_{cli}, involves values of X at overlapping gridcells exclusively (shown as gray cells in Fig. 1). Hence this term can be attributed to the difference in climate forcings between the two simulations. The second term, labeled δ_{sho} involves values of X at non-overlapping gridcells between the high- and low- resolution simulations. These gridcells are the red and blue gridcells from Fig. 1, and are associated with poor shoreline representation at low resolution.

1.2 Regionally averaged quantities

The variables FPC and LAI are averaged across the domain, rather than aggregated. The bias in this case is calculated as:

$$\begin{aligned}\delta &= \mu_X^{\mathrm{LR}} - \mu_X^{\mathrm{HR}} \\ &= \frac{\sum_{i,j} X_{ij}^{\mathrm{LR}} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\mathrm{LR},\overline{\mathrm{HR}}})}{\sum_{i,j} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\mathrm{LR},\overline{\mathrm{HR}}})} \\ &\quad - \frac{\sum_{i,j} X_{ij}^{\mathrm{HR}} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\overline{\mathrm{LR}},\mathrm{HR}})}{\sum_{i,j} A_{ij} (M_{ij}^{\mathrm{LR,HR}} + M_{ij}^{\overline{\mathrm{LR}},\mathrm{HR}})},\end{aligned} \tag{3}$$

where the first term is the low-resolution regional average, and the second term is the high-resolution regional average. Rearranging terms yields

$$\delta = \delta_{\text{cli}} + \delta_{\text{sho}}, \tag{4}$$

where

$$\begin{aligned}\delta_{\text{cli}} = &\frac{\sum_{i,j} X_{ij}^{\text{LR}} A_{ij} M_{ij}^{\text{LR,HR}}}{\sum_{i,j} A_{ij} (M_{ij}^{\text{LR,HR}} + M_{ij}^{\text{LR,}\overline{\text{HR}}})} \\ &- \frac{\sum_{i,j} X_{ij}^{\text{HR}} A_{ij} M_{ij}^{\text{LR,HR}}}{\sum_{i,j} A_{ij} (M_{ij}^{\text{LR,HR}} + M_{ij}^{\overline{\text{LR}},\text{HR}})},\end{aligned} \tag{5}$$

and

$$\begin{aligned}\delta_{\text{sho}} = &\frac{\sum_{i,j} X_{ij}^{\text{LR}} A_{ij} M_{ij}^{\text{LR,}\overline{\text{HR}}}}{\sum_{i,j} A_{ij} (M_{ij}^{\text{LR,HR}} + M_{ij}^{\text{LR,}\overline{\text{HR}}})} \\ &- \frac{\sum_{i,j} X_{ij}^{\text{HR}} A_{ij} M_{ij}^{\overline{\text{LR}},\text{HR}}}{\sum_{i,j} A_{ij} (M_{ij}^{\text{LR,HR}} + M_{ij}^{\overline{\text{LR}},\text{HR}})}.\end{aligned} \tag{6}$$

1 2.1.2 Precipitation

CHELSA considers only orographic precipitation (Karger et al., 2023), which is done by computing the wind effect index H for each high-resolution cell. This index reflects how much moisture gets pushed up towards the top of a mountain as well as rain shadow in its leeward direction, and it is computed using u-wind and v-wind components from CMIP6 data. Those components were interpolated to the high-resolution grid with a B-spline, and then were projected onto a world Mercator projection.

$$H = H_{W,L} \rightarrow d_{LH_i} < 0 \times H_{W,L} \rightarrow d_{LH_i} \geq 0, \tag{1}$$

$$H_W = \frac{\sum_{i=1}^{n} \frac{1}{d_{WHi}} \tan^{-1}\left(\frac{d_{WZi}}{d_{WHi}^{0.5}}\right)}{\sum_{i=1}^{n} \frac{1}{d_{LHi}}} + \frac{\sum_{i=1}^{n} \frac{1}{d_{LHi}} \tan^{-1}\left(\frac{d_{LZi}}{d_{LHi}^{0.5}}\right)}{\sum_{i=1}^{n} \frac{1}{d_{LHi}}} \tag{2}$$

$$H_L = \frac{\sum_{i=1}^{n} \frac{1}{\ln(d_{WHi})} \tan^{-1}\left(\frac{d_{LZi}}{d_{WHi}^{0.5}}\right)}{\sum_{i=1}^{n} \frac{1}{\ln(d_{LHi})}} \tag{3}$$

, where d_{WHi} and d_{LHi} denote the horizontal distances in windward and leeward direction, while d_{WZi} and d_{LZi} are the corresponding vertical distances. The summations in (2) and (3) are performed within a circle with the radius of 75 kilometers.

The H index is then corrected according to the atmospheric boundary layer height to account for the contribution of the surface pressure level to the wind effect. Lastly, the low-resolution precipitation p_{lr} is multiplied by the corresponding H indices and normalized to obtain high-resolution precipitations p_{hr}, so that within each low-resolution grid cell the sum of the values p_{hr} remains equal to p_{lr} (see section Methods in Karger et al. (2021)).

2 2.1.3 Surface downwelling shortwave radiation (RSDS)

The total shortwave radiation, measured in W/m^2 is represented as in (Karger et al., 2023), Sect. 2.2.2:

$$S_{\mathrm{n}} = S_{\mathrm{s}} + S_{\mathrm{h}}. \tag{4}$$

Here, S_{s} is direct solar radiation reaching the surface, computed according to the position of the Sun with respect to the high-resolution grid cell. Diffuse solar radiation S_{h}, which is the energy re-emitted by the atmosphere, takes into account the percentage of the sky observable from a grid cell.

Computation of S_s component starts with astronomical equations. For the sun elevation angle θ, sun azimuth φ, latitude λ, the solar declination angle δ, the Julian day number J, hour h, and the hour angle in degrees $\bar{\omega}$, we have the following:

$$\begin{aligned}
\sin\theta &= \cos\lambda\cos\delta\cos\bar{\omega} + \sin\lambda\sin\delta \\
\cos\varphi &= \frac{\cos\delta\cos\bar{\omega} - \sin\theta\cos\lambda}{\sin\lambda\cos\theta} \\
\delta &= 23.45 \cdot \sin\left(\frac{360^\circ[284+J]}{365}\right) \\
\bar{\omega} &= 15^\circ(12-h).
\end{aligned} \tag{5}$$

Using these identities, $\cos\gamma$ is computed as

$$\cos\gamma = \cos\beta \cdot \sin\theta + \sin\beta \cdot \cos\theta \cdot \cos(\varphi - \alpha), \tag{6}$$

where γ is the angle between the Sun beam and the normal to the terrain, while α and β are surface slope and aspect. Then, S_s is computed using constants $G_{sc} = 1367\ kW \cdot m^2$, $\tau = 0.8$, and air optical thickness m defined in formula (13) of Karger et al. (2023):

$$S_s(h) = \varsigma(h) \cdot G_{sc} \cdot \tau^m \cdot \cos\gamma. \tag{7}$$

Diffuse solar radiation is calculated as

$$S_h = (0.271 - 0.294\tau^m) G_{sc} \Psi_s, \tag{8}$$

where Ψ_s is the sky view factor computed as

$$\Psi_s = \frac{1}{N} \sum_{i=1}^{N} [\cos\beta\cos\varphi_i + \sin\beta\cos(\Phi_i - \alpha) \cdot (90 - \varphi_i - \sin\varphi_i\cos\varphi_i)] \tag{9}$$

for $N = 8$ azimuth directions Φ_i and the corresponding horizon angles φ_i.

$$rsds = \bar{S}_n(1 - 0.75 \cdot clt^{3.4}), \tag{10}$$

where $\bar{S}_n$ is an average of S_n over 24 hours, and clt is the cloud cover computed according to formulas (19)–(22) of Karger et al. (2023).

To summarize this procedure, we note that the S_{s} and S_{h} components are obtained by taking into account shadowing and obstruction of light, the position of the Sun, the slope and the aspect of the terrain, and cloud cover resulting from orographic precipitation formation.