Review of the paper „Ent Köppen-Geiger v1.0, a software tool for generating Earth System Model global vegetation boundary conditions for alternative climates“, authored by Nancy Kiang et al.

Kiang et al. presents „EntKGLUT“, a lookup-table tool for generating vegetation boundary conditions for climate models from Köppen–Geiger climate classes. It translates these classes into plant functional type (PFT) cover fractions, bare-soil fractions, canopy height, and monthly leaf area index (LAI), using statistics derived from modern satellite-based vegetation data. Weighted kernel density estimation identifies representative canopy properties, while optional post-processing rules adapt the vegetation composition for past geological periods by replacing PFTs considered inappropriate for the selected age.

The authors evaluate the tool in two applications. For modern conditions, they compare simulations performed in the ESM ModelE-2.1, using the original satellite-derived vegetation boundary conditions with simulations using the simplified LUT-derived conditions. The simulations produce broadly similar climates and relatively small differences in global carbon fluxes, although biomass estimates are affected by acknowledged canopy-height biases and a vegetation-model bug. For the Late Cretaceous at 85 Ma, they apply the tool to a ROCKE-3D simulation with 1500 ppm CO_2, initially run with bare-soil continents. Introducing the inferred vegetation warms global mean surface air temperature by 1.1 °C, with the strongest warming at high latitudes, and slightly increases global precipitation. A subsequent vegetation classification shows further changes, but a fully converged climate–vegetation state remains for future work.

The tool offers a practical way to construct vegetation boundary conditions where detailed reconstructions or dynamic vegetation simulations are unavailable. While in principle, I appreciate the development of EntKGLUT as a practical and promising approach to estimating vegetation boundary conditions, particularly where detailed vegetation reconstructions are unavailable. However, the tool relies on several assumptions and simplifications whose implications are not sufficiently evaluated or discussed. Its main limitations concern transferring modern climate–vegetation relationships to past environments, the simplified PFT replacement rules, and sensitivity to climate-model biases. The modern experiments demonstrate the climatic consequences of simplifying vegetation inputs; they do not independently establish the accuracy of the resulting paleovegetation. A clearer assessment of these limitations and the associated uncertainties is needed to establish the tool’s applicability, especially to paleoclimates. Therefore, I suggest major revision before publishing the paper in GMD:

1.) A Climate class is not necessarily the same as a vegetation class. Different vegetation types may occur within the same Köppen–Geiger class because of differences in soils, drainage, disturbance regimes, and land use. Please discuss how this within-class variability affects the representativeness of the assigned vegetation properties.

2.) Satellite-derived vegetation data may reflect modern soils, land use, disturbance regimes, and historical legacies, including substantial human impacts. Please clarify how these influences are treated when constructing the LUT, as they may otherwise be carried over into vegetation boundary conditions for palaeoclimates.

3.) The modern climate–vegetation relationships underlying EntKGLUT implicitly reflect vegetation under modern atmospheric CO_2 concentrations. Applying these relationships to the 85 Ma experiment at 1500 ppm accounts for CO_2-induced climatic changes through the climate

simulation, but does not necessarily capture direct physiological effects on vegetation cover, PFT competition, and LAI. Please clarify which vegetation properties can respond to CO_2 in this configuration and which remain prescribed from modern analogues, and discuss the implications for applications under different CO_2 concentrations.

4.) The tool does not explicitly represent dispersal limitations or migration lags following climate change. Although modern PFT mixtures reflect competition under contemporary conditions, the tool does not simulate how competitive outcomes change in other environments. These limitations are particularly relevant to transient applications and past environments without close modern analogues and should be acknowledged when defining the tool's intended scope.

5.) The simplified palaeovegetation replacement rules only partly account for regional evolutionary and dispersal histories. Replacing grasses with bare soil changes total vegetation cover, while transferring LAI and height from removed to replacement PFTs may produce functionally inconsistent combinations, particularly in seasonal LAI patterns. In addition, discrete climate-class boundaries may cause abrupt vegetation changes in response to small climatic shifts, potentially complicating convergence during asynchronous coupling. Please justify these assumptions and assess their implications.

6.) Validation and uncertainty assessment remain limited, and the conclusions should be qualified accordingly. Similar modern climates under the original and LUT-derived VBCs demonstrate limited climatic differences in this model configuration, but do not establish the accuracy of individual vegetation properties or their transferability to other climates and geological periods. Compensating errors or limited model sensitivity may contribute to this agreement. The Cretaceous experiment does not yet demonstrate full climate–vegetation convergence, and the comparison with fossil reconstructions requires a clearer assessment of agreement and uncertainty. Uncertainties associated with satellite data, KDE bandwidth, sparsely represented climate classes, climate inputs, and PFT replacement rules should be assessed and, where feasible, propagated into the resulting VBCs and simulated climate effects. The discussion should also distinguish asynchronous LUT coupling from process-based dynamic vegetation modelling.

Major and minor comments in detail:

- Köppen-Geiger defines climate classifications. This is not the same as biomes. Please check the document again and do not mix these definitions.

L81: The comparison between 'biome-based BIOME' and 'PFT-based BIOME3' is misleading. Biome 1 was also based on PFTs. The main advance of BIOME3 and BIOME4 lay in linking vegetation distribution and biogeochemical processes, and in modelling competition on the basis of productivity.

L83: "surface fluxes of a daily soil water balance" the water balance is no surface flux.

L84: BIOME4 evaluates PFTs on the basis of their net primary production and sustainable leaf area. Empirical rules then assign a biome based on the dominant PFTs and their LAI. Thus, combinations are not simply ranked on the basis of monthly values.

L85: remove "latest"

L89: „half-hourly physical time-step“. This is the case in ORCHIDEE; but probably not a general number.

L89: „drive“ is too strong, better: .... to exchange water and energy with the atmosphere

L93: What do you mean by „high“ ?

L97: „Weaknesses in the understanding and data“; better: "remaining uncertainties in ecological processes and limitations in available data"

L99: It does not automatically follow from the uncertainties inherent in process-based models that vegetation determined on the basis of climate data is more reliable. This approach, in turn, presupposes that the climate–vegetation relationships used are applicable to the period under investigation. Satellite data and attribute databases can improve the parameterisation, but they do not eliminate this assumption.

L103: MODIS is an instrument, not a satellite.

L113: canopy height and vertical structure help to characterise vegetation; however, the PFT composition cannot be unambiguously determined from these factors alone.

L122: The Köppen–Geiger classification is suitable for depicting observed and simulated climate zones. Shifts in these zones may indicate changes in climatic suitability for vegetation. However, it does not, on its own, predict actual shifts in vegetation.

LL135: Identical temperature and precipitation classes do not guarantee identical canopy heights or LAI. CO_2, soils, disturbances, land use and vegetation composition can alter these properties. For the future and deep geological history, the issue of transfer over time must also be considered; ‘conserved in space’ does not adequately describe this stronger assumption.

LL137: The evolutionary reconstructions are simplifications, not complete reconstructions. It is plausible to exclude C_4 grasses for the Cretaceous period. However, replacing them with C_3 grasses or bare ground requires further justification: the absence of a modern group of plants does not mean that the corresponding area was devoid of vegetation. Furthermore, evolutionary origin, first fossil occurrence and ecological spread all took place at different times.

LL142: A modern comparison does not automatically validate palaeoclimatic applicability. Similar climate simulations using original and LUT-based boundary conditions initially show that the simplified boundary conditions yield similar results in the modern test. They do not prove either the correct vegetation under different climates or the correct Cretaceous LAI and crown heights. A fossil comparison can only verify the characteristics that have actually been reconstructed.

L149: maximum annual LAI or annual maximum LAI?

L151: as a function of geological age ? Instead of „years before present“

L152: Please clarify whether the modern evaluation uses the same satellite data and period as the LUT construction. If so, this comparison assesses how well the LUT reproduces its source data and their simulated climatic effects, rather than providing independent validation. This is a useful consistency test, but it does not establish transferability to past or future climates. Please distinguish

these purposes and discuss the limitations of applying modern climate–vegetation relationships under different CO_2 concentrations and vegetation compositions.

L155: The fossil comparison only examines comparable characteristics. A match in vegetation types does not automatically confirm the assigned crown heights, LAI or PFT area proportions.

L 194: Check this sentence

L195: “extended version” is not explained here. Please refer at least to the Appendix.

LL203: Please check and explain it in more detail. The standard deviation is not identical to the width at half maximum of the peak.

L213: Please check this sentence.

L214: Clarify which grid cell is chosen and why. Do you mean KDE instead of regression?

L228: Maybe „vegetation albedo“, as the soil albedo is part of the tool.

LL274: Is bare soil realistic? The fact that grasses did not yet exist does not mean that these areas were devoid of vegetation. Other plants may have grown there that the model does not adequately represent.

L282: „are“ is missing: grasses are present.

L292: I understand that this is something you have to deal with, but it would be good to provide a scientific basis for this decision as well. The idea that these two functional types of deciduous trees were entirely replaced by evergreen conifers 66 Ma ago is a strong assumption that would need to be supported by paleobotanical evidence.

L294: linear ramp up?

LL308: Please clarify the functional consistency of the PFT replacement procedure. When the replacement PFT is absent from a grid cell, it inherits the LAI and height of the removed PFT. This may combine the replacement PFT’s prescribed parameters with structural and seasonal characteristics of the removed vegetation type. For example, replacing deciduous needleleaf trees with evergreen needleleaf trees while retaining the original monthly LAI could preserve a deciduous seasonal cycle. Sharing the same growth form does not necessarily justify this transfer, particularly for monthly LAI. Please explain how these inherited values interact with PFT-specific parameterizations in the receiving model and assess the sensitivity of the resulting boundary conditions to this assumption.

L337: „spinups were conducted….“ what does this exactly mean, 3 enemsble members, or 3 spin-up cycles?

L338: „transient conditions from 1850 to 2020“. during this transient run, the VBCs are constant. right? On modern level, right?

L338: How do you distinguish the effect of the vegetation from the effect of internal climate variability?

LL356: The two asynchronous updates provide a preliminary test of climate–vegetation feedbacks, but do not establish convergence towards a mutually consistent climate–vegetation equilibrium. Equilibrating climate under prescribed vegetation does not ensure that the vegetation subsequently inferred from that climate remains unchanged. Please report the duration and equilibrium criteria of the second simulation, quantify the changes between successive vegetation classifications, and justify whether further iterations are necessary. The sensitivity to the initial uniformly sandy, unvegetated surface and the robustness of the classification based on ten-year climatologies should also be assessed. Finally, comparison between iterations tests internal consistency; evaluating the accuracy of the predicted Cretaceous vegetation requires comparison with independent paleobotanical evidence.

L362: Please explain in more detail how you removed the land-use types

L368; „from with croplands“ → correct

L377: somewhere in the main manuscript should occur the abbreviations for the cover types used here

L440: the units are not clear. Is it a flux? Is it per grid-cell?

L441: „within the range…“ How large are these ranges of uncertainties. Please provide numbers and at least references

LL448: This bug limits the significance, as it can also influence the difference between E1 and E0 due to different PFT distributions

L469: Replacing all grass cover with bare soil introduces a strong assumption about total vegetation cover. Excluding a grass PFT does not establish that the corresponding area was unvegetated rather than occupied by other plants. Please justify this replacement using paleobotanical evidence and assess sensitivity to alternative vegetation assignments. The prescribed bright/dark soil mixture also changes albedo, so the resulting climate response combines changes in vegetation cover, PFT composition, and soil reflectance. Please distinguish these effects and clarify that changes in cover-weighted canopy height and LAI may reflect either increased bare-ground fractions or changes in the properties of the remaining vegetation.

L497: Please qualify the attribution of the 1.1 °C warming as “directly due” to reduced surface albedo. The experiment simultaneously changes vegetation cover, canopy height, and monthly LAI, and therefore does not isolate the albedo contribution. Changes in snow and ice may also represent feedbacks rather than solely the initial cause of warming.

L499: Please emphasize that the warming is relative to the prescribed bare-soil reference state and quantify the subsequent vegetation changes to assess whether further asynchronous iterations are needed.

L500: Discussion is probably the heading of the next section

L535: “an for research” → please correct.

L536: I guess, KDE cannot distinguish between which peak is environmentally valid and which is caused by a data error. It merely describes the available data. Particular caution is required regarding the aforementioned contamination of the MODIS-LAI data: if an LAI value for shrub or

grass areas partially includes the signal from neighbouring trees, this can result in an additional peak, I guess. Please clarify.

LL544: the sentence is underlined.

LL567: Please distinguish the effects of biases in the simulated climate from those introduced by the LUT and paleo-vegetation replacement rules. At 85 Ma, both deciduous broadleaf PFTs are explicitly replaced by evergreen needleleaf trees, whereas the cited fossil reconstruction includes deciduous broadleaf trees in high-latitude mixed forests. This component therefore cannot be reproduced after post-processing, even if the input climate becomes warmer. Please discuss this structural limitation alongside the high-latitude temperature bias and clarify whether the mixed forests mentioned in lines 575–576 refer to vegetation before or after PFT replacement.

L582: The mismatch with the fossil vegetation reconstruction may reflect not only biases in the simulated climate, but also the transfer of modern climate–vegetation relationships to elevated CO_2 conditions. The reconstructed vegetation developed under contemporaneous CO_2 levels, whereas the LUT-derived cover fractions and structural properties implicitly reflect modern conditions. Missing direct CO_2 effects could therefore contribute to discrepancies even under an accurately simulated climate. Please discuss this possibility separately from climate biases and the imposed PFT replacements, without attributing the mismatch to CO_2 unless supported by sensitivity tests.

LL610: A similar climate does not automatically confirm accurate vegetation boundary conditions. E0 and E1 produce a similar modern climate. This supports the claim that the simplification in this model and experiment has relatively little effect on the climate simulation. However, it does not prove that PFT fractions, LAI and height are correct: different vegetation errors could cancel each other out, or the model might be insensitive to certain differences. The applicability to any climatological data — particularly palaeoclimates — is therefore not sufficiently substantiated.

LL615: Asynchronous updating links climate and vegetation classification. However, it does not automatically incorporate the processes of a dynamic vegetation model, such as CO_2-dependent competition, establishment and vegetation dynamics. In addition to climate model errors, the limitations of modern LUTs and the substitution rules must therefore also be discussed.