

## Review of manuscript egusphere-2024-2170

***‘Uncertainty quantification for overshoots of tipping thresholds’*** by Lux-Gottschalk & Ritchie

The authors have addressed the comments, questions and suggestions included in my first review in their responses and revisions. In particular, restructuring the manuscript by describing the models and the deviation of the tipping probability before presenting the impact of uncertainty in the location of the critical threshold and the linear restoring force on the tipping response helps to follow and understand the key results of this study.

I have a few more specific comments, questions and suggestions which, in large parts, are related to the clarity of the language in the revised manuscript. These are listed below.

### Specific comments

L01 (and L22): I am not sure whether parts of the Earth system (including those potentially displaying tipping dynamics) can be considered to be currently in a stable state. The question that I posed in my first review may have been misleading. Maybe formulate more carefully. See also L432-L444, where this assumption is, in fact, discussed for the AMOC in this manuscript.

L04: Maybe replace “provides one possible mechanism for tipping” by “constitutes a possible mechanism for tipping”, as no other mechanism of tipping is described in the abstract.

L09: Maybe replace “look at” by “assess” or “explore”.

L09: Maybe replace “affect the probability of tipping” by “propagate to uncertainties in the probability of tipping” (or similar).

L15: I am not sure if “pathway believed to offer little danger of tipping” fits well here. Please consider to reformulate this phrase.

L20: Please consider to also add older references. Tipping dynamics in the Earth system have not just gained attention recently, see, for example, van Nes et al. (2016) and Lenton et al. (2008) (which is also cited at a later point in the manuscript).

L26-L28: I am not sure if this sentence adds any new information compared to the previous sentences, except introducing the AMOC as an example for a tipping system. I would suggest to reformulate along the lines of, for example, “One example for a part of the Earth system suggested to exhibit tipping behaviour is the Atlantic Meridional Overturning Circulation (AMOC)”. Please also consider to add a few more specific references on the possible tipping dynamics of the AMOC, e.g. based on Weijer et al. (2019) and references therein.

L29: Maybe replace “trigger” by “lead to”.

L32: Maybe replace “impact is” by “impacts are”.

L33: What is meant by “measures”? Please consider to be more precise in the wording here.

L34: Overshoots of tipping thresholds have not been introduced so far, but have only been described in the abstract (L05-L07). Please add a short explanation, and a motivation why overshoots should be studied in the introduction.

L36-37: Please consider adding “without a change in the external forcing”.

L39: Please consider adding “by (slowly) changing the external forcing”.

L43: It could be helpful to explain what you mean by “mitigation window” in the introduction. A possible explanation is given in L86-88 with respect to overshoots, but it is unclear whether / how this applies here.

L43-L45: How does this justification in terms of a “natural restriction” related to L34-35, describing that “it is important to understand which mechanisms can cause a system to tip” to “understand overshoots of tipping thresholds”. I think I understand what you mean, but wonder if clarity could be improved here.

L47: Maybe it is possible to clarify this sentence. Do you mean something along the lines of “climate model simulations have suggested that tipping might not occur despite an overshoot”?

L54-55: Maybe consider to rephrase along the lines of “emphasises that high-impact, low-likelihood climate outcomes, to which some tipping events belong, should be part of climate risk assessments”, if appropriate.

L57-58: Please consider to briefly describe what “indicators of critical slowing down” refers to for readers that are not familiar with early warning signals.

L60: Maybe replace “retain” by “return to”.

L79: Maybe replace “environmental parameter” by “external forcing” for clarity, as used previously in the manuscript and, for example, in the axis labelling in Figure 1.

L79: Maybe replace “have different impacts on the overshoot of the critical bifurcation (threshold) value” by, for example, “have different characteristics with respect to the overshoot of the critical threshold / bifurcation value”.

L81-82: I would like to suggest “These uncertainties introduce uncertainties in the time and peak overshoot distance...”.

L84: Maybe replace “thus includes a range of forcing parameters where the AMOC exhibits multistability” by “thus there exists a range of external forcings where the AMOC is characterized by multistability”.

L86-87: I would suggest to reformulate along the lines of “inverse-square law between the peak overshoots distance, exceedance time and system characteristics”.

L98: Previously, “peak external forcing” was described as “peak overshoot distance”. I would suggest to stick to one of these terminologies throughout the manuscript for clarity.

L105: “affects the tipping behaviour” or “affects the uncertainty in the tipping behaviour”?

L111-112: Since a general reference to bifurcation theory has already been given in the introduction, this may not be needed here.

L115: To me, this reads like a repetition of the first sentence of this paragraph. Please consider removing or bringing both sentences together.

L116: I am not sure whether the reference point for the distance to the basin boundary is clear. Alternatively, this could maybe be replaced by “the size of the basin of attraction”.

L116: Comma may not be needed.

L117: Does “all system characteristics” refer to the linear restoring force and the distance to the basin boundary? Please consider to clarify here.

L117: To me, this formulation is confusing after reading the preceding sentence. Maybe replace “remain unchanged” by “are the same”.

L117: What is meant by “system’s”? Does this refer to (two) systems that differ in terms of the location of their tipping threshold (that is “systems”)? Or the “systems’ states”?

L117-118: If I understand correctly, the basins of attractions, however, have different sizes for the two example systems considered here, when the external forcing approaches zero again after the overshoot (and the systems are not at the same distance to their respective thresholds). Does the argument described in L193-L196 for the case of the uncertain linear restoring force also play out for uncertainties in the location of the tipping threshold?

L119: Maybe consider to remove “develop this idea further”.

L124: Maybe replace “for less time” by “of shorter duration” (or similar).

L125: If possible, please formulate “contrasting consequences” more specific.

L127: “curves show” instead of “curvesshow”

L135: Are “the probability of tipping” and the “critical tipping threshold location” equivalent? Maybe also consider to find another phrase for “critical tipping threshold location”. I am not sure whether this phrase is intuitive, given that it could be understood as the critical threshold of an example system, while (if I understood correctly) it here, for a given overshoot profile, refers to *the* critical threshold of *the* system that separates tipping from not tipping (L133).

L158: Please consider to reformulate along the lines of “For lower thresholds, however, an overshoot of the threshold...”

L159: Maybe replace “larger peaks” by “larger peaks in the external forcing” or similar.

L164: If I understood correctly, in L126-L127 the same trajectory is described to “tip due to the large and long overshoot”. This formulation may seem a bit contradictory to “tipping is nearly avoided”. Please check the formulation.

L172: Please consider to expand this sentence to improve clarity, for example, as “For example, forcing profiles that were very unlikely to result in tipping (10% probability of tipping) for

the initial distribution are now exceptionally unlikely ( $< 1\%$ ) to give rise to a critical transition given the knowledge-based distribution.”

L180: After introducing Section 2, this explanation of the strength of the linear restoring force may not be needed here. I would suggest to move it to Section 2, where this quantity is formally introduced, or to integrate into the sentence for better readability.

L189: Please consider to expand this sentence for clarity, for example, as “... does not cross the unstable branch (representing the boundary of the basin of attraction) when reducing the external forcing...”, if appropriate.

L193: I think I understand what is meant here, but would suggest to reformulate. I am not sure whether a system can “take longer to realise it is over the edge”?

L193-194: I would suggest to reformulate along the lines of “In addition, for a weaker restoring force proportionality factor, the boundary of the basin of attraction is further away from the initial stable state (blue dot in Fig. 3b)”.

L202: Maybe replace “then plots” by “shows”.

L203: Maybe replace “not unlike before this is now fixed” by, for example, “which is now fixed, in contrast to Sect. 3.1”.

L203-205: I think this could be removed here, and should rather be part of the figure caption.

L212: Maybe replace “whereas” by “in contrast”.

Figure 4: I was wondering whether Figure 4b might be easier to understand if it would be split into two panels, showing the tipping probability for (1) the initial and knowledge-based distribution of the restoring force probability factor, and (2) the initial and alternative distribution of the restoring force probability factor?

L244: Please integrate “a proxy for the strength of the AMOC” into the sentence to improve readability.

L264-267: If I understand correctly, these sentences are also relevant for the description of Figure 5c (L281-LL283). Would it be possible to bring these parts together?

L288: How do “arbitrarily” and “the advective timescale is relatively well constrained” / “within the reasonable physical range” fit together? Please clarify or reformulate.

L290: If I understand correctly, the red dashed line indicates the value chosen for the advective timescale (see caption of Figure 5). Please move the reference to the red dashed lines accordingly (e.g. to L288), or clarify.

L316: To me, “We now again follow a similar approach as for Figures 1 and 3” reads like a repetition of the first sentence of this paragraph. Please consider removing or bringing both sentences together.

L345-346: I would suggest to phrase this sentence the other way around for clarity, i.e. “The large uncertainty in the system parameter, here in terms of the diffusive timescale, again causes large uncertainties in the tipping behaviour”.

L359: Please consider specifying as “posterior distribution of the diffusive timescale”.

L359: Please define “mitigation window”, see previous related comment.

L365-366: Please consider simplifying this sentence.

L384-388: If I understand correctly, this section describes the effects of an increasing standard deviation of the normally distributed diffusive timescale on the tipping uncertainty (resulting in a larger tipping uncertainty), while L 393-401 describes the effects of a decreasing standard deviation (potentially leading to a reduction in the uncertainty of tipping behaviour). Please consider to bring both sections together.

L398-L399: Please add a reference to Figure 8a for clarity.

L405: Maybe remove “individually” or reformulate. “isolated” and “individually” may not be needed together.

L458: Maybe reformulate along the lines of “the strong influence of the uncertainty in the tipping threshold location”.

L459: Maybe replace “seen” by “shown”.

L461: Maybe replace “if we want to avoid the tipping of elements of the climate system” by “to avoid critical transitions in parts of the climate system” or “to avoid tipping in parts of the climate system”.