Earth Syst. Dynam. Discuss.,
https://doi.org/10.5194/esd-2020-43-AC2, 2020

Earth System Dynamics Discussions

Interactive comment on "Collapse of the Atlantic Meridional Overturning described by Langevin dynamics" *by* Jelle van den Berk et al.

Jelle van den Berk et al.

jelle.van.den.berk@knmi.nl

Received and published: 8 September 2020

This is an interesting paper, which aims to describe the collapse and hysteresis of the AMOC observed in intermediate complexity climate models subject to freshwater forcing by low-dimensional Langevin dynamics as a stochastic bifurcation of a double well potential. Substantial revisions are necessary to improve the clarity of the manuscript and to support the conclusions.

R: We thank the reviewer for the detailed comments and valuable suggestions. Below are our follow up comments and actions.

General comments 1.) It is not clear what the purpose of the paper is. The authors do not state what their model is able to explain or predict. Is the purpose to predict the

Printer-friendly version

Discussion paper

exact parameter value of a collapse? Or at least to develop a method to do this?

R: Our aim is to investigate whether it is possible to model the outcome of complex models with a low dimensional model and thereby enhance understanding of the physics of an AMOC collapse and its hysteresis behaviour. The Langevin model defines a low-dimensional manifold that captures the essential collapse characteristics. To the extent that the low-dimensional model is successful in capturing the more complex model this investigation can indeed be seen as the development of a method to predict the parameter range where in a model a collapse would occur. The method is thus partly geared toward providing a means of prediction, but (at present) mainly to provide some characterisation of the collapse that will allow comparison between climate models. If a good fit can be found, then we can further explain the non-dynamical nature of the AMOC variations. As we show, this is partially the case.

Are there prospects to apply the method to observational data?

R: Since the forcing is a freshwater anomaly in the North Atlantic, we would need to estimate the counterpart in the real world. Moreover, the forcing values at the bifurcations have to be known. At present, this is not the case for the real world, and in models it is model-dependent. An attempt could be made to relate the forcing to an indicator that can be linked to the bifurcation points, for instance Mov. This should first be tested. If this can be done, then from the transient change in mu and with knowledge of sigma from observed AMOC variations, a predictive model for the likelihood of an AMOC collapse could be developed.

Or is an aim to understand dynamically what is happening in realistic climate models? This should be stated in the introduction

R: We would like to see this paper as a first step to say more about the behaviour of complex climate models. In particular, we propose a simplified low-dimensional model that is able to explain (and predict) bifurcation (tipping) points and abrupt change in the AMOC. Ultimately, this could be used to investigate abrupt behaviour in CMIP models, and how likely abrupt changes would be, even if not simulated by the model. Without a-priori knowledge of the freshwater-forcing values associated with the model's bifurcation points, first must be investigated whether mu can be linked to a general indicator like Mov (see the comment above). We will state the purpose of the paper more clearly in the introduction and add the outlook above to the discussion section.

(P2L29ff). It is also unclear whether they want to only/mostly model the AMOC collapse (as stated at some points in the paper) or also the resurgence.

R: We are mainly interested in modelling the collapse, not the resurgence. We will note this here as "Although the hysteresis loops of the AMOC include both a collapse and a resurgence at, we will only attempt to model the collapse from the the stable on branch to the stable off branch."

2.) Regarding the conclusions, how can the authors say that the model successfully captures the dynamics? They don't compare with other models of higher or lower complexity, nor do they have any metric that shows goodness of fit or anything similar. This would be necessary to make such a conclusion.

R: We thank the reviewer for this comment. We will add the posterior spread to show the goodness of fit. To explore and compare with other low-dimensional models than the Langevin model is beyond the scope of this paper (the six included already form a multi-model ensemble). A possible next step could be to apply the Langevin model to a transient run where mu depends on time. The AMOC should eventually collapse and we can have greater confidence in the applicability of our our approach.

3.) The manuscript is not very well written and hard to follow. The terminology is often unclear. (E.g. what is a "track", and how does the use of "stability landscape" apply here? See specific comments.) Some corrections are given under "technical corrections", but the language and terminology has to be generally improved throughout the manuscript. Furthermore, I believe the manuscript can be shortened severely. What the authors want to get across can be said more efficiently. Many things are mentioned twice or more (see specific and technical comments for suggestions). Finally, the labels in multiple figures are unreadable.

R: We critically reviewed the text. We comment on identified issues with the specific comments below. In short, the terminology will be clarified, and parts of the text will be removed. The figures will be improved with larger and consistent fonts for the labels.

4.) The data acquisition seems problematic. I am not sure whether it is viable for this journal to present a data analysis based on visually extracted data from a figure of another publication. Accordingly, the quality of the data is a major drawback of the study (e.g. arbitrary smoothing and AMOC metric). Their main problem in fitting the data might be due to the specific metric that is shown in the Rahmstorf et al. (2005) figures, so it is a shame that the authors are not able to resolve that.

R: Indeed, the data has been obtained from the figures of the paper. However, it is not visually extracted as the reviewer suggests. The figure we used is a vector graphic and the dataset can be retrieved from it by inverting the plot matrices used to map the original data to the values in the graph. We can replicate the data in this manner. In order to validate this method, we asked for the data from Rahmstorf et al. (2005) but have not received a reply at present. Less smoothing, and presumably larger noise levels, would likely show a stochastic collapse more easily. We will elaborate on the data acquisition in the paper but want to emphasize that the main goals is to develop and test a method to capture complex model behaviour with a simple low-order model. The methodology described here is not affected by smoothing, only the assessment how well the method works is somewhat hindered by this.

5.) The description of their method contains many errors, and is incomplete. An explicit expression for the likelihood, as well as details of the Metropolis-Hastings implementation are missing.

R: We agree that more detail could be given, but we believe a description of the Metropolis-Hastings algorithm is too much detail for this paper as it is well established and already described in many textbooks. We will add a reference to the textbook of Bernardo & Smith which described it.

In the discussion, the authors name difficulties in the numerical implementation as a possible reason for the failure of their fit to describe the lower AMOC branch, but it is for the reader not possible to assess whether this is relevant, since no details or robustness tests are given. Furthermore, it is not stated how many data points the respective data sets contain, and it is not mentioned that the authors assume successive data points to be independent.

R: Indeed, it is important to assess the robustness of our implementation. To further detail the validity methodology and outcome, we will add a table with model characteristics and state that each point is independent.

It is also not mentioned how the maximum of the posterior parameter distributions is picked.

R: The most likely value is the mean of the posterior distribution. This does assume that the posterior distributions are unimodal. We will discuss this in the text.

6.) Finally, several questions regarding the methodology. a) Why do the authors not try to estimate sigma with their Bayesian method?

R: In principle this can be done. The variation in the hysteresis loops appears constant and can therefore be estimated more easily by other means. This does add to the computational costs and expands the search space, making it more difficult to find solutions. Therefore, we did not follow this method. We will mention this in the text. Why not include observational noise?

R: Observational noise of the real AMOC would have a larger spread. Synthetic series on the basis of the found parameters could be generated with such a noise level. However, we intended to fit the intermediate complexity model outcome, using the data published. AMOC collapse and its likelihood at a given point in parameter space is model-dependent. One of the essential parameters in this, is the model-dependent sigma. Therefore, we preferred to estimate the sigma of the particular model.

This could handle the fact that the data is filtered arbitrarily. It could also completely change the locations of the inferred bifurcation points.

R: The bifurcation points are determined by the limit (non-stochastic) solutions. A noise driven transition could occur, however, and push the points that bound the hysteresis curve further inwards, towards each other. We will mention this limitation more clearly in the paper.

b) To make the paper more understandable it would be good to note explicitly early in the manuscript that the movement of mu is actually known.

R: Agreed, we will add "The forcing values of mu are known and the same for each climate model." However, the values of mu+ and mu- are not, and model-dependent.

c) Why not try multiplicative noise? (see also e.g. Das/Kantz Phys. Rev. E 101, 062145, 2020) This should relatively easily give a model that describes the asymmetric behavior.

R: Multiplicative noise is state dependent, while we have made the assumption that the noise level is constant; therefore, we used additive noise. We will discuss this point in the text.

d) It should be noted explicitly that there is no time dependency of the data. I wonder why they choose not to fit to time series instead? This would allow to treat the non-equilibrium nature of the data. Also, it would be much more applicable to observational data and to make predictions.

R: Indeed, we will add a remark that each data point corresponds to a fixed value of mu and is not time dependent. Each mu value represents a separate climate model run which has run (more or less) to equilibrium. We need to be able to estimate mu+ and mu- from the models, and this can only be done from a hysteresis-curve which indeed contains equilibrium solutions and no time-dependence. We agree that a logical next step is to apply the method on time-dependent runs, but to validate the model it is needed that we know the equilibrium bifurcations points as well for that model, and that the time-dependent runs are based on a slowly-changing mu and include an AMOC collapse. There are not many models available that answer all these requirements.

e) Why not only move along alpha at a certain fixed beta? Is moving both parameters supported by the data significantly better?

R: Early attempts with a fixed beta resulted in worse fits; therefore, we opted not to restrict beta. This will be mentioned in the revised text.

Specific comments Abstract: “Machine learning”: To my knowledge MCMC is not considered a machine learning technique. The abstract needs to be expanded to better reflect the motivation of the study, what their method enables them to do, and their conclusions.

R: We will remove mentioning machine learning. We will also add to the abstract: “The Langevin model allows for comparison between models that display an AMOC collapse. Variation between climate models studied here is mainly in the strength of the stable AMOC and the strength of the response to a freshwater forcing.”

P2L42-45: This is a not a very clear explanation of the salt advection feedback. The main point is that North Atlantic salinity anomalies (positive/negative) are amplified by their effect on the overturning flow (strengthening/weakening), the strength of which controls the North Atlantic salinity. This is thus a positive feedback and leads to bistability with the associated possibility of abrupt transitions.

R: We will rewrite as suggested.

P2L53: “. . .number of solutions for a given value of the freshwater forcing goes from 2 to 3. . . “. Should say “goes from 3 to 1” as the bifurcation point is crossed. (There are 2 solutions precisely at the bifurcation point, but I think this saddle-node fixed point is not relevant here.)

R: We will correct this as suggested.

P3 Caption Fig.1: The terminology of this figure is not appropriate and furthermore not understandable at this point within the manuscript. No trajectory is shown, but a bifurcation diagram.

R: Correct, we will rename to "bifurcation diagram".

They have to be more specific with what they mean by a deformation of the "trajectory". Also, at this position within the manuscript, it is completely unclear what they mean with "trench of the distribution". Either leave out or explain in the main text.

R: We will remove the use of "trench" from the text.

Furthermore, I suggest to use the term "resurgence point" for mu-, and use that terminology throughout the paper.

R: Suggestion in agreement with our other reviewer, we will replace with "resurgence point".

Note that e.g. in P5L91, mu+/- are being referred to as "collapse points".

R: We will replace "collapse points" with "bifurcation points".

P4L64: Can the authors elaborate why they think a double well potential has mainly been studied qualitatively? I would argue that this simple and general mathematical model has been studied quantitatively to an exceptional degree.

R: The reviewer is correct that this model has been extensively studied and applied. But to our knowledge it has not been quantitatively applied to AMOC collapse in complex numerical models or observational data in the way as we present. We will replace "studied mainly in a qualitative way (within catastrophe theory)" with "studied mainly qualitatively in connection with the Langevin equation." Bolton &a (Boulton, C., Allison,

L. & Lenton, T. Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model. Nat Commun 5, 5752 (2014).) do study an AMOC hysteresis loop qualitatively, but do not mention the Langevin equation. We will add Bolton &a to our references and discuss its relevance as a quantitative study of AMOC bistability. Specifically, the transient run studied in that paper and how it could relate to the Langevin model.

P4L65ff: It is a bit confusing when the authors first say that 2 parameters are enough to describe bi-stability, but then use another 2 parameters to scale and shift to the AMOC variable. Maybe it would be better to first explicitly say that by a shift and scale of the variable x, one can eliminate the third order term as well as the fourth order coefficient.

R: Correct, that is their purpose: to reduce the polynomial to a smaller set such that only the minimal number of parameters remain. We will add a clarification as suggested.

Both of these transformation do not influence the global bifurcation behavior. Then, they can state that a shift and scaling is considered when fitting to the climate model data.

R: True, because the topology is not affected. We will add a remark as suggested.

P4L78-81: Can the authors elaborate why they obtain these rough estimates for the parameters, and how they are insensitive to other parameter values?

R: These are not estimates but interpretations that can be linked to the bifurcation diagram. We will replace “The value of nu is roughly . . . “ with “In the bifurcation diagram the value of nu is roughly . . . .” And likewise for lambda.

P5L90-91: When speaking about “solution” what exactly do the authors mean?

R: In this instance we mean that only the trivial solution exists: only 0 as the value for all variables.

P5L92-97: This section is a bit unclear. Can the authors define a “track”, and what does it mean to be one-dimensional?

R: We will remove the notion of a "track". It being 1-dimensional means that because the hysteresis loop is 1-dimensional the values (alpha, beta) as a function of mu are as well.

The fact that alpha and beta are called normal and splitting factor is better mentioned earlier.

R: Second mention to be removed. A more clear distinction of "parameter" and "variable" would be appropriate.

R: We will alter the text such that the Rahmstorf set has data mu and psi, alpha and beta are the stability parameters (which in turn are expressed as a rate and offset), nu and lambda are the scaling parameters, and mu+/- the bifurcation points.

P5L101: This argument is unclear to me. The fact that the AMOC is scalar variable should not constrain the path through the stability landscape in any way. Do the authors rather want to say that in the climate model experiments there is only a single control parameter mu, and that by assuming a linear dependency of both alpha and beta on mu, they can express some parameters by the extremal values of mu?

R: Indeed, mu is the only control parameter. By assuming linearity a reduced set of equations can be determined later on.

P8L127-129: Maybe the authors can elaborate more specifically on why these arguments are relevant in order to neglect a non-linear change of either mu or alpha/beta?

R: We will remove these lines, they are not needed for the argument.

P8L141-146: Improve this explanation. When introducing stochasticity, the asymptotic dynamics for each parameter value give rise to a stationary density.

R: Correct, to be added.

In the case of the scalar potential, this distribution can be given analytically up to a normalization factor. Thus, the distribution can be used as a likelihood function (if I understand correctly) for parameter inference with MCMC.

R: Correct, to be added.

P9L157: What is the “sampled” bi-modality region?

R: Sampled as in where the dataset has values. We will drop “sampled”.

P9L160: “These changes correspond directly to the potential functions in Fig. 2.” What is meant by this?

R: We mean that the distribution functions can be linked to the different characteristic shapes of the polynomial catalogued in fig 2. We will rewrite as “Each distinct shape of the distribution can be linked to one of the potential functions in Fig 2.”.

P10L174: Can the authors explain why lambda < nu in general?

R: The offset (lambda) cannot exceed the scaling factor (nu) because the offset needs to be roughly in the middle of the two stable branches.

Caption Figure 5: The terminology is unclear. What is meant by a “singular” maximum?

R: We mean unimodality, to be rewritten as such.

What is meant by the dominant and the weak mode, and what is the inversion?

R: For a bimodal distribution there is a mode with more probability mass than the other which we call the dominant mode and the model with less mass the weak mode. And inversion is where these modes switch in strength: the dominant turn weak and the weak turns dominant.

There are also grammatical errors (“...in the middle and inversion from weak. . .”).

R: “in the middle and inversion from weak to dominant takes place” to be removed from caption.

P11L180: I think a more precise statement would be that they estimate the posterior probability distribution of the parameters, given the data psi(mu). Furthermore, the following equation does not define the likelihood but the posterior distribution.

R: Yes, correct; will be rewritten as suggested: the equations states that the posterior distribution is proportional to the likelihood multiplies with the prior distribution.

P11L183: Why linearised?

R: To be removed, this related to linearisation of beta0, delta beta, etc

P11L184-188: This statement of Bayes' rule is not correct, please revise. The right hand side is not called Bayes' factor (which arises in model comparison).

R: We will remove L187-188.

P12L209-219: The constants muS+ etc. are not properly introduced and should be shown in one of the figures.

R: These are given in the caption of fig 6. We will add a reference to the figure on the text.

The footnote 2 needs to be explained better.

R: This is a technicality in how we defined the optimiser. We will rewrite to explain that this is useful to avoid solutions that intersect the B1 or B2 twice (see also P17L308-310 below).

Caption Table 1: Why is a linear function used and not a higher order polynomial? This does not seem to be very suitable to the data.

R: We only fitted the beginning (left part) of the upper branch and assumed a constant sigma. We will add this as a clarification to the text.

P13L228-230: This is not very precise wording. What do the authors mean with "unstable" and "more stable" solutions?

R: We will remove “more”. Stability relates solely to the attractors and repellor.

P14L242-245: What do the authors want to say here? It comes as a surprise to me that suddenly only the data for mu > mu+ should be relevant?

R: We ignore the data on the lower branch before the collapse point because we did not want it to influence the fits, especially because we are only interested in the collapse from the upper branch.

And why do they now claim that the model C-GOLDSTEIN does not “appear” to show a collapse?

R: To be removed: our intent was to comment of the smoothness of the trajectory, but it is unnecessary,

P14L252-253: Unclear what the authors are trying to say.

R: We will remove these lines, they are redundant.

P14L259: What is meant by “non-linear degradation”?

R: We mean the part of the hysteresis loop after the collapse point: before that point the change was fairly linear, but after it is strongly non-linear.

P15L267: In what way is the model sufficient to describe the data? Certainly the “re-invigouration” is not well captured.

R: The aim was to model the collapse; the resurgence appears more difficult. We will add a clarification to emphasise this point and estimate a goodness of fit.

P17L308-310: Unclear what is meant here. What are “non-admissible” solutions?

R: An inadmissible solution is one where the curve through (alpha, beta) space intersects one of the subspace B1,2 twice. Because B1,2 are concave this is a possibility. To be discussed in the revised text.

P17L313-315: Unclear. Smoothing might be due to other reasons?

Printer-friendly version

Discussion paper

R: Perhaps, but is not mentioned in Rahmstorf (2005).

P17L323: What is meant by “direct numerical stochastic integration”?

R: This remark was originally intended to point out another way to perform the calculations: by solving the SDE directly. We now believe this remark to be redundant and will remove it.

P17L330: How exactly does this paper present a step forward to assess the likelihood of a future collapse of the AMOC? The method presented here relies on previously modeled collapses of the AMOC with realistic climate models. How does the method generate additional information?

R: If the characterisation has predictive values, more complex models can be used to derive a collapse point if the freshwater forcing at the bifurcation points can be estimated. It is also a way to compare the collapse characteristics of various models. If the freshwater forcing at the bifurcation points can be generalised and linked to a robust indicator (such as, perhaps, Mov), the method can be applied to the real world as well. We agree, there are quit some steps in between the method outlined here and its extension to the modelled and observed timeseries. We will expand the discussion on this point.

Technical corrections

R: We are in agreement with all corrections below

P1L15: last glacial maximum and early holocene -> last glacial period

R: To be replaced as suggested.

P2L26: . . . which are presumed to be functions . . .

R: To be corrected.

C6P2L30: ...and tipping points in the climate, it has not been . . .

R: To be corrected.

P2L38: . . . or an increased surface freshwater flux by changes in precipitation minus evaporation. . .

R: To be corrected.

P2L46: scalar variable obtained by integrating . . . To be corrected.

P2L52: . . . one of the two basins of attraction vanishes . . .

R: To be corrected.

P7L126: rather “(delta alpha, delta beta)”?

R: Correct: changes in alpha, beta: to be corrected. P8L142: Remove: As shown by Cobb (1978), this distribution belongs to the exponen- tial family.

R: We will remove this.

P8L144: The polynomial potential introduced in the previous section, we had already. . ., gives the probability . . .

R: To be corrected.

P8L148: Note that C = C(alpha,beta), which does not have. . .

R: To be corrected.

P8L150: . . . because of the scaling . . .

R: To be corrected.

P8L152 and Fig. 4 caption: In what way is this a sample collapse trajectory, or an example trajectory?

R: This should indeed be example, not sample.

P9L161: . . . a change in only one . . .

R: To be corrected

P9L163-165: Why not say this at P8L148? It is a bit redundant otherwise.

R: We will move these lines to P8L148.

P10L168: arrived at -> described

R: To be corrected

P10L171: independent of each other

R: To be corrected.

P11L180: The method used is not considered machine learning.

R: To be removed.

P11L190: Does not seem to be relevant, as it is not done here.

R: This sentence relates to "These resultant posterior distributions can, in turn, be used as prior distributions, yielding a chain of sampled parameter vectors." It is roughly how the sampler works, but we will remove this because it is redundant.

P11L194-196: This is partly redundant, and it is not clear why the authors mean that the model can be fit with uninformative priors.

R: We will remove the part about uninformative priors: it is unnecessary.

P11L199-200: Redundant.

R: We will replace "With ν and λ introduced earlier, the state variable x undergoes an affine transformation and normalises the polynomial. These ..." with "The parameters ν and λ . . . "

P11L207: Redundant.

R: To be removed.

Printer-friendly version

Discussion paper

P11L207-208: An overview of priors is: The following prior distributions are used:

R: To be corrected.

P14L243: do -> to

R: To be corrected.

P14L254: Why “sample” paths? The authors are showing distributions, which is exactly contrary to showing sample paths.

R: We will remove “ . . . for linearly parametrised sample paths through the stability space” . . . “

P14L255: Redundant.

R: We will remove “The β parameter changes linearly and α follows from the constraints in Eqs. 4 and 5. Blue and red lines indicate the prior bounds for $\mu-$ and $\mu+$, respectively.”

P15L265: couples -> models. Why “additionally”?

R: This is not needed: we will remove “Additionally, a linear parameterisation through state space couples to the freshwater applied to the North Atlantic subtropical gyre region.”

P16L275: Do they mean arg(max(|psi|)) ?

R: We mean max(|psi|), to be corrected.

P17L309 till its -> until it

R: To be corrected.

Interactive comment on Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2020-43, 2020.