We would like to thank the two referees for their detailed discussion of our manuscript and their valuable suggestions. We hope that in the following we can clear some misunderstandings and answer the questions raised in a comprehensive manner.

Reviewer 1:

1. ***The manuscript could use more citations, contrasting its methodology and findings to previous work on probabilistic AMOC projections. It advertises itself as "the first probabilistic assessment of the future AMOC behavior using a calibrated conceptual model and global mean temperature data for the RCP3-PD and RCP4.5 emission scenarios". This wording is dangerous, as it can be interpreted to mean that this manuscript is the first to present probabilistic AMOC projections. There is an existing literature on this subject.***

Answer:

We thank the reviewer for the comprehensive list of suggestions, which is highly appreciated. We added several references in the introduction section at page 1 and 2. For further clarification, we changed the mentioned paragraph to:

> *"In summary, we presented a probabilistic assessment of the future AMOC behaviour using a calibrated conceptual model and global mean temperature data for the RCP3-PD and RCP4.5 emission scenarios."*

We also changed a similar paragraph in the conclusion section from

> *"This is to our knowledge the first probabilistic projection of dynamic sea-level rise and an example for the potential of a modular approach in climate system projections within the limits of interpolation."*

to

> *"This probabilistic projection of dynamic sea-level rise is an example for the potential of a modular approach in climate system projections within the limits of interpolation."*

2. ***There is additional literature that could be reviewed concerning the dynamic sea level effects of AMOC variations.***

Answer:

Several references discussing this relation in further detail are added in section 7 on page 6 of the manuscript.

3. ***It is not entirely clear to me how the predictive envelopes for the various projections are obtained, given the different AMOC emulators. Are the projections from the five different emulators superimposed assuming each is equally likely, and a common predictive envelope constructed for this mixture of models?***

Answer:

Yes, the models have been given the same weight each. While there are observations from which the overall strength of the overturning can be inferred, we find that available data does not allow to discriminate between models with respect to their ability to model future changes in the AMOC. We thus assume that our five models can serve as independent realisations of future AMOC response to external forcings. We combined the 600 random representation pathways of the MAGICC output with each of our five models leading to 3000 different AMOC mean pathways that are considered equally likely. Figure 4, 5 and 6 are based on this data set. To make that point clearer, we modified the following paragraph (section 5 on page 5) of our manuscript

> *"We then combine each of the 600 realisations with our five AMOC emulator settings to obtain a distribution of AMOC responses under these scenarios."*

to

> *"We then combine each of the 600 realisations with each of our five models leading to 3000 different AMOC mean pathways that are considered equally likely."*

4. ***There are several points at which the manuscript fails to propagate uncertainties (or at least, fails to mention uncertainties being propagated).***

Answer:

We agree that there is a need to clarify the method used in our manuscript more explicitly and are grateful for this comment. We are indeed unable to propagate all errors in this paper and have made this more explicit in the text.

5. ***Forcing the AMOC box model with MAGICC temperature projections, without first adding noise to the temperature projections to simulate natural variability***

Answer:

We appreciate the suggestion by the reviewer to add noise to the system prior to the temperature projections. We have decided against it for the following reason. The idea of our study is to provide estimates of the expected mean change not considering internal variability. In fact, conceptual box models as the Stommel-type model used here are not capable of capturing year-to-year variability but are designed to describe long term trends of the mean AMOC strength. Introducing noise on a subdecadal time scale would thus not be consistent with the physical processes included in the model, while decadal noisy variability would be problematic in a century long projection.

6. ***Fitting the AMOC box model to GCM output without propagating the parametric uncertainty in the fits. Figure 2 and 3 indicate a large amount of noise in the GFDL model and one could imagine a wider range of its fits being possible, at least for that model. For example, the validation plot in Figure 3 shows a predicted AMOC decline that is compatible with, but probably stronger than the actual decline visible in the GFDL output. Could other, almost-as-good fits to the GFDL output in Figure 2 produce a weaker decline in Figure 3 that are also compatible with the validation data?***

Answer:

What is obvious from Figure 2 and 3 is that the inter-model variability of the AMOC strength is much larger than the model specific inter-annual variability. The ensemble spread of the equilibrium AMOC strength ranges from 25 Sv to 15 Sv and also the difference in the freshwater response is significant. The GFDL model for example weakens from a strong equilibrium state of 25 Sv by about 8 Sv within the 100 years forcing period and recovers within 50 years afterwards, whereas the NCAR model weakens only by 4 Sv from 16 Sv to 12 Sv, but does not fully recover in the 100 years after the forcing has stopped. We hence assume that fitting the box model to the different AOGCMs accounts for the major part of the parametric uncertainty of the box model, while the AOGCM specific parameters uncertainties seem to be much smaller. To capture these great dynamical differences in our non-linear model with a narrow convergence-space we had to adjust our parameter manually.

However, in order to underline that the associated uncertainties are small in comparison to other sources of uncertainty considered by us, we provide an estimate of the possible variation in the projections of the AMOC strength by “almost-as-good” fitting parameters estimated from experiment (b) (see Figure 2 of the manuscript). As an example we focus on that parameter, since it turned out to have dominant influence on the simulated

AMOC strength under global warming. Likewise, we concentrate here on the GFDL model, since it shows the strongest variability. Oscillatory behaviour is much less pronounced in the other four ensemble members (compare Fig. 3).

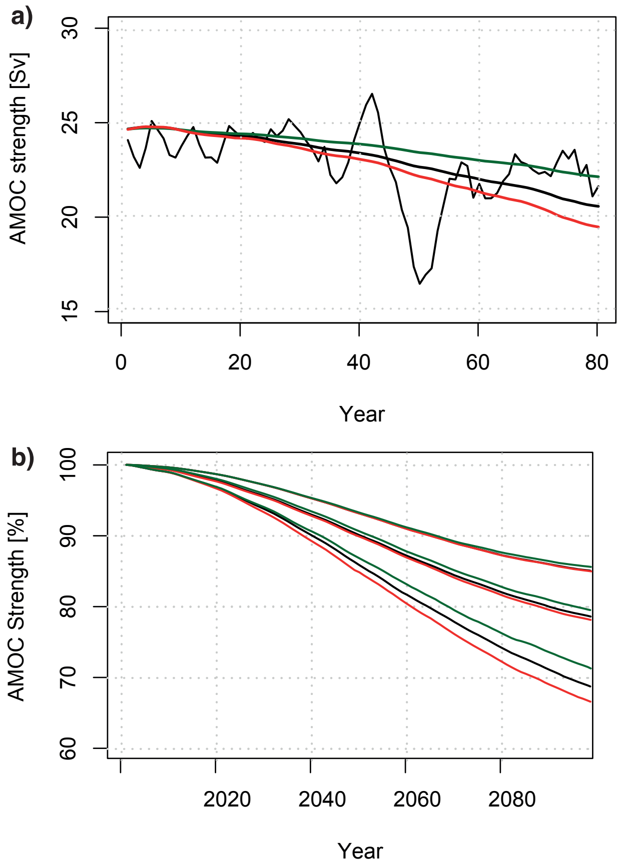

Figure R1. Parameter variations for GFDL, $p = 0.4$ (control,black), $p = 0.25$ (green), $p = 0.5$ (red)
a) Calibration results as in Fig. 2b) of the manuscript,
b) Ensemble emulation results for RCP4.5 as in Fig. 4 a) of the manuscript, 90th percentile (upper), median (middle), 10th percentile (lower)

We chose p=0.25 and p=0.5 as a lower and upper bound, where the reproduction of the AMOC behavior in the warming experiment is already rather poor as shown in Fig. R1 a) that corresponds to Fig. 2b) in the manuscript. The influence of this varied parameter on the ensemble behaviour under the RCP4.5 scenario is shown in Fig. 1 b). Even strong variations of the most influential parameter change the ensemble median by less than 1 %. Only the 10th percentile shows a more pronounced change by 2 % whereas the upper bound remains virtually unchanged. This is remarkable, since in this case we vary the parameter for one model by 25 %, which already results in a strong reduction of the model emulation capacity, whereas the same parameter varies between the models by a factor of 9. It appears that the impacts of parameter variations of single models are of secondary importance compared to the variations between the different models, that differ up to 6 % in the median projection for the RCP4.5 scenario as shown in Fig. R2. We will provide this figure also as supplementary material to the manuscript.

Therefore, we would like to argue that our approach of using a five-model ensemble with inter-model parameter differences up to one order of magnitude covers most of the uncertainty related to the emulation parameters. Additionally, we restricted ourselves to interpolations within the calibration range, to avoid uncertainty amplification.

We introduced the following modifications in the manuscript:

Page 3, Section 4:

> *"... and the parameter-set was manually optimised to reproduce each AOGCM output (Table 2)."*

Page 4, Section 4: The following text is added at the end of section 4:

> *"The uncertainty associated with single emulation parameters of the different models is much smaller than the inter-ensemble spread. Thus, we account for the major parametric uncertainty component when assuming all five emulator configurations obtained here as equally likely representations of the AMOC and quantifying the resulting spread."*

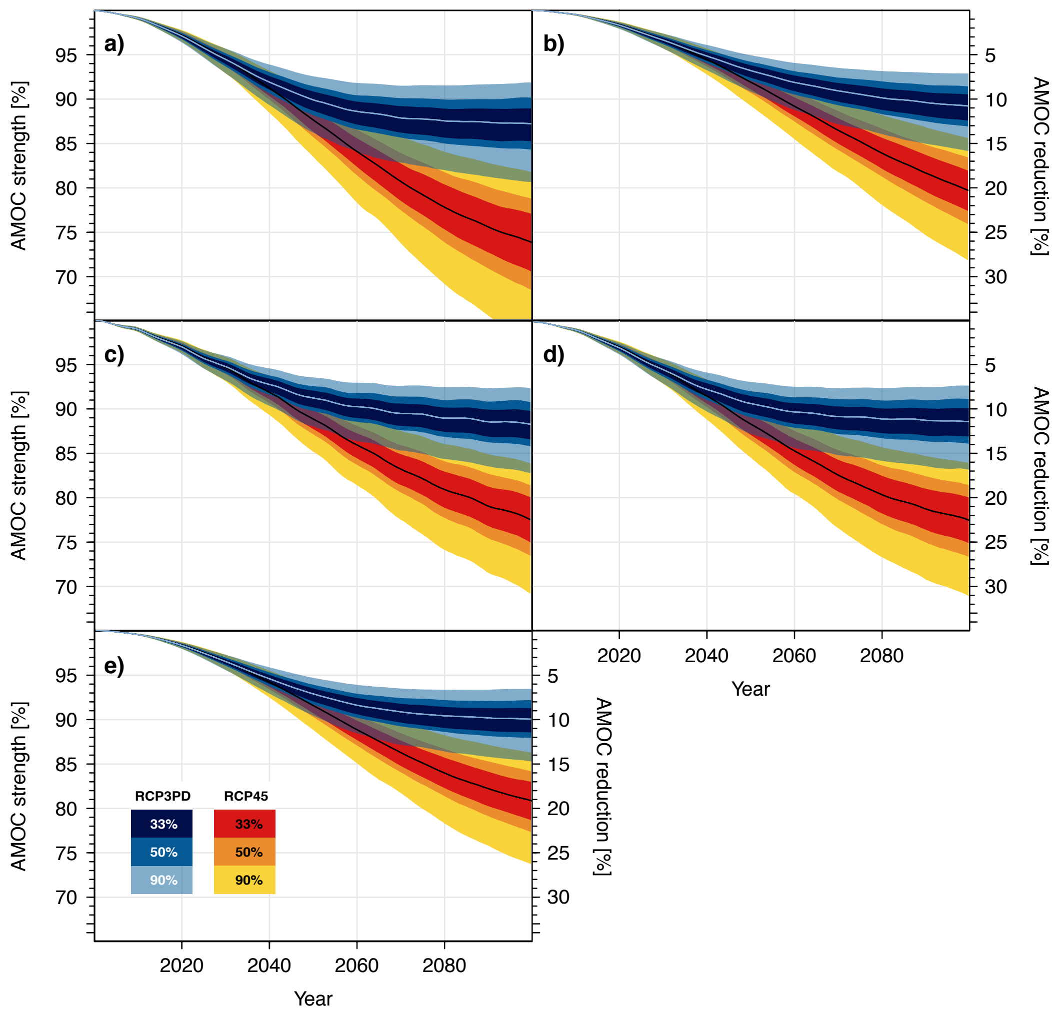

Figure R2. The projected AMOC weakening as in Fig. 4 of the manuscript for each ensemble member separately: **a)** GFDL_R30, **b)** MRI_CGCM_2.3, **c)** MPI_ECHAM5, **d)** MIROC3.2, **e)** NCAR_CCSM2.0.

7. Projecting the AMOC weakening from the box model output without adding noise to the AMOC projections to simulate natural variability

Answer:

Continuous AMOC measurements are just available since 6 years which makes it very difficult to make profound statements about the current state of the circulation,. Furthermore, these observations reveal an inter-annual variability of the AMOC ranging from 10 to 30 Sv (Cunningham et al. 2010), and a strong interdecadal variability, sometimes associated with the Atlantic Multidecadal Oscillation (e.g.. Park et al. 2008; Dong et al., 2005 and references therein).

In our opinion, it is by no means justified to assume that all of these effects are covered by all the five AOGCMs used as a basis for our emulation, which means that taking the models internal variability and adding it as a noise signal could greatly underestimate the uncertainty. Furthermore, whether or not this variability might change under global warming scenarios is very uncertain and assuming a constant variability might be also misleading in this case.

Therefore, we decided to restrict ourselves to projections of the AMOC mean, as it is done in similar studies, e.g. Zickfeld et al., (2004) and Urban and Keller (2010).

For further clarification, we changed all phrases denoting *"project ... the further evolution of the AMOC"* or phrases with similar meaning to *"project ... the further evolution of the AMOC mean strength"*

e.g. in the abstract:

"we project the future evolution of the AMOC mean strength within the covered calibration range"

8. Projecting dynamic sea level rise from a regression of sea level rise on MOC weakening, without propagating the uncertainty in the regression. (This properly should use the regression prediction interval, not to be confused with the parametric confidence interval of the regression slope coefficient.)

Answer:

As the reviewer suggested, we have repeated our SLR projections including uncertainty propagation of the regression procedure, the updated Fig. 6 of our manuscript is shown in Fig. R3. The reviewer suggested to use the regression prediction interval, but since we restrict ourselves to projections of the mean behaviour as discussed above, propagation of the individual parameter uncertainties is sufficient. These uncertainty intervals are shown in an updated version of Table 3 of the manuscript in Tab. R1.

	a [cm/Sv]	b [cm]
GFDL CM2.1	1.68 ± 0.08	3.30 ± 0.56
MIROC MEDRES	2.81 ± 0.14	1.95 ± 0.62
MPI ECHAM5	2.74 ± 0.26	2.63 ± 0.65
IPSL CM4	2.58 ± 0.15	2.32 ± 0.67
MIROC HIRES	1.45 ± 0.21	4.01 ± 0.59

Table R1. Update of Tab. 3 of the manuscript

To account for the parameter uncertainty of the linear regression parameters, we randomly picked a value out of a Gaussian distribution with a standard deviation of the parameter uncertainty and combined it for each of the five models that were analysed in Yin et al. (2009) with our 3000 AMOC representations.

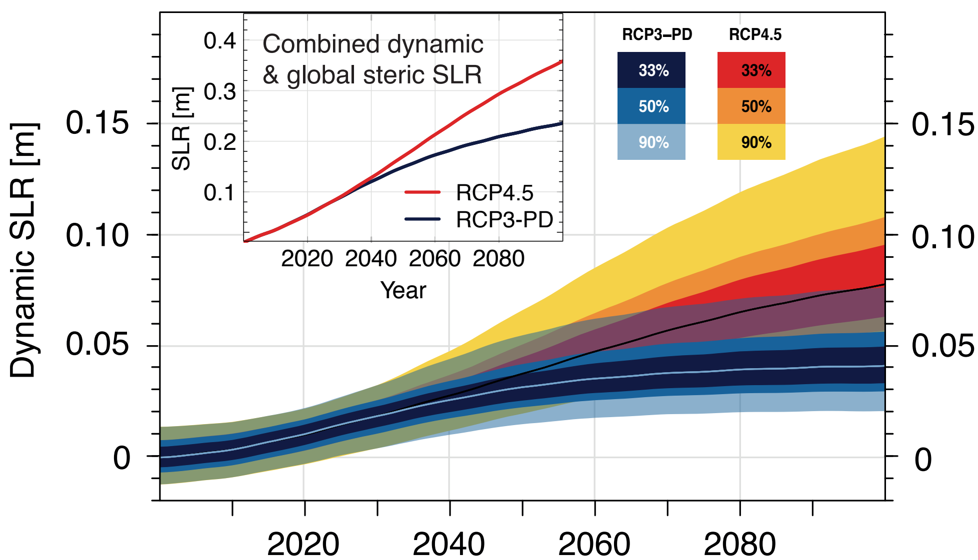

Figure R3. Update of Fig. 6 of the manuscript

The main difference between Fig. 6 of the manuscript and the updated version is a broader uncertainty envelope for the 2000-2050 period due to the uncertainty of the offset-parameter b, whereas the projection-envelopes for 2100 remain virtually unchanged (maximum changes are less than 1mm). Again, most of the parameter uncertainty is already covered by the inter-model variability of our ensemble.

We accordingly updated Fig. 6 in the manuscript with Fig. R3 shown above and the following paragraphs of the manuscript have been modified:

Section 7, Page 6:

> *"By combining these linear regression results of all five models with our AMOC projections for different emission pathways we can provide probabilistic projections of the dynamic sea-level rise in the New York City region (Fig. 6)."*
>
> *"To account for the parameter uncertainty of the linear regression parameters shown in Table 3, we randomly picked a value out of a Gaussian distribution with a standard deviation of the parameter uncertainty and combined it for each of the five models with our 3000 AMOC representations. Thus, we provide probabilistic projections of the mean dynamic sea-level rise in the New York City region (Fig. 6)."*

Caption of Fig. 6:

> "Probabilistic projections of the dynamic sea-level rise at the New York City coastline for the RCP3-PD (blue) and the RCP4.5 (orange) emission pathway till 2100."

to

> "Probabilistic projections of the *mean* dynamic sea-level rise at the New York City coastline for the RCP3-PD (blue) and the RCP4.5 (orange) emission pathway till 2100."

9. What is responsible for the oscillatory behavior of the MAGICC temperature projections in Figure 4b?

Answer:

The signal is due to the solar cycle that is flattened in Meinshausen et al. (2009). We updated this figure to the flattened version in the manuscript to avoid confusion.

Reviewer 2:

10. One point to make on section 2, concerning terminology. The word "emulator" tends to be reserved (in statistics and more widely) for stochastic representations of deterministic functions, for example as fitted using Gaussian processes. Simpler deterministic models fitted to more complicated deterministic models tend to be called "surrogates". The box model would be a surrogate for the AOGCM.

Answer:

We appreciate the reviewers concern, but would like to keep the notation as we think that it will be best understood in the community that we aim at. Fielding this community the term "emulator" is usually used to describe simplified models describing more complex systems, so it should be more intuitively understandable then surrogate. For example the simple climate model MAGICC providing our global mean temperature projections is also considered to be an AOGCM emulator (compare Meinshausen et al. (2009)). In a very similar context to the presented one a simple box model is as well described as an AMOC emulator (Zickfeld et al. (2004)).

11. I would like to have more information about how the box model was initialized for each AOGCM. I would like a lot more information concerning the statement "the parameter set was optimized to reproduce each AOGCM output". Inspecting Figure 2 I guess that the authors have minimized the sum of squared residuals between the box model output and the AOGCM. Why would that be a good idea? Well, the resulting parameter estimate will be the maximum likelihood estimate if it were the case that the residuals are IID Gaussian with mean zero. With the exception of GRDL_R30 this would seem to be a reasonable assertion about the residuals. I would like to make two further points. First, the marginal variance of the residuals appears to be increasing with the AMOC strength. This suggests to me that a better criterion would be to minimize the sum of squared residuals of the logarithms. Second, the authors have missed an opportunity to use the very detailed theory of maximum likelihood estimators to provide an uncertainty estimate for the parameters.

Answer:

As explained in the answer to reviewer 1, our model parameters are manually adjusted to capture the great dynamical differences between the ensemble AOGCMs with our non-linear Stommel model. The Stommel model as it is reported in the literature (e.g. Stommel (1961), Rahmstorf (1995)) shows a strong non-linear behaviour with regard to freshwater forcing with a very narrow, in our case five dimensional, convergence-space. An implicit description of the convergence space is not available.

Within our model ensemble we nevertheless cover a wide range of different parameter configurations that account for the majority of the parameter uncertainty as discussed in detail above (Answer to point 6 of reviewer 1).

12. Section 5 also raises some interesting questions. It appears as though the five surrogates, standing in for the five AOGCMs, have been averaged together. What is the epistemic principle behind which justifies this? One can see that the five AOGCMs vary widely (e.g. Figure 2 and 3) so it follows that some of them match the recent observations better than others. So, on this basis it would have been more prudent to show each AOGCM separately.

Answer:

As discussed above, only a few years of observational data is available, which does not allow for a weighting of the different AOGCMs according to their agreement with the observations. We instead assumed each of the five models to be equally likely representations of the AMOC and the ensemble spread to cover the parameter uncertainty. The individual AOGCM are shown separately in Fig. R2. We will provide this figure also as a supplement to the manuscript.

13. I am also concerned that Figure 4a underestimates the uncertainty. If the tuned box model is standing in for the AOGCM, then the error in its approximations must be incorporated into the simulation. But the text does not suggest that this has been done; and, indeed, because it needs estimates of sigma, I doubt that it has. The authors need to add a Gaussian IID error to the box model output with standard deviation sigmahat, in order for their simulation to account for the difference between the tuned box model and the AOGCM. This could add a few Sv to the range of uncertainty at 2100.

Answer:

The question raised here is somewhat similar to point 6 and 7 of the first reviewer. We showed above that strong deviations of the parameter with the highest impact would only alter the ensemble projections by less than 1 %, whereas the inter-model spread is about 6 %. Therefore, we would like to argue that we cover most of the uncertainty connected with parameter variations within our ensemble. The variation of the temperature scaling coefficient of the GFDL model provides a variation in the ensemble median for the RCP4.5 scenario of 0.125 Sv and of at maximum 1 Sv for the 10^{th} percentile assuming an equilibrium AMOC strength of 25 Sv. As these values show, we have to disagree that we significantly underestimate the uncertainty of our projection.
Concerning the reviewer's statement of adding a few Sv to the uncertainty to account for the residual variability we want to underline again that we just want to describe the mean AMOC strength (see e.g. our more detailed answer to point 7 of reviewer 1).

14. ***Continuing in the same vein, the authors should not use the point estimate thetahat for each run of their box model, but should sample from the uncertainty about thetahat, in order to incorporate parametric uncertainty.***

Answer:

The parametric uncertainty of the point estimate thetahat of the individual model fits is not included in our uncertainty estimate as it is small compared to the inter-AOGCM variability as described in more detail above (answer to point 6 of reviewer 1). Therefore we did not sample from the associated distribution.

References

Park, W., and M. Latif (2008), Multidecadal and multi-centennial variability of the meridional overturning circulation, Geophys. Res. Lett., 35, L22703

Dong, Buwen, and Rowan T. Sutton, (2005): Mechanism of Interdecadal Thermohaline Circulation Variability in a Coupled Ocean–Atmosphere GCM. J. Climate, 18, 1117–1135.

Cunningham, A., and Marsh, R., (2010): Observing and modeling changes in the Atlantic MOC, Wiley Interdisciplinary Reviews: Climate Change, 1,2,180-191