Thank you very much for your comments and very relevant questions. Before giving detailed answers to each of your comments, we would like to emphasize the aims and scope of our paper. Our methodology aims to provide country-level estimates for the future damage from tropical cyclones, consistent with climate model projections, with mainly economic and financial applications in mind. Examples of applications are estimating the impact of cyclones on the creditworthiness of government debt; providing a physical risk module for integrated assessment models; stress testing the resiliency of the financial system at a country level and at a global level to physical risk etc. Given this aggregate country-level analysis objective, our model is certainly a simplification compared to state-of-the-art cyclone dynamics models, does not aim for precise prediction of individual cyclone tracks, and does not integrate a bottom-up description of damage to individual assets. At the same time, our paper improves earlier studies of cyclone damage on the aggregate level (for example, compared to Mendelsohn et al. (2012): our model uses several climate scenarios with state-of-the-art bias correction, uses SSPs to project population and GDP, is based on precise asset value distribution etc.) We will include these clarifications in the introduction of the revised version of the manuscript. For reviewer’s convenience, we attach a supplement with further technical elements and figures.

Q1 - Cyclone description accuracy It is not clear that this tropical cyclone model leads to accurate forecasts of storms. Changes in wind patterns can have no effect in this model. Is the cyclone model in this paper as accurate as the models developed by Emanuel? Or is this model a step backwards?

We reiterate that building precise cyclone forecasts is not the aim of CATHERINA. We propose an algorithm designed to assess the future cost distribution for country-level damage assessment. The cyclone intensification process used is inspired from STORM model (Bloemendaal et al., 2020), which includes a single climate variable, and extended following Holland (1997) and Emanuel (1988) to encompass 2 more variables. We found that this extension provides statistically significant instrumental variables in the description of tropical cyclone intensification, which is the aim of the algorithm. Another step forward in our methodology is the use of state-of-the-art bias correction module for
integrating climate model projections.

Consequently, even if some thermodynamical processes have been simplified in this approach, our approach is still a step forward with respect to state-of-the-art in the context of integrated assessment models (IAM) for climate impact analysis. Indeed, our approach can integrate any CMIP simulation with limited set of available variables (only few vertical levels, some only available at monthly time scale, with some variables not always available). The adaptability of our algorithm to any CMIP exercise and simulation comes with a constraint implying necessary simplifications. Our approach combined with a bias correction module makes our algorithm easy to implement, more sophisticated in terms of processes included with respect to existing IAM and bias-corrected. We will emphasize this message in the section 3.2.3 of revised version of the manuscript.

Q2 - Damage and sub-perils The estimates of the effect of each hurricane are crude. The model assumes that all damage is from wind whereas only 40% of cyclone damage is wind related. Another 40% of cyclone damage is from storm surge. But storm surge strikes largely just the coastline. The remaining 20% of damage is from excess precipitation which often falls far from where the cyclone strikes land.

The model does not distinguish sub-perils, associated with key thermodynamical processes of cyclones (heavy precipitation, storm surge and associated flooding, strong winds) but instead uses a statistical relationship to estimate the global damage induced by a cyclone from a proxy variable given by the maximum wind speed (which is the proxy used in Saffir-Simpson Hurricane wind scale to define the intensity of a cyclone). We will include these elements in the third section (3.2.1) of the paper. The damage function is fitted on multiple events from the total damage reported in the global disaster database EM-DAT (Guha-Sapir et al., 2018). This database, used in most studies on the topic, accounts for the total reported damage (sum over all sub-perils) and does not distinguish damages from sub-perils. So, despite the relevance of the reviewer’s comment, for our application (see our comment in the introduction of our reply, which are now included in the manuscript introduction to clarify the context of our study), distinguishing the sub-perils generating the impacts is not needed.

Q3 - Damage function calibration The model depends a great deal on the damage function. But it is not clear how this damage function was estimated.

We use region-specific damage functions from Eberenz, Lüthi, et al. (2020). This method uses a parametric function following Emanuel (2011). The parameters are estimated for each region with machine learning techniques from the reported damage estimates in the International Disaster Database (EM-DAT) Guha-Sapir et al. (2018) crossed with cyclone tracks (IBTrACS), and geographic and socio-economic information along these tracks.

We reiterate the main step of the optimization performed by Eberenz, Lüthi, et al. (2020) and Lüthi (2019) to define the regional damage functions. The authors first defined the event damage ratio (EDR) as a fraction error between normalized reported (NRD) and simulations (SED) for each cyclone and the total damage ratio (TDR) is defined in each region summing over events. For each event, there is a value for vh allowing to optimally calibrate the explicit damage function described in Emanuel (2011). Then, the authors proposed two complementary optimization methodologies to find the value of vh maximizing the prediction of the regional damages Eberenz, Lüthi, et al. (2020):

- Root mean square fraction (RMSF), minimizing the spread of the event damage ratios (EDR) – defined as the ratio of simulated damage vs. reported damage.
- Total damage ratio (TDR), finding the value of vh, such as the ratio of total simulated damage – obtained summing over event damage – and total reported damage tends to 1.
We will clarify the calibration of these functions in section 4.3. In particular, we will review the approach of Eberenz, Lüthi, et al. (2020) to find the values of $v_h$ (c.f. technical supplement).

**Q4 - National Asset** The estimates of how national assets are distributed across space are crude. Light times population is not going to allocate national assets carefully. I am specifically concerned about how well they model the assets near the coast.

We chose to build our model based on state-of-the-art estimates, in such a way that the methodology is uniform country-wise. This dataset (Eberenz, Stocker, et al., 2020) is also used for the calibration of damage functions in Eberenz, Lüthi, et al. (2020) (discussed in Q3). Therefore, the use of this data allows to estimate the exposure in a consistent manner. To verify the accuracy of estimation, a back-test has been performed (Section 4.4). As we mention in the beginning, the only way to improve the estimates of asset value distribution would be to use the actual asset distribution from asset-level databases, but such databases are not yet available at the global scale. We will add this explanation when introducing the exposure dataset section 2.4.

**Q5 - Spatial distribution dynamics** The model appears to assume the spatial distribution of assets are fixed within a country.

The model assumes that the spatial distribution varies with population changes proposed in the Socio-Economic Data Application Center dataset presented by Jones and O’Neill (2017, 2020). In particular, the spatial distribution of the population is different in varying shared socioeconomic pathways (SSPs). These projections are available with a one-eighth degree resolution. Figure 1 (in the supplement) represents this multiplicative factor in the SSP2 (1a), SSP3 (1b), SSP4 (1c) and SSP5 (1d) in 2100. The revised version will include a subsection to better describe the exposure dynamics (spatial and temporal) lacking in the current manuscript (c.f. technical supplement).

**Q6 – National asset dynamics** The paper does allow national assets to change over time, but they do not describe how this is done.

To estimate future exposures along the cyclone track in each scenario, we use the downscaled estimation for the exposed wealth and the coefficients representing the change between the current state and the future scenario. We use the most granular projections of GDP per capita variation curves (Figure 2 – Data Source: https://tntcat.iiasa.ac.at/). Binding the two (regional GDP per capita and local population) we build a dynamic projection of exposure factor. Similarly as for Q5, the asset-exposure dynamics will be further detailed in the final version of the paper (c.f. technical supplement).

**Q7 - Adaptation** There is no effort to measure adaptation by the country being hit or how that might change over time.

Indeed, we left this question for further research. Supposing that adaptation increases with time alone would not be a relevant hypothesis. However, this question could be one of the direct applications of the model. For example, measuring the investment costs required to shift the value of $v_h$ or $v_t$ – and thus reduce the risk of future damage – can be a research question derived from this model simulations. In the revised paper, we will present more clearly the possible application of the model integrating the adaptation scenario, changing the values for the vulnerability parameter ($v_h$ and $v_t$) in the section 4.2.

**Q8 - Bias control** The initial forecasts of windspeed from the climate models are very inaccurate. The corrections appear to matter a great deal. However, these corrections
have been made are on the historic data. So once they adjust historic data to actual historic outcomes, they do fine. But how well the model predicts future wind speeds is unclear.

Our bias correction approach is the standard in the climate community (see http://ccafs-climate.org/bias_correction/)[1]. We do not have reanalysis data for the future. Therefore, there is no ‘reference’ value to evaluate the prediction of the model. This is why we control the bias using the past distributions, where we can compare climate models and reanalysis and assume that errors between the two are similarly distributed in the future. We reiterate that this assumption is relatively classical in the climate community and we will integrate these precisions in the paper in section 5.1.

**Q9 - Results: quantiles vs. expected** Figure 19 suggests the model predicts a small probability of very large damage but an expected value that is quite small. What explains this large tail to the distribution of damage? Is this simply the probability of a large storm striking a large coastal city? What is the expected value of damage?

We ran the 7 models over 300 representative years to obtain these distributions. There is an effect due to certain large coastal cities exposure for the ‘very unlikely’ band (between 95 to 99 percentile) of annual damages. However, given the scale observed more than one city have been hit by storms. Because the aim of the model was also to stress test the resiliency of the financial and economic systems, looking at the expected value of damage was less interesting that studying the quantile value especially in the context of events with large tail risk. Coronese et al. (2019)[2] investigating the increase of economic damage due to extreme natural disasters supports this thesis showing that the impact of climate change is particularly striking for extreme events (See for example, Coronese et al. 2019, Figure 2A). The table containing the expected value of damage after bias correction is in the technical appendix. The revised version will integrate this summary table with the expected value of the damage in the section 5.2 as well as the precisions above to explain the focus on quantiles in the visualization.

**Q9 - Results: SSP vs. RCP components** Why does going from historic (1980-2020) to RCP2.5 lead to more damage than going from RCP2.5 to RCP8.5? Going from historic temperature to RCP2.5 is a 1C increase whereas going from RCP2.5 to RCP8.5 is going from 2C to 5.4C? Given the assumption that wind speed increases more rapidly as sea surface temperature rises, this outcome is hard to understand.

Socio-economic change leads to wider differences than climate change, and this was expected (cf. Mendelsohn et al. (2012), Figure 3 for example). The explanation for this is contained in the dynamics of (i) GDP and (ii) population in SSPs. In the revised version we add further explanation about this result including more references to discuss the results of our simulations.

**Q10 - Results: Countries damages** How much confidence do the authors have that they understand the relative damage caused by tropical cyclones at the end of the century across countries? How much of this is simply assuming the same distribution as today?

Thank you for this very interesting question. We can see in Figure 4 (20 in the paper) that the distribution across countries is different from one SSP to another. For example, we have sensibly the same distribution in SSP2 and SSP5 with a higher expected damage in SSP5 because of the growth hypothesis this scenario relies on. However, SSP3 (rocky road) or SSP4 (inequality) are distributed differently. The scenario emphasizing inequalities –and its interpretation by scientists in terms of (i) socioeconomic developments (Riahi et al., 2017) and (ii) population distribution (Jones & O’Neill, 2017) – increases damage concentration in the United-States. On the other hand, the rocky-road scenario, linked to higher and more rural population, lower GDP and national rivalry sees
the damage more equally distributed on other nations. We integrate this precision in the final version.

**Q11 - Overall critics** It is not likely that anyone could design adaptation measures from this study given the crudeness of both the tropical cyclone predictions as well as the damage predictions. Is there any reliable prediction of a change in tropical cyclone outcomes from current outcomes other than they will get uniformly more powerful?

The current dataset – with low resolution data, and maybe not entirely sufficient number or realizations – might not be accurate enough to calibrate properly adaptation measures. However, we believe that the framework presented here is perfectly adapted to project a dense set of trajectories, compute expected and damage percentile over the next decades and therefore measure the investment required to either adapt, mitigate or include a migration factor in global economic modeling in the next fifty years. This work also reflects a practical exercise not carried out until now, which makes it possible to cross-reference the latest data sets developed, putting into perspective both the socio-economic and climatic development hypotheses, and to carry out a bottom-up, rather than top-down, damage calculation. The conclusion of the revised manuscript will mention the limits of the current application and better explain the scope of applicability of the model.


Please also note the supplement to this comment: https://gmd.copernicus.org/preprints/gmd-2021-384/gmd-2021-384-AC1-supplement.pdf