Dear Anonymous Referee #1,

We would like to start by thanking you for your positive and constructive comments. Please find below a documented list of changes we have made to the manuscript (marked R: in blue font). We have edited figures, added the recommended table and flowchart describing the GCM and synthetic hurricane rainfall event production, and provided additional explanation in the manuscript and SI. We hope that these clarifications will improve the reader's understanding of our work.

Kind Regards,

Leanne Archer

Anonymous Referee #1 Comments

The manuscript provides a valuable analysis of rainfall-driven flood risks from hurricanes in Puerto Rico under current and projected climate warming scenarios. The topic is timely and of significant importance, given the vulnerability of Puerto Rico to climatic extremes. The manuscript is well-written and methodologically sound and provides a significant contribution to the understanding of climate impacts on flood risks in small island regions. However, there are a few aspects of the manuscript that require further attention and clarification before publication. I have provided a few editorial comments on the PDF. Here are my main comments.

R: Many thanks for these comments!

1. Please specify the future climate scenarios (CMIP5/6; SSPs/RCPs).

R: We have added the RCP scenarios used to represent 1.5°C and 2°C into the text where the HAPPI ensemble members are introduced in Section 2.1, starting on Line 168, Page 6:

"Representative Concentration Pathway (RCP) 2.6 was used for model boundary conditions at 1.5°C, using a weighted combination of RCP2.6 and RCP4.5 at 2°C."

2. Lines 416-418: The basic information on these models is critical for interpreting the results. I recommend that the authors summarize the basic information about the GCMs, including their resolution. This information will enhance the reader's understanding of GCMs applied in the study.

R: A table outlining the horizontal resolution, the number of simulated years of climate model data for each climate scenario, and the reference for each of the four HAPPI Global Climate Models has been added to Section 2.1, on Line 175, Page 6:

HAPPI Climate Model	*Horizontal Resolution*	*Number of simulated years of climate model data*			Reference
		Present day	*1.5°C*	*2°C*	
CanAM4	*2.81° x 2.81°*	*332*	*346*	*332*	Wehner et al., (2014)
CAM5-1-2-025degree	*0.31° x 0.23°*	*409*	*365*	*396*	Von Salzen et al., (2013)
ECHAM6-3-LR	*1.88° x 1.88°*	*427*	*378*	*383*	Stevens et al., (2013)

NorESM1-HAPPI	*1.25° x 0.94°*	*423*	*382*	*351*	Bentsen et al., (2013) Iversen et al., (2013) Kirkevåg et al., (2013)

We have also now referenced this table in the above mentioned section of text to make this clearer to the reader, now on Line 466, Page 16:

"However, one climate model (CanAM4) shows the opposite trend above the 10-year return period (see ***Error! Reference source not found.*** *7). One key reason for this is likely to be the differences in resolution of the underlying Global Climate Model (GCM) data: CanAM4 GCM has a coarser resolution (2.81°x2.81°) than the next most coarse GCM ECHAM6-3-LR (1.88°x1.88°) (see* ***Error! Reference source not found.****)."*

And on Line 181, Page 7:

"For each climate model, the number of simulated years was calculated as the sum of the number of simulated events per year divided by the simulated annual frequency of events in the climate model data (see ***Error! Reference source not found.****)."*

And on Line 483, Page 17:

"The lower bound here represents the results from the CanAM4 model, which has the lowest GCM resolution (see ***Error! Reference source not found.****)."*

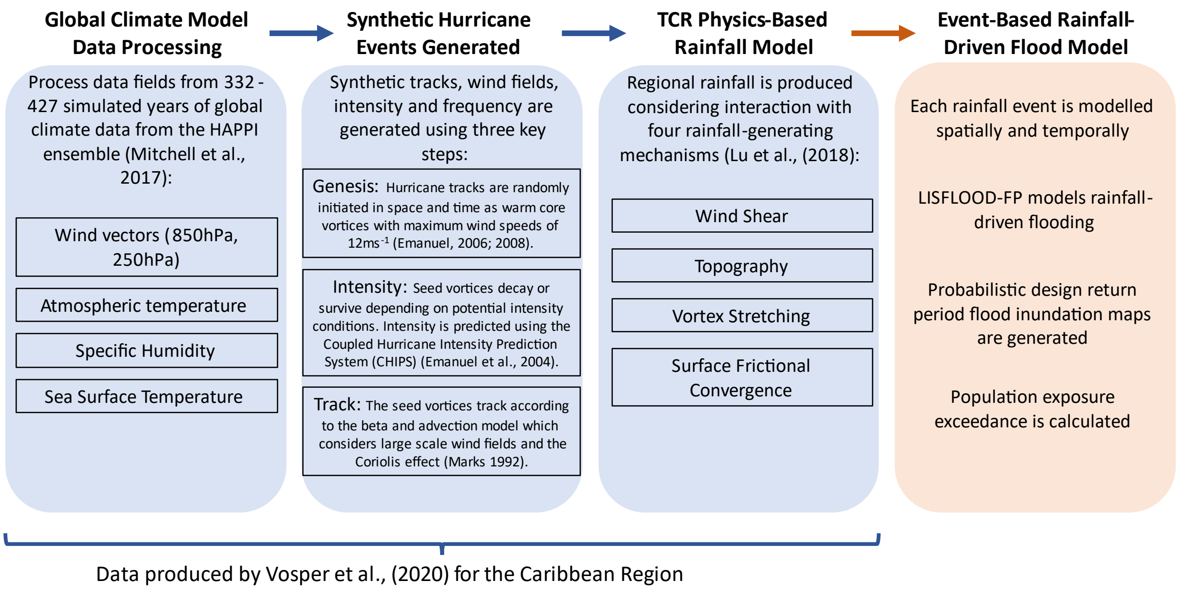
3. This study does not incorporate storm surge component in the modeling process. Rappaport (2014) has highlighted that storm surge accounts for roughly half of the fatalities from Atlantic hurricanes in the US between 1963 and 2012. The authors should offer a comprehensive rationale for excluding storm surge from the model and discuss the implications this may have on the validity and applicability of the study's findings.

R: We agree with your comment that storm surge associated with hurricanes is an extreme hazard that has a large impact on exposed populations. To include storm surge in our model, we would need to additionally simulate the storm surge associated with each synthetic hurricane rainfall event using a storm surge model, before then additionally modelling the associated flood hazard and its interactions with rainfall-driven flooding on the island. Due to the computational costs that would be involved in this - and given the large event set and computational simulation time already involved in modelling rainfall-driven flooding - it was therefore not feasible to also include storm surge in the modelling process in this study. As a result, we decided to focus on inland flooding driven by rainfall, as there are several papers that have demonstrated the considerable impact rainfall has on flooding in Puerto Rico (Smith et al., 2005; Hernández Ayala et al., 2017). However, we do agree that this approach means that the population exposure estimates may be underestimating the absolute numbers of people exposed to overall flooding from tropical cyclones, as coastal flooding is not modelled. As a result, we have added a few sentences in Section 4.3 of the Discussion where we outline the uncertainties in absolute population exposure numbers to highlight this, starting on Line 633, Page 23:

"On the other hand, as this study does not include estimates of coastal flooding, the population exposure estimates may also be an underestimate. This means that it is important to consider that the exposure estimates outlined in this study are for inland rainfall-driven flooding only."

4. The manuscript would be improved by the inclusion of a flowchart that delineates the methodology for generating synthetic hurricane rainfall events. Given that the terms 'synthetic hurricane event' and 'hurricane rainfall event' can be easily misunderstood, a visual representation would greatly assist in distinguishing these terms and clarifying the approach used in the study.

R: We have now included a flowchart in Section 2.1 as Figure 2 starting on Line 160, Page 6 which demonstrates how the synthetic hurricane rainfall events are generated, as well as how these are used as input to the event-based hydrodynamic model:



5. The choice of the future data period (2106 - 2115) needs further explanation. The rationale behind selecting this particular decade should be elucidated to justify its relevance to the scenarios under investigation.

R: We have added the rationale for the future data period in Section 2.1, starting on Line 185, Page 7:

"This future time period was selected in the HAPPI climate ensemble as the future time slice, chosen to represent a 1.5°C and 2°C world at around 2100 (which was the generally accepted time period for these temperature scenarios in the IPCC Special Report on 1.5°C (IPCC, 2018)*), whilst also providing 100 years of simulated GCM data following the present day time slice (2006-2015)* (Mitchell et al., 2017)*."*

6. While IMERG and NCEP datasets are used as observational data sources, there is no discussion about their reliability or the uncertainties associated with satellite data. Authors should elaborate on these aspects, possibly including error metrics or validation studies, to affirm the credibility of these data sources and their suitability for this analysis.

R: We have added a detailed description of the credibility and suitability of these datasets in two additional places in the manuscript, outlining findings from studies that have validated and compared these products. Firstly, we now discuss the capacity of IMERG and NCEP Stage IV to represent extreme rainfall from tropical cyclones in Section 2.4, starting on Line 323, Page 10:

"IMERG has been widely compared to gauge-based rainfall data over many locations globally, demonstrating good performance in estimation of total rainfall (Freitas et al., 2020; Pradhan et al., 2022)*, as well as good representation of temporal* (Yu et al., 2021) *and spatial event structure* (Omranian et al., 2018; Rios Gaona et al., 2018; Pradhan et al., 2022)*. For example,* Rios Gaona et al., (2017) *shows IMERG has a low relative bias over the Netherlands (-1.51%), and* Tan et al., (2017) *reports a correlation coefficient of 0.78 against radar and gauge-based observations in the United States. IMERG has also been shown to perform well at capturing rainfall from tropical cyclones* (Rios Gaona et al., 2018; Yu et al., 2021)*. For example,* Omranian et al., (2018) *found IMERG correctly predicted 62% of rainfall from Hurricane Harvey. Nonetheless, some studies have identified a tendency for IMERG data to underpredict rainfall intensity during extreme rainfall events* (Freitas et al., 2020; Mazza and Chen, 2023; Tian et al., 2018; Yu et al., 2021)*. For example,* Yu et al., (2021) *found that extreme precipitation rates from IMERG were 7.53% lower than gauge data for Typhoon Lekima in 2019.*

NCEP Stage IV is a ground-based gauge and radar observation product that is often used in multi-product comparison studies as the baseline observed dataset (Nelson et al., 2016)*. These studies have demonstrated that NCEP Stage IV produces good representation of overall rainfall rates across the United States* (Nelson et al., 2016; Prat and Nelson, 2015)*, as well as the spatial and temporal structure of rainfall* (Habib et al., 2009)*; including for tropical cyclones* (Gao et al., 2020; Villarini et al., 2011). Prat and Nelson, (2015) *compare annual rain rate for the conterminous United States using NCEP Stage IV against gauge data, finding a correlation coefficient of 0.93 (R^2).* Gao et al., (2020) *show that NCEP Stage IV only overestimated rainfall from Hurricane Harvey by 2%. However, underestimation of extreme rainfall has been shown in some studies due to an increase in the number of missed events as rain rate increases* (Habib et al., 2009; Prat and Nelson, 2015)*. For example,* Prat and Nelson, (2015) *report that NCEP Stage IV has a tendency to underestimate rainfall in comparison to surface observations across the conterminous United States (-14%- +1% depending on location). This is likely a product of the inherent limitations of radar-based precipitation products (see* Nelson et al., (2016)*).*

The model used to produce the synthetic hurricane rainfall event set utilized in this study has previously been compared to NCEP Stage IV data over Puerto Rico, showing very good agreement (Feldmann et al., 2019)*. This demonstrates the suitability of the use of NCEP Stage IV as an observation dataset for comparison against in this study.* Omranian et al., (2018) *showed IMERG was able to represent 62% of rainfall from Hurricane Harvey in comparison to NCEP Stage IV, thus suggesting that IMERG is also likely capable of adequately representing extreme rainfall associated with Hurricane Maria. However, the performance of IMERG and NCEP Stage IV data can be dependent on the number of gauge-based observations available* (Tang et al., 2018; Tian et al., 2018)*. 14 out of 24 USGS gauges were damaged during Hurricane Maria in Puerto Rico* (Bessette-Kirton et al., 2020)*. As a result, this is a key limitation of using observed data products to estimate tropical cyclone rainfall that should be considered when drawing conclusions about the accuracy of flood hazard associated with these rainfall products."*

Secondly, we have also expanded on a previous point in Section 4.1 of the Discussion to better explain the reasons for the differences in the observed rainfall datasets, starting on Line 548, Page 20:

"Moreover, there are limitations of the observation precipitation datasets used, which propagate into the flood estimates. Many studies have compared the performance of NCEP Stage IV and IMERG rainfall data (Li et al., 2022; Mazza and Chen, 2023; Omranian et al., 2018; Villarini et al., 2011)*. Tropical cyclone precipitation in the conterminous United States between 2002-2019 was much higher in NCEP Stage IV than in satellite products such as IMERG* (Mazza and Chen, 2023)*. Other studies support this conclusion and find that the explanation for this difference is more likely an underestimation of other products, and not an overestimation bias in NCEP Stage IV itself* (Villarini et al., 2011)*. For example, IMERG is likely to underestimate orographic rainfall, which could explain why the flood extent using IMERG is lower than using NCEP Stage IV (see* ***Error! Reference source not found.****). This provides an incentive for the event set approach outlined in this study, as it allows a consideration of a wider range of plausible events to get a greater understanding of uncertainty than just the observed."*

7. In Figure 6, it is recommended that the range of the x-axis be consistent across all four results to facilitate a straightforward fair comparison. This will improve the figure readability and allow for a clearer interpretation of the data presented.

R: We have updated Figure 6 (now Figure 7) to reflect a consistent x-axis across the four subplots, starting on Line 473, Page 17.

8. Section 5 should be entitled "Conclusions" (plural).

R: This has been changed on Line 658, Page 23.

9. In the SI, please clarify what satellite imagery was used when evaluating the HWMs.

R: We have now clarified this in the SI on Page 1:

"*The satellite imagery used was* (ESRI, 2023) *(ArcGIS/World Imagery) at 0.5m resolution, as this was readily available for comparison with the HWMs in GIS applications."*

References

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate, Geosci Model Dev, 6, 687–720, https://doi.org/10.5194/GMD-6-687-2013, 2013.

Bessette-Kirton, E. K., Coe, J. A., Schulz, W. H., Cerovski-Darriau, C., and Einbund, M. M.: Mobility characteristics of debris slides and flows triggered by Hurricane Maria in Puerto Rico, Landslides, 17, 2795–2809, https://doi.org/10.1007/s10346-020-01445-z, 2020.

World Imagery - Overview: https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9#!, last access: 3 November 2023.

Feldmann, M., Emanuel, K., Zhu, L., and Lohmann, U.: Estimation of Atlantic Tropical Cyclone Rainfall Frequency in the United States, J Appl Meteorol Climatol, 58, 1853–1866, https://doi.org/10.1175/JAMC-D-19-0011.1, 2019.

Freitas, E. da S., Coelho, V. H. R., Xuan, Y., Melo, D. de C. D., Gadelha, A. N., Santos, E. A., Galvão, C. de O., Ramos Filho, G. M., Barbosa, L. R., Huffman, G. J., Petersen, W. A., and Almeida, C. das N.: The performance of the IMERG satellite-based product in identifying sub-daily rainfall events and their properties, J Hydrol (Amst), 589, 125128, https://doi.org/10.1016/J.JHYDROL.2020.125128, 2020.

Gao, S., Zhang, J., Li, D., Jiang, H., and Fang, Z. N.: Evaluation of Multiradar Multisensor and Stage IV Quantitative Precipitation Estimates during Hurricane Harvey, Nat Hazards Rev, 22, 04020057, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000435, 2020.

Habib, E., Larson, B. F., and Graschel, J.: Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in south Louisiana, J Hydrol (Amst), 373, 463–478, https://doi.org/10.1016/J.JHYDROL.2009.05.010, 2009.

Hernández Ayala, J. J., Keellings, D., Waylen, P. R., and Matyas, C. J.: Extreme floods and their relationship with tropical cyclones in Puerto Rico, Hydrological Sciences Journal, 62, 2103–2119, https://doi.org/10.1080/02626667.2017.1368521, 2017.

IPCC: Summary for Policymakers, in: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, edited by: Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T., Cambridge University Press, Cambridge, 1–24, 2018.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjansson, J. E., Medhaug, I., Sand, M., and Seierstad, I. A.: The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections, Geosci Model Dev, 6, 389–415, https://doi.org/10.5194/GMD-6-389-2013, 2013.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., and Schulz, M.: Aerosol–climate interactions in the Norwegian Earth System Model – NorESM1-M, Geosci Model Dev, 6, 207–244, https://doi.org/10.5194/GMD-6-207-2013, 2013.

Li, Z., Tang, G., Kirstetter, P., Gao, S., Li, J. L. F., Wen, Y., and Hong, Y.: Evaluation of GPM IMERG and its constellations in extreme events over the conterminous united states, J Hydrol (Amst), 606, 127357, https://doi.org/10.1016/J.JHYDROL.2021.127357, 2022.

Mazza, E. and Chen, S. S.: Tropical Cyclone Rainfall Climatology, Extremes, and Flooding Potential from Remote Sensing and Reanalysis Datasets over the Continental United States, J Hydrometeorol, 24, 1549–1562, https://doi.org/10.1175/JHM-D-22-0199.1, 2023.

Mitchell, D., Achutarao, K., Allen, M., Bethke, I., Beyerle, U., Ciavarella, A., Forster, P. M., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, N., Massey, N., Fischer, E., Schleussner, C.-F., Scinocca, J., Seland, Ø., Shiogama, H., Shuckburgh, E., Sparrow, S., Stone, D., Uhe, P., Wallom, D., Wehner, M., and Zaaboul, R.: Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design, Geosci. Model Dev, 10, 571–583, https://doi.org/10.5194/gmd-10-571-2017, 2017.

Nelson, B. R., Prat, O. P., Seo, D. J., and Habib, E.: Assessment and Implications of NCEP Stage IV Quantitative Precipitation Estimates for Product Intercomparisons, Weather Forecast, 31, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1, 2016.

Omranian, E., Sharif, H. O., and Tavakoly, A. A.: How Well Can Global Precipitation Measurement (GPM) Capture Hurricanes? Case Study: Hurricane Harvey, Remote Sensing 2018, Vol. 10, Page 1150, 10, 1150, https://doi.org/10.3390/RS10071150, 2018.

Pradhan, R. K., Markonis, Y., Vargas Godoy, M. R., Villalba-Pradas, A., Andreadis, K. M., Nikolopoulos, E. I., Papalexiou, S. M., Rahim, A., Tapiador, F. J., and Hanel, M.: Review of GPM IMERG performance: A global perspective, Remote Sens Environ, 268, 112754, https://doi.org/10.1016/J.RSE.2021.112754, 2022.

Prat, O. P. and Nelson, B. R.: Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002-2012), Hydrol. Earth Syst. Sci, 19, 2037–2056, https://doi.org/10.5194/hess-19-2037-2015, 2015.

Rios Gaona, M. F., Overeem, A., Brasjen, A. M., Meirink, J. F., Leijnse, H., and Uijlenhoet, R.: Evaluation of Rainfall Products Derived from Satellites and Microwave Links for the Netherlands, IEEE Transactions on Geoscience and Remote Sensing, 55, 6849–6859, https://doi.org/10.1109/TGRS.2017.2735439, 2017.

Rios Gaona, M. F., Villarini, G., Zhang, W., and Vecchi, G. A.: The added value of IMERG in characterizing rainfall in tropical cyclones, Atmos Res, 209, 95–102, https://doi.org/10.1016/J.ATMOSRES.2018.03.008, 2018.

Von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., Verseghy, D., Reader, M. C., Ma, X., Lazare, M., and Solheim, L.: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes, Atmosphere-Ocean, 51, 104–125, https://doi.org/10.1080/07055900.2012.755610, 2013.

Smith, J. A., Sturdevant-Rees, Paula., Baeck, M. Lynn., and Larsen, M. C.: Tropical cyclones and the flood hydrology of Puerto Rico, Water Resour Res, 41, 1–16, https://doi.org/10.1029/2004WR003530, 2005.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, J Adv Model Earth Syst, 5, 146–172, https://doi.org/10.1002/JAME.20015, 2013.

Tan, J., Petersen, W. A., Kirstetter, P. E., and Tian, Y.: Performance of IMERG as a Function of Spatiotemporal Scale, J Hydrometeorol, 18, 307, https://doi.org/10.1175/JHM-D-16-0174.1, 2017.

Tang, G., Behrangi, A., Long, D., Li, C., and Hong, Y.: Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products, J Hydrol (Amst), 559, 294–306, https://doi.org/10.1016/J.JHYDROL.2018.02.057, 2018.

Tian, F., Hou, S., Yang, L., Hu, H., and Hou, A.: How Does the Evaluation of the GPM IMERG Rainfall Product Depend on Gauge Density and Rainfall Intensity?, J Hydrometeorol, 19, 339–349, https://doi.org/10.1175/JHM-D-17-0161.1, 2018.

Villarini, G., Smith, J. A., Baeck, M. L., Marchok, T., and Vecchi, G. A.: Characterization of rainfall distribution and flooding associated with U.S. landfalling tropical cyclones: Analyses of Hurricanes

Frances, Ivan, and Jeanne (2004), Journal of Geophysical Research: Atmospheres, 116, 23116, https://doi.org/10.1029/2011JD016175, 2011.

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C. T., Paciorek, C., Gleckler, P. J., Sperber, K. R., Collins, W. D., Gettelman, A., and Jablonowski, C.: The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1, J Adv Model Earth Syst, 6, 980–997, https://doi.org/10.1002/2013MS000276, 2014.

Yu, C., Hu, D., Di, Y., and Wang, Y.: Performance evaluation of IMERG precipitation products during typhoon Lekima (2019), J Hydrol (Amst), 597, 126307, https://doi.org/10.1016/J.JHYDROL.2021.126307, 2021.