Reducing Uncertainty in Sea-level Rise Prediction: A Spatial-variability-aware Approach

Subhankar Ghosh
Dept. of Computer Science & Eng.
University of Minnesota, Twin Cities
Minneapolis, USA
ghosh117@umn.edu

Shuai An
Dept. of Computer Science & Eng.
University of Minnesota, Twin Cities
Minneapolis, USA
an000033@umn.edu

Arun Sharma
Dept. of Computer Science & Eng.
University of Minnesota, Twin Cities
Minneapolis, USA
sharm485@umn.edu

Jayant Gupta
Dept. of Computer Science & Eng.
University of Minnesota, Twin Cities
Minneapolis, USA
gupta423@umn.edu

Shashi Shekhar
Dept. of Computer Science & Eng.
University of Minnesota, Twin Cities
Minneapolis, USA
shekhar@umn.edu

Aneesh Subramanian
Dept. of Atmospheric & Oceanic Sciences
University of Colorado, Boulder
Boulder, USA
aneeshcs@colorado.edu

Abstract—Given multi-model ensemble climate projections, the goal is to accurately and reliably predict future sea-level rise while lowering the uncertainty. This problem is important because sea-level rise affects millions of people in coastal communities and beyond due to climate change's impacts on polar ice sheets and the ocean. This problem is challenging due to spatial variability and unknowns such as possible tipping points (e.g., collapse of Greenland or West Antarctic ice-shelf), climate feedback loops (e.g., clouds, permafrost thawing), future policy decisions, and human actions. Most existing climate modeling approaches use the same set of weights globally, during either regression or deep learning to combine different climate projections. Such approaches are inadequate when different regions require different weighting schemes for accurate and reliable sea-level rise predictions. This paper proposes a zonal regression model which addresses spatial variability and model inter-dependency. Experimental results show more reliable predictions using the weights learned via this approach on a regional scale.

Index Terms—climate change, sea-level rise, spatial variability, forecasting, machine learning, regression

I. INTRODUCTION

Sea-level rise is one of the most pressing environmental challenges facing the world today. In 2021 Intergovernmental Panel on Climate Change (IPCC) [26] projected that the global sea level is likely to rise by 10 to 32 inches (26 - 82 cm) by 2100, relative to 1986-2005. However, these global projections mask significant regional variation: sea level is projected to rise more in the Arctic than in the tropics and more in the mid-latitudes than in the subtropics. For example, in Chesapeake Bay (near Maryland and Washington DC), projections based on tide gauges, satellite observations and geo-physical models indicate that sea level will probably rise between 12-20 inches (30 - 50 cm) by 2050 (starting from 2005) [25]. The need for regional sea-level rise projections is becoming increasingly urgent. Coastal communities worldwide are already experiencing the impacts of sea-level rise, such as flooding, erosion, and saltwater intrusion. These impacts are expected to worsen as the sea level continues to rise. Regional sea-level rise

projections are essential for coastal planning and adaptation. They can help communities identify areas that are most at risk and develop strategies to mitigate and adapt to these impacts.

Given a multi-model ensemble of climate projections from state-of-the-art climate models of the present, the goal is to accurately and reliably predict future sea-level rise while lowering the prediction uncertainty. A spatially-variable weighting scheme assigns different sets of weights to the candidate models in different regions unlike global climate models (GCMs), which assign the same set of weights across different regions. Using a spatially variable weighting scheme ensures that higher priority is given to more reliable models in a specific region when performing predictions and reducing uncertainty.

Predictions from climate models are improved when obtained via multi-model ensembles. The result is more accurate and produces more reliable predictions with lower uncertainty. The earliest works generally used an average of the simulated predictions to obtain the final result. However, this method would not entirely capture the prediction skill of the models in different regions. More recent approaches [7], [8] use a global weighting scheme based on regression or deep learning methods to reduce the discrepancy between observations and the weighted prediction obtained from the simulations. These models assign greater weight to more reliable predictions on a global scale but do not adequately handle the spatial variability arising from regional factors which affect changes in sea level. They can approximate the thermal expansion of the oceans but cannot properly account for regional factors like tsunamis, tides, boundary waves, river runoff, vertical land movement, etc. The spatial variability arising from such regional factors is a significant challenge that can lead to different weighting schemes based on the model's reliability in that specific region.

This paper explores a zonal regression model to address spatial variability in sea-level rise prediction. This approach assigns weights to models based on their prediction skills in a
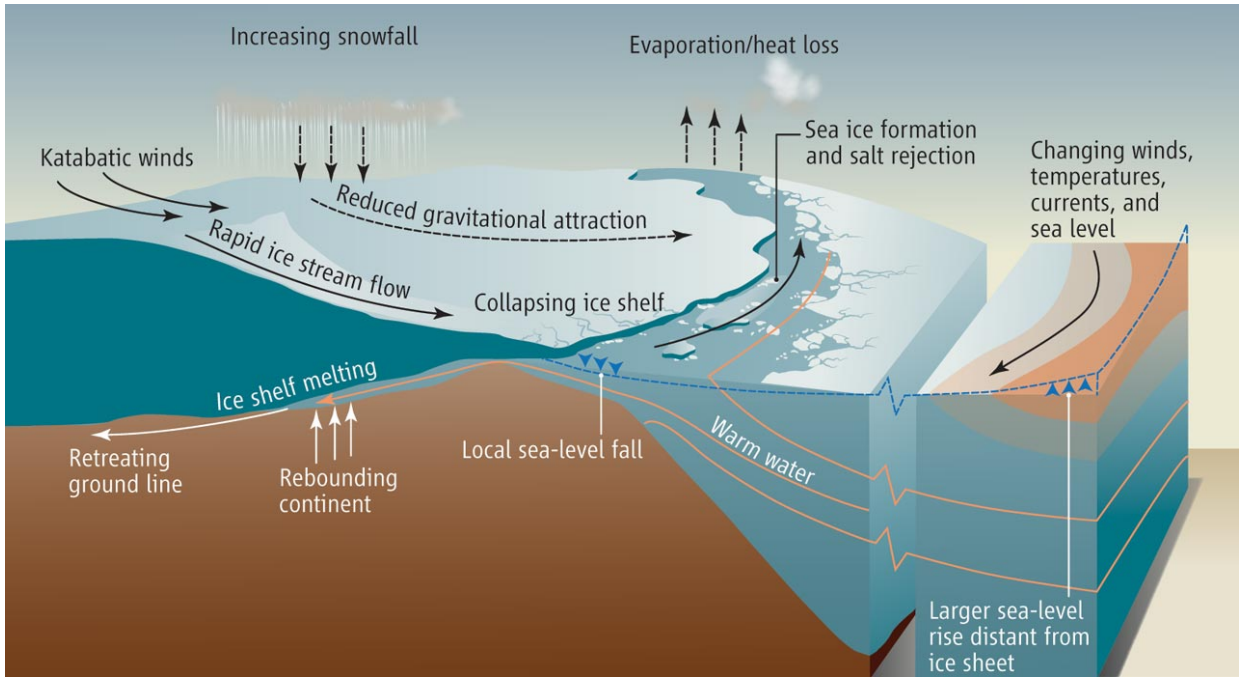
Fig. 1. Interaction between ice sheets, ocean, and atmosphere on the Antarctic and Greenland ice sheets. [1]

specific region, looking for improved predictions. Traditional weighting strategies use an inverse of the root mean square between the observations and the simulations to assign the weights on a global scale. This paper proposes a spatial-variability-aware regression model to learn the weights from historical predictions in a region. Empirical results in Figure 5 using root mean squared errors from the different model predictions, the baseline method, and the proposed approach display the effectiveness of the approach in reducing the prediction error.

II. BACKGROUND

Global sea levels can be affected by factors like warming oceans and the addition of freshwater from continental ice (Figure 1). Ocean dynamics can bring warm waters in contact with glaciers, leading to the decay of ice sheets. Changes in atmospheric temperature can increase surface temperatures over continents leading to further loss of ice sheets. This reduction in the ice mass balance can impact the sea level in different regions of the earth. Other factors such as tsunamis, tides, boundary waves, river runoff, and vertical land movement can also impact regional sea levels.

Figure 2 provides a comparison of the significant factors contributing to changes in the global mean sea level (GMSL). Different satellite systems measure different GMSL properties. For example, mass-driven changes such as ice melt and land water storage are measured by satellite altimeters, like GRACE/GRACE-FO [15]. Argos [16] measures ocean heat content by monitoring factors such as temperature and salinity changes in the ocean.

III. PROPOSED APPROACH

This paper explores a zonal regression method to represent spatial variability in weighting multi-model ensemble climate projections. CMIP-6 projections obtained from the National Center for Atmospheric Research (NCAR) are used as the simulated data, while Copernicus sea-level anomaly gridded data are used for the ground truth observation.

Given a set of spatio-temporal tensors representing climate model projections (P_r) over a region r, the proposed weighting strategy aims to find a set of coefficients (w_r) for the multi-model ensemble such that the combination of projections with weights ($\sum_i P_{r,i} w_{r,i}$) improves the accuracy of the sea-level rise prediction while simultaneously reducing the prediction uncertainty. Here, i represents the model under consideration.

We formulate the problem as a non-negative least squares problem where the goal is to find the set of weights w_r that minimizes the prediction error:

$$\underset{w_{r,i}}{\operatorname{argmin}} ||O_r - \sum_i P_{r,i} w_{r,i}||, \tag{1}$$

where $w_{r,i} >= 0$ and $\sum_i w_{r,i} = 1$. O_r represent the ground truth observation data over region r.

Simultaneously, we also need to reduce the prediction uncertainty:

$$minimize \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} (E_t - \bar{E})^2}, \tag{2}$$

where $E_t = ||O_{r,t} - \sum_i P_{r,t,i} w_{r,i}||$. This can be considered as a multi-objective optimization problem in which we want to minimize both the prediction error and the prediction
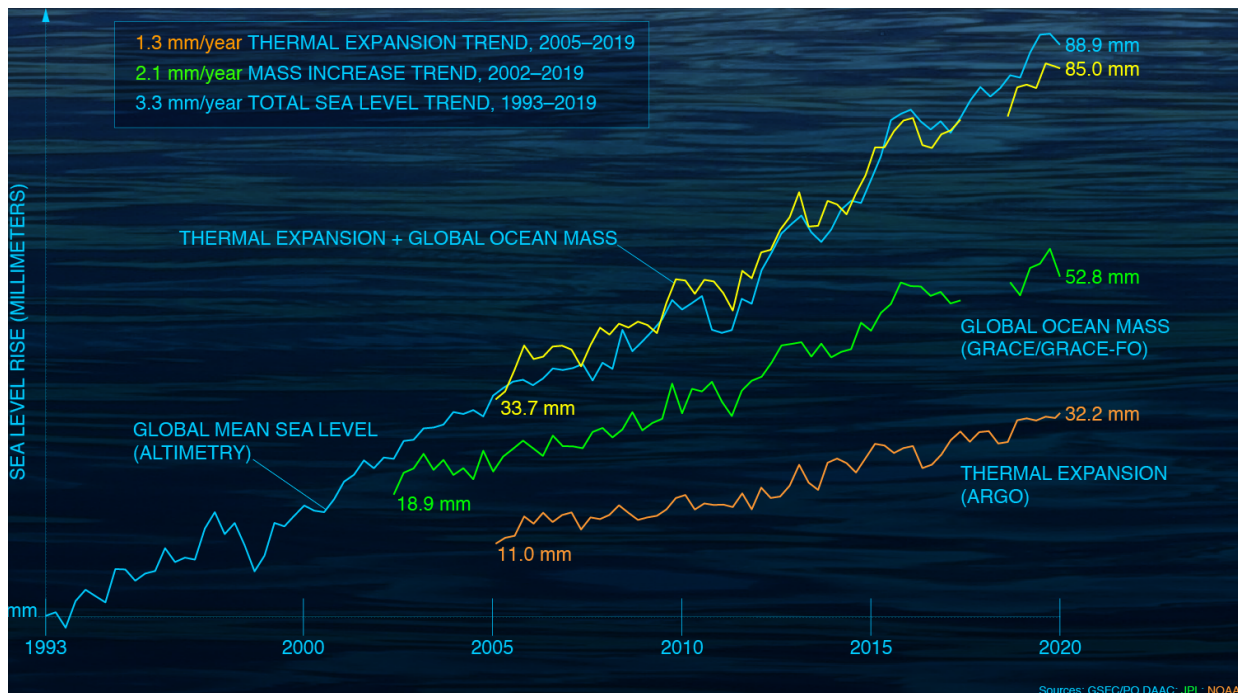
Fig. 2. Contribution to sea-level rise from thermal expansion and ocean mass. [6]

uncertainty. One technique for solving this problem is to use Pareto optimality. The main goal in this technique is to find solutions such that each quantity (prediction error and prediction uncertainty) is minimized without affecting the other. Thus we want to find the Pareto frontier (i.e. set of Pareto efficient solutions) as in Figure 3. A point $y'' \in \mathbb{R}^m$ is preferred to (dominates) another point $y' \in \mathbb{R}^m$ is referred to as $y'' \succ y'$. In Figure 3, both A and B dominate C in reducing the prediction error and the uncertainty.

Model similarity can lead to projections biased towards the largest set of similar models and the underestimation of uncertainties [23]. To address this, the proposed method down-weights models with a higher covariance in their predictions. The final weights obtained are a product of the weights obtained from the regression coefficients and the coefficients if the model is down-weighted. This zonal regression formulation overcomes the drawbacks of one-size-fits-all (OSFA) models such as GCMs which do not account for spatial variability.

IV. EXPERIMENTAL FRAMEWORK

A. Experimental Evaluation

A zonal regression model is employed to learn a weighting scheme for the simulated projections such that the output of the weighted predictions reduces the discrepancy with the observations. The main experimental goal was to compare the sea-level rise predictions obtained with the proposed spatial variability-aware zonal regression approach against the sea-level rise predictions from a one-size-fits-all method that doesn't adapt to zones. Figure 4 shows the overall validation framework. The metric for comparison with the

baseline [7], [8] was solution quality, specifically the root mean squared error between the combined prediction and the ground truth observations. The experiments were performed with the CMIP-6 projection models of sea-level height and the Copernicus satellite data for ground truth.
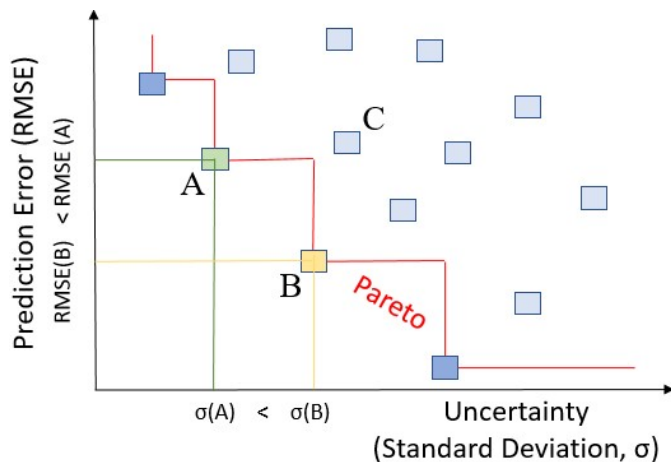

Fig. 3. Pareto frontier for simultaneous reduction of prediction error and uncertainty.

B. Data

This work primarily uses Copernicus [9] sea-level gridded satellite data for ground truth observations and zos (Sea Surface Height Above Geoid) data from CMIP-6 (Coupled Model Intercomparison Project 6) for the simulated projec-
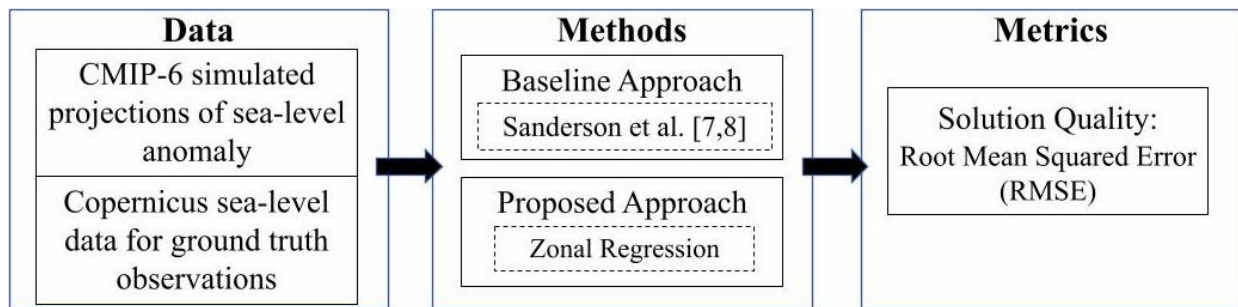
Fig. 4. Experiment Design

tions available from NCAR (National Center for Atmospheric Research).

The sea level satellite observation data [9] from Copernicus provides daily and monthly global estimates of mean sea level anomaly based on satellite altimetry measurements from 1993 to 2012. Here sea level anomaly refers to the height of the sea surface in comparison to the mean sea level at a given time and region. The dataset is generated using a two-satellite merged constellation where one satellite is used as a reference while the other is used to improve accuracy.

The Coupled Model Intercomparison Project (CMIP) [10] is a collaborative framework for collecting, organizing, and distributing output from multiple coupled climate models performing common sets of experiments. It is focused on global climate model (GCM) simulations of the past, current and future climate. The major components include (1) DECK (Diagnostic, Evaluation, and Characterization of Klima), which are some common experiments, and CMIP historical simulations (1850–present); (2) common standards, coordination, infrastructure, and documentation; (3) ensembles of CMIP-Endorsed Model Intercomparison Projects (MIPs) specific to a phase of CMIP (CMIP6) that build upon the DECK and CMIP historical simulations. Some data variables of interest available from CMIP-6 include air pressure at sea level, air temperature, surface temperature, sea-level anomaly, surface runoff flux, etc. The experiments mentioned in this paper focus on sea-level anomaly (zos) data.

Experiments mentioned in this paper were focused on the Eastern North America (ENA) region due to the availability of more reliable ground truth data in the ENA region. The experiments were performed on four CMIP-6 models obtained from NCAR, namely CESM2 [17], CESM2-FV2 [18], CESM2-WACCM [19] and CESM2-WACCM-FV2 [20].

V. PRELIMINARY RESULTS

Figure 5 provides a comparison of the root mean squared errors (RMSE) in sea-level rise predictions from four simulation models, the baseline method, and the proposed approach. The RMSE obtained from each model prediction was higher than the weighted combination of the predictions. In the ENA region, the proposed method had a median RMSE of 0.35 while the baseline method [7], [8] had a median RMSE of 0.36 across the different time scales. The standard deviation

of the prediction was 0.0177 for the proposed method while the standard deviation of the baseline method was 0.0158. The outliers in a boxplot represent datapoints that are located outside the whiskers of the box plot. For example, outliers in Figure 5 are data points 1.5 times outside the interquartile range (IQR) above the upper quartile ($Q3$) and below the lower quartile ($Q1$), i.e. ($Q1 - 1.5 * IQR$ or $Q3 + 1.5 * IQR$). This shows that accounting for spatial variability leads to more reliable predictions.

VI. RELATED WORK AND DISCUSSION

The Fourth National Climate Assessment (2018) mentions a weighting strategy used by the Intergovernmental Panel on Climate Change for differential weighting of climate projection models. The approach, based on works by Sanderson et al. [7], [8], incorporates both skills in climatological performance and the inter-dependency of models due to factors such as common parameterizations. A set of weights are obtained via a ranking scheme of the climate models based on the inverse of the root mean square error (RMSE) between the observations and the predictions. The final output is a single set of weights that can be used to find a weighted average of the climate projections. Inter-dependency between models is determined based on inter-model RMSE between each pair of models. A model is down-weighted if its inter-model RMSE with another model is significantly lower than a predetermined threshold. The final weight for each model is a product of its skill and its inter-dependency weights. This final set of weights obtained is used for combining projections irrespective of region and thus the strategy ignores the spatial variability from various regional physical forcings.

Spatial variability is receiving increasing attention in pattern detection [12], [13], [27], [28] and prediction studies [11], [29]. Recent approaches [14] towards predicting regional sea-level rise with machine learning require explicit input values for the regional physical factors, which might be difficult to obtain in certain cases. Such bottlenecks can be avoided by employing the proposed approach.

VII. CONCLUSION AND FUTURE WORK

This paper discussed a new spatial variability-aware zonal regression approach towards weighting sea-level rise models. Rather than relying on a model's global projection skill, this
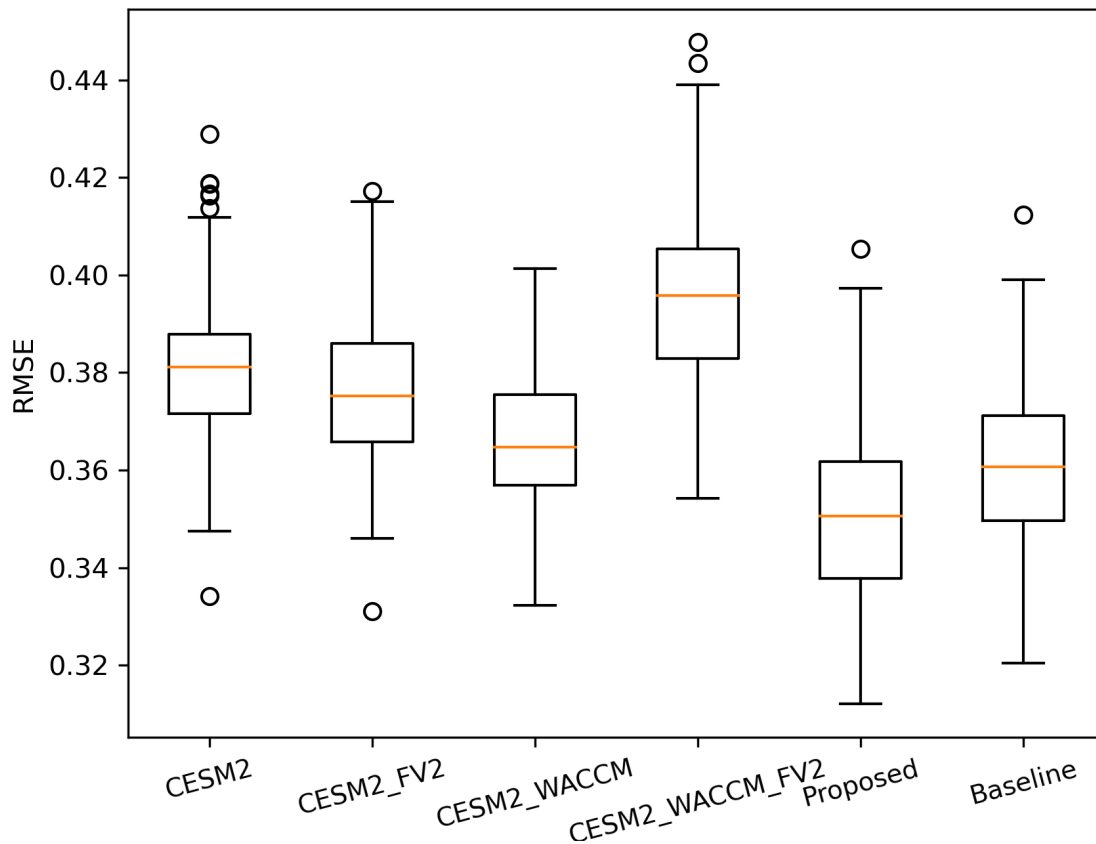
Fig. 5. Root mean squared errors of each model.

method learns a weighting scheme that accounts for regional factors in sea-level rise predictions.

Future Work: We plan to extend this work into more sophisticated approaches such as physics informed models using GWR (geographically weighted regression) or spatial-variability-aware deep learning methods [11] and work on novel ways to handle the inter-dependency between the models. A potential future direction of this work would be to make sea-level rise predictions at a higher resolution.

Sea-level rise is directly tied to the rise in greenhouse gas emissions. Therefore, we can improve the quantification of uncertainty in sea-level rise projections by considering different scenarios that are likely to affect levels of greenhouse gas emissions in the future. For example, with no human intervention, scientists project emissions will double by 2100. If all countries comply with the Paris Agreement, emissions should reach net zero after 2050. A third scenario assumes countries only meet their current national commitments for reducing emissions. In our future work, we plan to explore how such emissions scenarios can be used to reduce uncertainty in sea-level rise projections.

REFERENCES

[1] J. Willis and J. Church, "Regional sea-level projection," Science, vol. 336, pp. 550–551, 2012.

[2] Intergovernmental Panel On Climate Change, "Climate change 2007: The physical science basis," Agenda, vol. 6., no. 07, 2007, pp.333.

[3] M. Vermeer and S. Rahmstorf, "Global sea level linked to global temperature," Proceedings of the national academy of sciences, vol. 106, no. 51, National Acad Sciences, 2009, pp. 21527–21532.

[4] W. Pfeffer, J. Harper and S. O'Neel, "Kinematic constraints on glacier contributions to 21st-century sea-level rise," Science, vol. 321, no. 5894, American Association for the Advancement of Science, pp. 1340–1343, 2008.

[5] R. Nicholls et al., "Sea-level rise and its possible impacts given a 'beyond 4 C world'in the twenty-first century," Philosophical transactions of the Royal Society A: mathematical, physical and engineering sciences, vol. 369, no. 1934, The Royal Society Publishing, pp. 161–181, 2011.

[6] NASA, "Understanding Sea Level," https://sealevel.nasa.gov/understanding-sea-level/global-sea-level/overview.

[7] B. Sanderson, R. Knutti and P. Caldwell, "A representative democracy to reduce interdependency in a multimodel ensemble," Journal of Climate, vol. 28, no. 13, American Meteorological Society, pp. 5171–5194, 2015.

[8] B. Sanderson, M. Wehner and R. Knutti, "Skill and independence weighting for multi-model assessments," Geoscientific Model Development, vol. 10, no. 6, Copernicus GmbH, pp. 2379–2395, 2017.

[9] Copernicus, "Sea level gridded data from satellite observations for the global ocean from 1993 to present," https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-sea-level-global.

[10] V. Eyring et al., "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization," Geoscientific Model Development, vol. 9, no. 5, Copernicus GmbH, pp. 1937–1958, 2016.

[11] J. Gupta et al., "Spatial variability aware deep neural networks (SVANN): a general approach," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 12, no. 6, ACM New York, NY, pp. 1–21, 2021.

[12] S. Ghosh, J. Gupta, A. Sharma, S. An and S. Shekhar, "Towards geographically robust statistically significant regional colocation pattern detection," Proceedings of the 5th ACM SIGSPATIAL International Workshop on GeoSpatial Simulation, pp. 11–20, 2022.

[13] S. Ghosh, J. Gupta, A. Sharma, S. An and S. Shekhar, "Reducing False Discoveries in Statistically-Significant Regional-Colocation Mining: A Summary of Results," In 12th International Conference on Geographic Information Science (GIScience 2023)., vol. 277, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 3:1–3:18, 2023.

[14] V. Nieves, C. Radin, G. Camps-Valls, "Predicting regional coastal sea level changes with machine learning," Scientific Reports, vol. 11, no. 1, Nature Publishing Group UK London, pp. 7650, 2021.

[15] NASA Jet Propulsion Lab, "Measuring Earth's Surface Mass and Water Changes," https://gracefo.jpl.nasa.gov.

[16] Argos (satellite system), "https://www.argos-system.org".

[17] Community Earth System Model 2 (CESM2), "https://www.cesm.ucar.edu/models/cesm2".

[18] Community Earth System Model version 2 - FV2, "https://www.cen.uni-hamburg.de/en/icdc/data/cryosphere/cmip6-sea-ice-area/cesm2-fv2.html".

[19] Whole Atmosphere Community Climate Model, "https://catalogue.ceda.ac.uk/uuid/108b11504ef44f4fb0bb66273e261709".

[20] CESM2-WACCM-FV2, "https://www.wdc-climate.de/ui/cmip6?input=CMIP6.CMIP.NCAR.CESM2-WACCM-FV2.historical".

[21] NASA-GISS-E2, "https://data.giss.nasa.gov/modelE/cmip6/".

[22] AWI Climate Model (AWI-CM), "https://www.awi.de/en/science/climate-sciences/atmospheric-physics/climate-modelling.html".

[23] R. Pathak, H. Dasari, K. Ashok, I Hoteit, "Effects of Observational Uncertainty and Models Similarity on Climate Change Projections," Research Square Platform LLC, 2023.

[24] M. Iturbide, "An update of IPCC climate reference regions for sub-continental analysis of climate model data: definition and aggregated datasets," Earth System Science Data, vol. 12, no. 4, Copernicus GmbH, pp. 2959–2970, 2020.

[25] D. Boesch et al., "Sea-level Rise Projections for Maryland 2023," UMCES reports on Sea-Level Rise Projections for Maryland, 2023.

[26] J. Lee et al., "Future global climate: scenario-based projections and near-term information," Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, Cambridge University Press, pp. 553–672, 2021.

[27] A. Sharma, J. Gupta and S. Ghosh, "Towards a tighter bound on possible-rendezvous areas: preliminary results," In Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp. 1-11, 2022.

[28] M. Farhadloo et al., "SAMCNet: towards a spatially explainable AI approach for classifying MxIF oncology data," Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining, pp. 2860–2870, 2022.

[29] Y. Li et al., "Eco-PiNN: A Physics-informed Neural Network for Eco-toll Estimation," Proceedings of the 2023 SIAM International Conference on Data Mining (SDM), pp. 838-846, 202.