Resilient Anomaly Detection in Ocean Drifters with Unsupervised Learning, Deep Learning Models, and Energy-Efficient Recovery

Claire Guo[1,3], Jiachi (Sophia) Zhao[2,3]

Lynbrook High School, 1280 Johnson Ave, San Jose, CA 95129[1]; Los Altos High School, 201 Almond Ave, Los Altos, CA 94022[2]; Boston University, Boston, MA 02215[3]

Introduction

- Ocean currents play a vital role in climate regulation, marine ecosystems, and pollutant transport

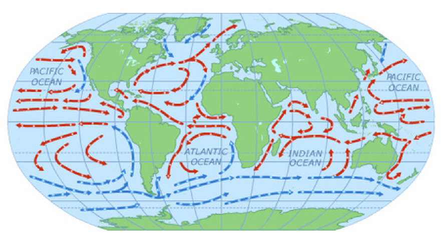

Figure 1. Global ocean currents

- Previous research has used ocean drifter data to model large-scale current patterns, but few have addressed the detection of localized anomalies in real time
- Studies have shown that changes in drifter trajectories can reflect external disturbances like cyclones, eddies, or oil spills
- However, most existing work focuses on trajectory forecasting, not anomaly detection without labels
- We hypothesize that analyzing drifter paths over time can reveal sensor errors and environmental anomalies, even when disruptions occur far from the monitored region

Methods

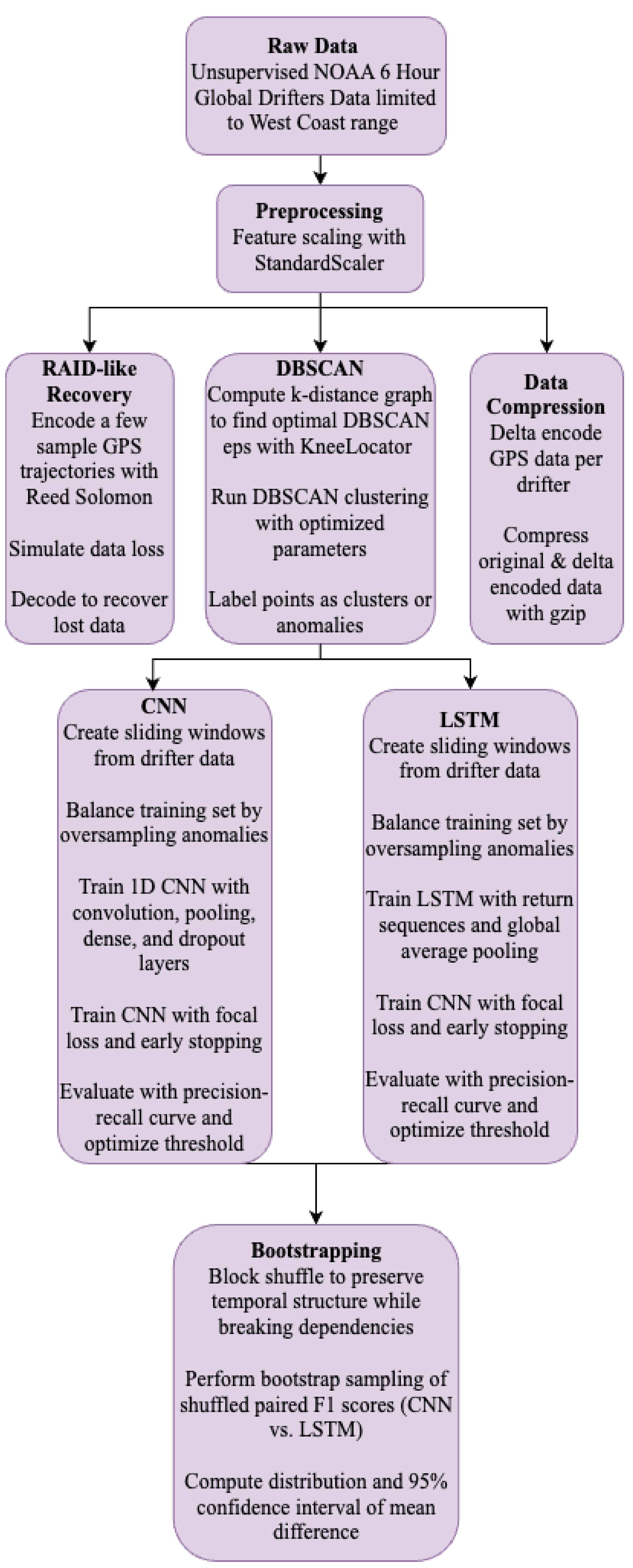

Figure 2. Flowchart of entire methodology

Results

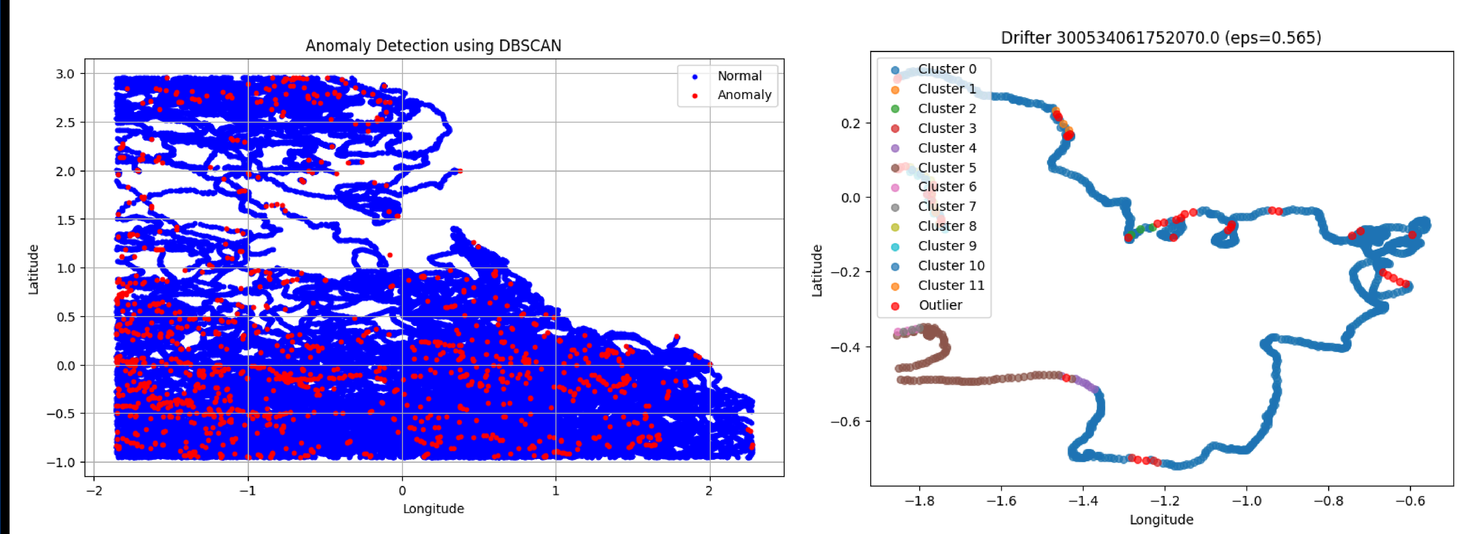

Figure 3. DBSCAN-based anomaly detection on ocean drifter data. Left: Overall results show anamoly points (red) separated from normal trajectories (blue) across all drifters. Right: Example of a single drifter.

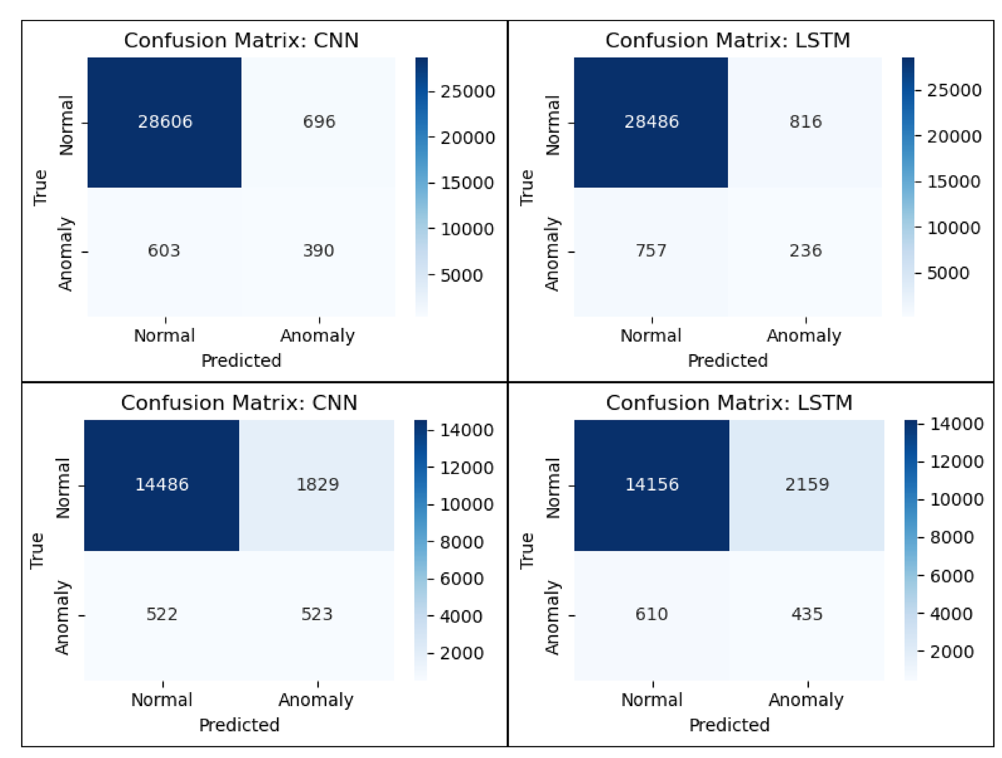

Figure 4. Confusion matrices for CNN and LSTM models. Top: Results on 2023 data; Bottom: results on 2010 data. Each plot shows true vs. predicted labels for normal vs. anomaly classifications. The 2010 confusion matrices show an increase in true positive anomaly predictions. CNN outperforms across all entries of the confusion matrix.

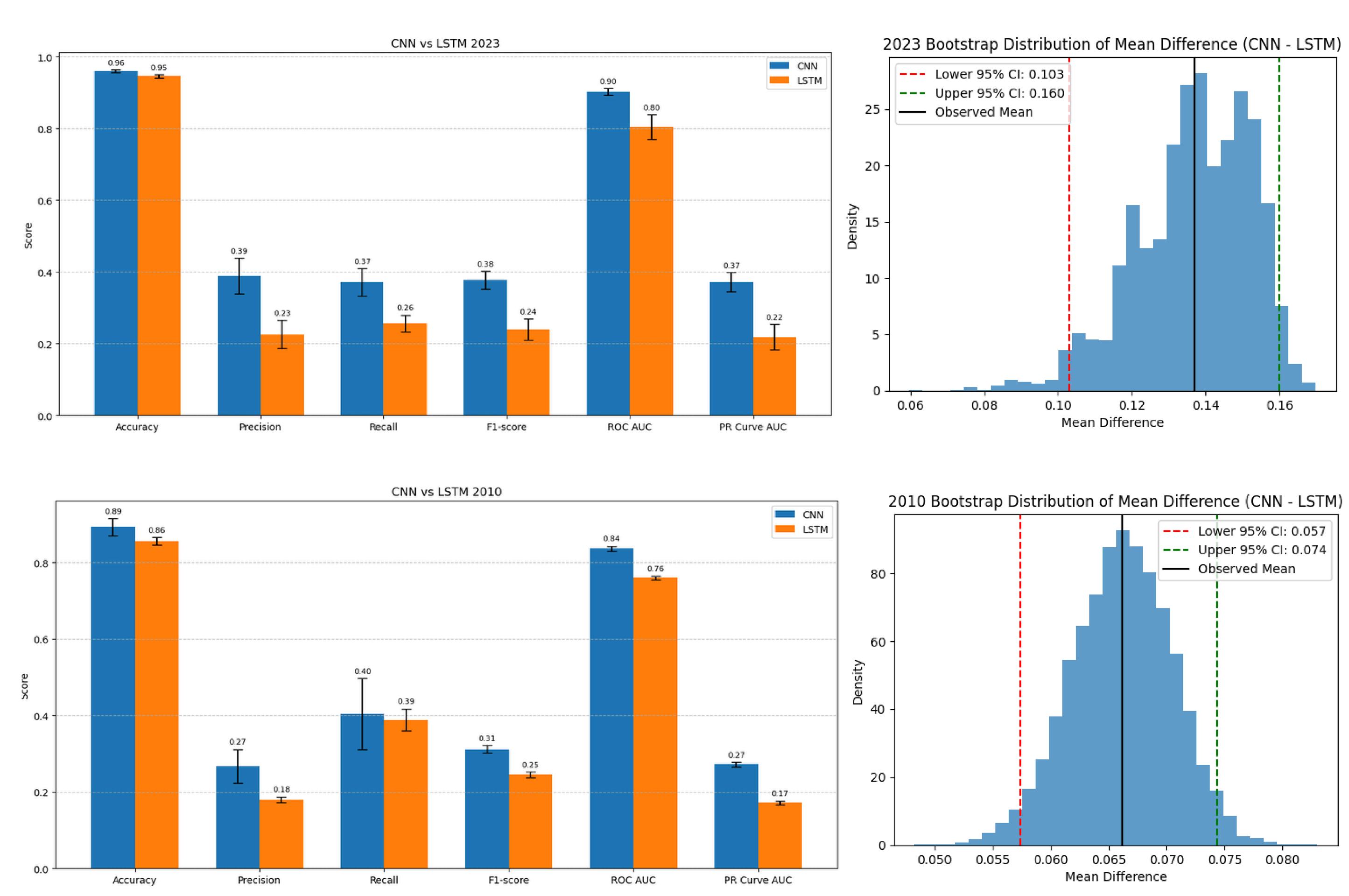

Figure 5. Bar graphs and bootstrap distributions for CNN and LSTM models. Top: Results on 2023 data; Bottom: results on 2010 data. CNN demonstrates superior average performance across 10 runs, as shown in the summary bar charts. The 95% confidence interval for the difference in F1 scores (CNN – LSTM) excludes 0, indicating a statistically significant advantage for the CNN model.

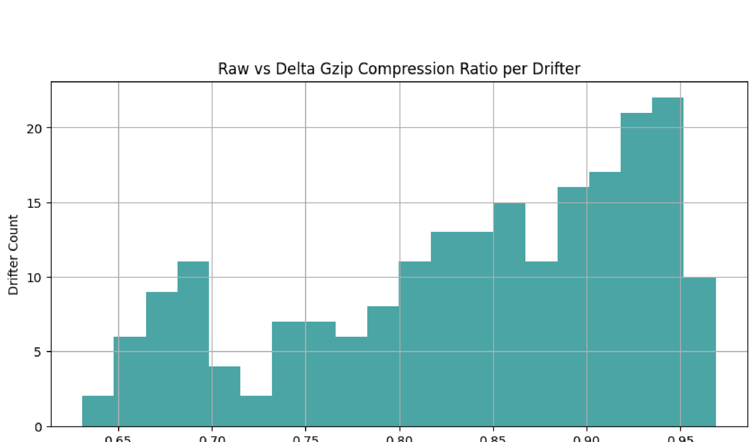

Figure 6. Histogram of gzip compression ratios (raw to delta-encoded) per drifter, averaging 0.84, showing delta encoding reduces compressed size by ~16%, enabling estimated energy savings of 27% during transmission.

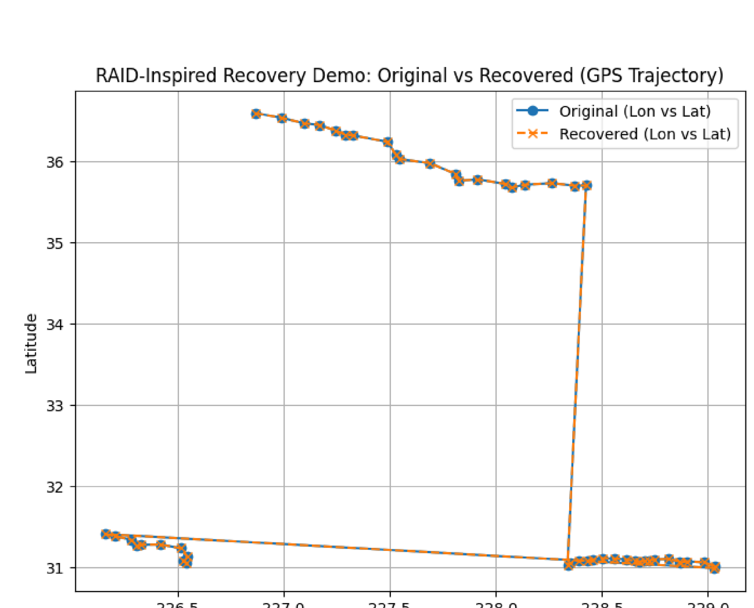

Figure 7. Original (circles) and recovered (crosses, dashed) GPS trajectories for 5 drifters over 20 steps. Data encoded with Reed-Solomon and 5 dropout rows corrupted. Recovery yields zero reconstruction MSE.

Discussion

Conclusions:

- Deep learning models trained on DBSCAN-labeled windows capture temporal anomaly patterns.
- CNN consistently outperforms LSTM on F1 score. The 95% confidence interval excludes zero, indicating a statistically significant advantage for CNN.
- RAID recovery and compression methods improve data robustness and energy efficiency
- Increased anomalies during Deepwater Horizon event suggest detection of environmental disruptions far from the spill zone

Limitations:

- No ground truth due to labels from unsupervised DBSCAN
- Limited scope: Study focuses on U.S. West Coast
- RAID/compression assumptions: Assume structured loss and smooth trajectories

Future Work:

- Expand detection to global ocean datasets
- Add environmental inputs like temperature, salinity, and satellite imagery
- Explore the performance of BiLSTM compared to current models of CNN and LSTM

References

Beron-Vera, F. J.; Olascoaga, M. J.; Goni, G. J. Surface Ocean Mixing Inferred from Different Multisatellite Altimetry Measurements. J. Phys. Oceanogr. 2010, 40 (11), 2466–2480. https://doi.org/10.1175/2010JPO4458.1.

Beron-Vera, F. J.; Olascoaga, M. J.; Haller, G.; Farazmand, M.; Trinanes, J.; Wang, Y. Dissipative Inertial Transport Patterns near Coherent Lagrangian Eddies in the Ocean. Chaos 2015, 25 (8), 087412. https://doi.org/10.1063/1.4928693.

Rollenbeck, R.; Bendix, J.; Fabian, P.; Boy, J.; Wilcke, W.; Dalitz, H.; Oesker, M.; Emck, P. Comparison of Different Techniques for the Measurement of Precipitation in Tropical Montane Rain Forest Regions. J. Atmos. Oceanic Technol. 2007, 24 (2), 156–168. https://doi.org/10.1175/JTECH1970.1.

Vieira, G. S.; Rypina, I. I.; Allshouse, M. R. Uncertainty Quantification of Trajectory Clustering Applied to Ocean Ensemble Forecasts. arXiv 2020, arXiv:2008.12253. https://doi.org/10.48550/arXiv.2008.12253.

Zika, J. D.; McDougall, T. J.; Sloyan, B. M. A Tracer-Contour Inverse Method for Estimating Ocean Circulation and Mixing. J. Phys. Oceanogr. 2010, 40 (1), 26–47. https://doi.org/10.1175/2009JPO4208.1.

Acknowledgements

We are deeply grateful to everyone who supported our project and contributed to the incredible experience in RISE. We especially thank Eugene Pinsky for his invaluable guidance in shaping our study. We also appreciate Patrick Bloniaz and other teaching fellows for their lectures and help. Finally, we extend heartfelt thanks to our families for their encouragement and support throughout this journey.