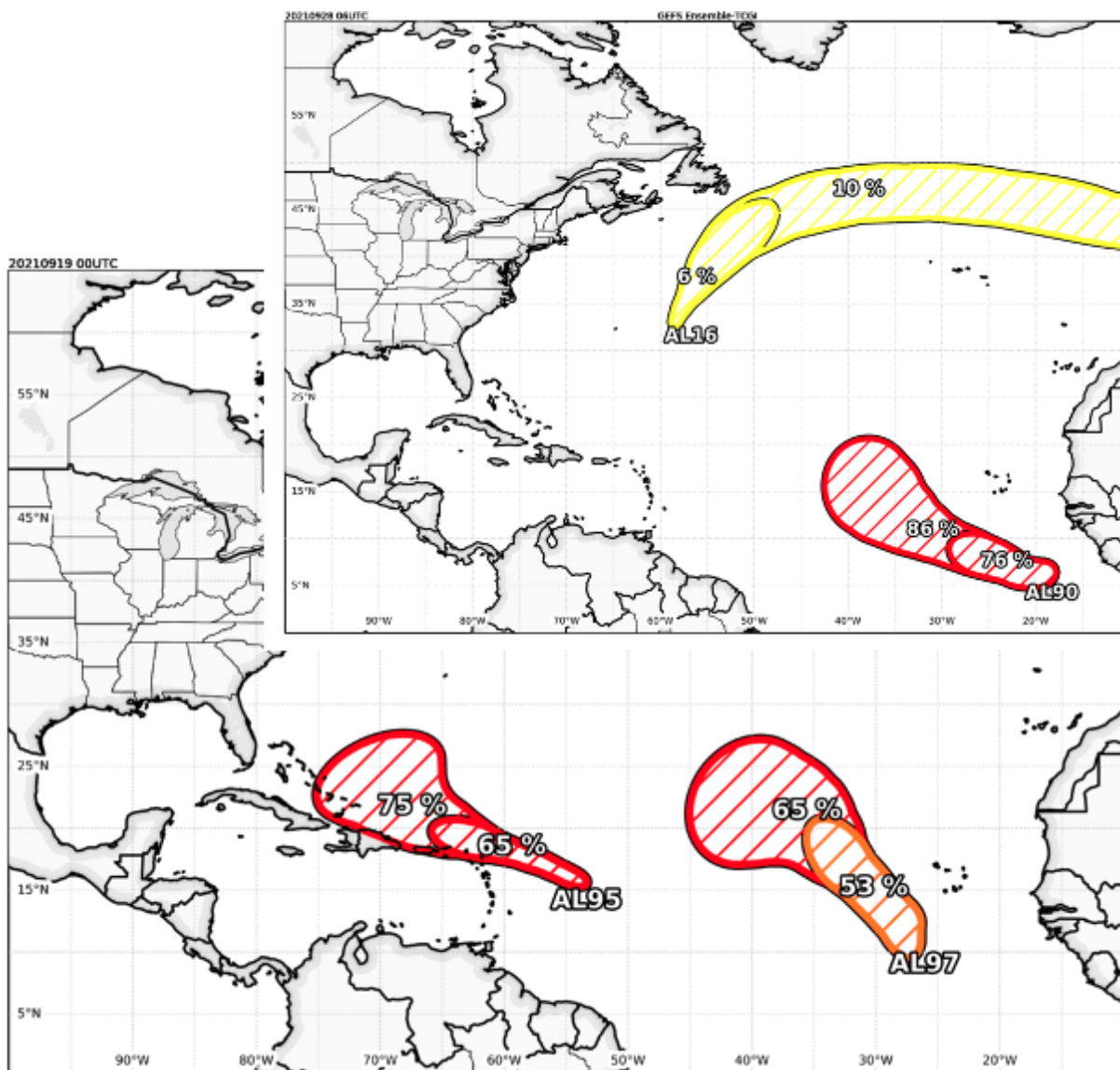


Machine Learning and Predicting Tropical Cyclogenesis

Tropical cyclogenesis is the birth of a tropical cyclone. Tropical cyclogenesis is a rare event—at a given location on Earth at a specific time, the chance of tropical cyclogenesis is less than 0.1%. Thus, tropical cyclogenesis can be difficult to forecast; however, early predictions of tropical cyclone formation can help with disaster preparation and emergency management. We are developing a machine learning model that builds upon [existing](#) statistical tropical cyclogenesis forecasting products that make predictions using linear discriminant analysis (LDA). Preliminary results suggest that the machine learning model has the same hit rate as the currently used LDA algorithm while substantially reducing the false positives. We hope to continue improving this product and to extend the tropical cyclogenesis forecasts from 2 days to 5 or 7 days.



Because identifying tropical cyclone genesis remains difficult, we are also working on developing an [Ensemble Global Genesis Index](#). For this ML-aspect of this work, we are utilizing ensemble model output combined with machine learning models to improve cyclogenesis predictions at 2 and 5 days. We have tested a suite of different ML algorithms, ranging from linear regression to random forests to deep neural networks. Preliminary work suggests that reliability and skill can be gained from ML techniques (such as k-nearest neighbor or random forest) compared to the legacy models that rely on logistic regression or threshold definitions. We are currently in the process of making this product operational, and example real-time displays are shown.

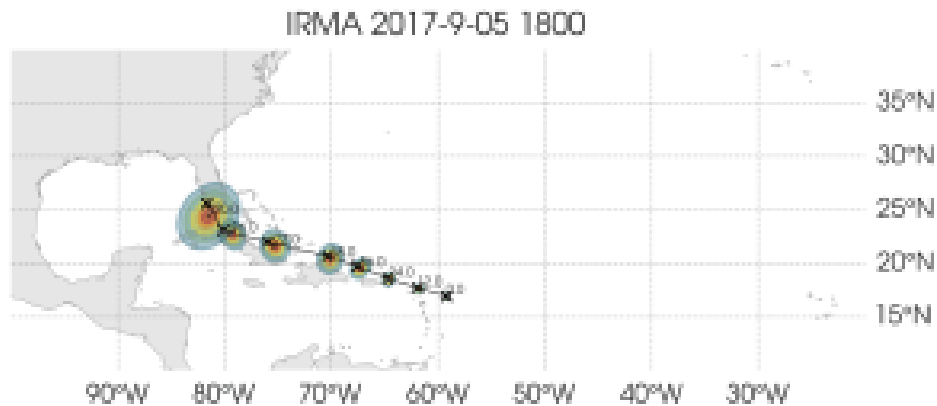
Machine Learning and TC Rapid Intensification

Rapid intensification refers to a process in which a tropical cyclone's intensity increases substantially in a short period of time. Rapid intensification is difficult to forecast as it is a rare event with complex physical processes. Rapid intensification can present serious threats to life and property, especially when it occurs shortly before a storm makes landfall. Currently, CIRA produces statistical forecasts of rapid intensification such as the Rapid Intensification Prediction Aid ([RIPA](#)). Currently, we are developing machine learning models to add to the rapid intensification forecasting tools with the goal of improving rapid intensification forecasts. Early results indicate that the addition of a machine learning model improves the existing consensus forecast. We plan to run real-time machine learning forecasts for summer 2023.

Machine Learning and Tropical Cyclone Forecast Uncertainty

Operational tropical cyclone forecasts have included uncertainty information for the past 40 years to aid decision makers in mitigation activities. The original products implemented in 1983 only provided track forecast uncertainty, but were replaced by wind speed probabilities that included the uncertainty in track, intensity and wind structure forecasts. Although these products have been very successful, the forecast uncertainty is still primarily based on climatological error distributions. Current dynamical model ensemble forecast products don't always capture the full range of forecast uncertainty (particularly for intensity forecasts), and often have limited sample sizes, and operational statistical forecast products are limited to predicting the average expected error of a consensus forecast from linear regression. To address these limitations, CIRA and collaborators at CSU Atmospheric Science have developed the **Tropical Cyclone Artificial Neural-network Error (TCANE) model**, a neural network model that can predict error distributions for tropical cyclone intensity and track forecasts in the Atlantic and East/Central North Pacific basins. The input to TCANE includes output from several deterministic track and intensity forecast models and measures of the tropical cyclone environment such as SST and vertical wind shear. The TCANE model's error predictions depend on the current forecast situation and are well-correlated with the National Hurricane Center forecast errors. Because TCANE estimates the full error distributions of track and intensity, it can be used to estimate the probability of rapid intensification. The rapid intensification forecasts

compare favorably with those issued by the NHC. Going forward, a real-time demo is planned for the 2023 Hurricane and Ocean Testbed (HOT) at the NHC, where we will develop prototype graphical products and explore using explainable AI techniques. Below is an example of a prototype product to depict the track uncertainty for a forecast for Hurricane Irma in September of 2017. The predicted errors were fairly small through 96 h, but the uncertainty increased by 120 h as Irma was forecast to turn to the north as it interacted with an upper-level trough approaching from the west.



Physics-Informed Machine Learning for Tropical Cyclones

Machine learning models can make accurate predictions of weather and climate; however, these data-driven models can also produce physically inconsistent results. Physics-guided machine learning methods seek to enforce analytic constraints, derived from physical knowledge, in these data-driven models. A new collaboration with colleagues at the University of Lausanne is focused on developing a physically-scaled model for tropical cyclone intensity prediction. We are [exploring](#) the use of potential intensity theory to scale a tropical cyclone's intensity by its maximum potential intensity. Our hope is that a scaled tropical cyclone intensity prediction model will produce more robust predictions of tropical cyclone intensity that are generalizable across basins, seasons, and from year to year.

Machine Learning for Tropical Cyclone Structure

Timely information regarding the convective structure of tropical cyclones is a critical part of forecasting, and one source of information that forecasters rely on is passive microwave imagery (PMI) data at 22-, 37-, and 89-GHz. PMI can penetrate through clouds to reveal the internal storm structure. In this imagery, the land appears warm relative to water surfaces and low-level moist air masses act to warm brightness temperatures over water surfaces, while deep convection appears relatively cold due to scattering from ice.

Since PMI provides information regarding the storm structure, it can be used for several hurricane prediction tasks. The first is in the identification of the circulation center, which is a critical step in initiating tropical cyclone advisories. The second is finding the position of the

storm in difficult situations, such as during the early development stages and at night. Third, knowing information about the convective structure can provide information regarding the convective rain bands, which can provide intensification information. And finally, it can be used to monitor structural changes, such as eyewall formation and eye replacement cycles. However, one drawback to PMI is that it is obtained from low Earth orbiting satellites, such as AMSR2. Unfortunately, this data comes at the cost of low spatial and temporal resolution, meaning that forecasters could miss important structural changes.

In our work, we are using machine learning to generate synthetic PMI from geostationary satellites, which provide high resolution spatial and temporal coverage. Starting with input data from geostationary satellite imagery (including visible, near-IR, and IR data), we feed this data into different ML algorithms to create synthetic imagery. For training data, we are using the Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset ([TC PRIMED](#)), which is a comprehensive 19-year dataset of global TC satellite PMI and coincident geostationary satellite visible and IR imagery (Razin et al., 2022a, 2022b). Currently, we are using random forests (RF), fully-connected neural networks (ANN), and convolutional neural networks (CNN) to create synthetic 89-GHz microwave imagery.

Publications

Razin, M.N., C.J. Slocum, J.A. Knaff, P.J. Brown, and M.M. Bell, 2022a: Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED), v1.0, NOAA National Centers for Environmental Information, <https://doi.org/10.25921/dmy1-0595>.

Razin, M. N., C. J. Slocum, J. A. Knaff, P. J. Brown, and M. M. Bell, 2022: Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). *Bull. Amer. Met. Soc.*, **Early Online Release**. <https://doi.org/10.1175/BAMS-D-21-0052.1>.