15C.2 AUGMENTING THE TROPICAL CYCLONE LOGISTICAL GUIDANCE FOR GENESIS (TCLOGG) BY INTRODUCING A MOST LIKELY TIME OF GENESIS

Ryan M. Remondelli[1] *, Daniel J. Halperin[2], Robert E. Hart[1]
[1]Florida State University, Tallahassee, FL
[2]Embry-Riddle Aeronautical University, Daytona Beach, FL

1. INTRODUCTION

Recent decades have seen significant advancements in operational Numerical Weather Prediction (NWP). While NWP models in the late 1990s and early 2000s have had limited ability to offer beneficial TC genesis guidance to forecasters (Beven 1999; Schumacher et al. 2009), improvements in NWP have led to new genesis tools and guidance. These products include the direct use of NWP output, via post-processed forecast guidance, greatly improving genesis prediction capabilities. Innovations such as the Genesis Potential Parameter (GPP; Kotal and Bhattacharya 2022), the Tropical Cyclone Genesis Index (TCGI; Brammer et al. 2022; 2023), and the Tropical Cyclone Logistic Guidance for Genesis (TCLOGG) have leveraged environmental indicators and NWP outputs for more accurate forecasts. Additionally, the integration of advanced machine learning techniques has further enhanced forecast accuracy and operational application potential.

TCLOGG (Halperin et al. 2017; 2020) has undergone significant development over the last decade, employing logistic regression to offer genesis probabilities across all global basins. However, TCLOGG and similar tools fall short in providing guidance for when TC genesis is likely to occur within the typical 3-day, 5-day, and 7-day TC genesis forecast periods. This lack of information is problematic for forecasters, as TC genesis, intensification, and landfall can all occur within a single Tropical Weather Outlook (TWO) period (e.g., 3 days). Developing a Most Likely Time of Genesis (MLTG) product for TCLOGG will help in filling a critical gap in information for forecasters.

Operational Global NWP Mean Absolute and Mean Genesis Timing Errors (hours) 2015-2021			
North Atlantic	*MAE*	*Mean*	*95% CI*
GFS	26.87	11.36	-8.48 ‣ 14.22
ECM	29.29	9.40	7.05 ‣ 11.74
CMC	31.87	12.94	9.41 ‣ 16.47
NAV	30.35	7.81	2.33 ‣ 8.50
UKM	28.16	-2.71	-4.91 ‣ -0.51
East Pacific	*MAE*	*Mean*	*95% CI*
GFS	22.23	4.28	2.28 ‣ 6.28
ECM	28.25	21.46	19.21 ‣ 23.71
CMC	26.39	5.41	2.33 ‣ 8.50
NAV	23.07	13.99	11.87 ‣ 16.11
UKM	22.20	5.89	3.73 ‣ 8.04

Table 1: The North Atlantic and East Pacific model specified MAE, mean error, and mean error 95% CIs for the 2015-2021 period. The 95% CIs are shaded by confidence, with intervals less than 5 hours in green, between 6 and 10 hours in yellow, and above 10 hours in red.

2. GENESIS TIMING BIAS CONFIRMATION AND DATASET DEVELOPMENT

This study aims to refine TCLOGG by incorporating a MLTG component to improve upon the biases in genesis timing predictions. An in-depth exploratory analysis of TCLOGG outputs from 2015 to 2021 identified biases in forecasted genesis times across various NWP models, including GFS, ECMWF, CMC, UKMET, and NAVGEM.

To accurately assess genesis timing and location errors, the study used the International Best Track Archive for Climate Stewardship (IBTrACS) for case verification. Non-hit cases were excluded from the dataset as a genesis verification time is required. Preliminary findings showed significant variability in timing errors across models and years, with some models displaying Mean Absolute Errors (MAE) exceeding 24 hours. Given the context of three or seven-day TWO forecasts, these MAEs are subjectively considered high. However, Table 1 also reveals 95% confidence intervals (CI) centered around zero.

These confidence intervals are narrow, indicating that the mean errors are closely clustered around zero. Considering the uncertainty in genesis time within NWP grids can be up to five hours, the CIs may suggest that in basins with good agreement, producing a single MLTG value could be challenging.

Using these identified cases from TCLOGG, a developmental dataset for the MLTG project was generated by retrieving known predictors crucial for TC genesis from each NWP model for all global basins. The MLTG developmental dataset variables were generated based on a storm-centered 5°x5° grid. Table 2 details the variables extracted from NWP model fields considered for potential inclusion in the MLTG developmental dataset, as identified in Halperin et al. (2017).

MLTG Developmental Dataset Variables Calculated from Storm Centered 5°x5° Grid	
Forecast Hour	Mean Sea Level Pressure
Latitude	Longitude
925 hPa Wind Speed Maximum	200 hPa - 850 hPa Average Vertical Wind Shear
250 hPa-850 hPa Thickness Maximum	250 hPa-850 hPa Thickness Perturbation from Average
850 hPa Relative Vorticity Maximum	850 hPa Relative Vorticity Perturbation from Average
700 hPa Average RH	700 hPa RH Perturbation from Average
850 hPa Average Divergence	

Table 2: The variables used from NWP model fields for potential inclusion in the MLTG developmental dataset. These variables were shown to be relevant predictors for TC genesis in H17.

3. RANDOM FOREST REGRESSOR MODEL DEVELOPMENT

The majority of the MLTG-based guidance utilized the Random Forest Regressor (RFR; Brieman 2001), a sophisticated ensemble-based machine learning algorithm known for handling non-linear relationships, as the estimator of choice.

The RFR models utilized cases from 2015-2020 (80% of the developmental dataset) for model training. First, forward predictor selection with cross-validation, using relevant predictors (Table 2), was completed. This step ensured that only the most important, non-redundant predictors were included, minimizing the risk of overfitting and maximizing skill.

Hyperparameter tuning occurred thereafter. This involved the optimization of several parameters important to the performance of the RFR. These include the number of trees in the forest (n_estimators) and the maximum depth of the trees (max_depth). This optimization process sought to find a balance between model complexity and prediction accuracy, ensuring the model was sufficiently detailed to capture the underlying data patterns without becoming overly complex.

Once the best features and parameters were found, the RFR model was fitted on the training dataset. The remaining 20% served as a validation dataset used to assess the model's accuracy in predicting the actual forecast hour of genesis.

Corresponding author address: Ryan M. Remondelli, Florida State University, Department of Earth, Ocean, and Atmospheric Sciences, Tallahassee FL, 32308; email: rremondelli@fsu.edu.

The results from the RFR-based MLTG methodology exhibited reduced MAE and root mean squared errors (RMSE) across various models and basins on the validation dataset. This result was shown particularly within the JTWC's area of responsibility. This improvement demonstrates the value of the models and allowed for the experimental implementation for the 2023 season.

4. CONSENSUS XGBOOST REGRESSOR DEVELOPMENT

Consistent with the original TCLOGG methodology, a consensus MLTG (CMLTG) needed to be created for all basins. However, a more advanced machine learning algorithm needed to be used. This is because for some cases, predictors are not available, and must be filled with NaN values.

As a result, the XGBoost Regressor is used. XGBoost, or Extreme Gradient Boosting, is a machine learning algorithm operating on the principles of ensemble based learning and gradient boosting (Chen and Guestrin 2016). XGBoost sequentially builds an ensemble of shallow decision trees to iteratively correct errors and improve overall predictive performance. XGBoost also effectively prevents overfitting by incorporating L1 and L2 regularization.

Further, XGBoost optimizes the interchange between a loss function and regularization to fine-tune model parameters through gradient descent (Chen and Guestrin 2016). A key benefit for developing the consensus MLTG, XGBoost Regressor incorporates a mechanism for handling missing data. During the training process, XGBoost can internally handle missing values in the input features. Rather than requiring explicit imputation, XGBoost intelligently learns the optimal input strategy for missing data as part of its optimization process. It achieves this by considering the missing values separately and creating a split in the decision trees that directs instances with missing values to one of the branches. This not only simplifies the preprocessing steps for users, but also contributes to the algorithm's overall resilience and effectiveness in scenarios where data incompleteness is prevalent (Chen and Guestrin 2016).

Utilizing a compiled dataset of timing forecasts for the 2016-2020 period (AL, EP), and 2015-2019 period (JTWC), a XGBoost Regressor model was trained utilizing the raw TCLOGG and RFR MLTG timing forecasts as input predictors.

An exhaustive grid search to hyperparameter tune the XGBoost model was applied on the developmental dataset. Subsequently, the model was trained on a distinct 80% partition of the developmental dataset, split by year, to prevent data leakage into the validation portion of the dataset.

The resulting CMLTG predictive skill, when tested on the validation dataset, is on par with the remainder of the RFR MLTG models. Accordingly, the updated CMLTG was experimentally implemented alongside the other MLTG guidance for the 2023 season to test overall skill.

5. 2023 OPERATIONAL RESULTS

With the verification of the experimental MLTG guidance for the 2023 season, it is important to note that this verification is based on a single season and utilizes a completely independent dataset from that used in the development of the RFR and XGB models. This is crucial, as some basins, notably the Indian Ocean (IO), have very low case counts, making them potentially unrepresentative of overall MLTG performance.

For the Atlantic (AL; Figure 1) Basin, the results indicate a minor level of improvement in models such as NAV, CMC, and GFS, suggesting some effectiveness of the MLTG in this region. Comparatively, minimal degradations are observed in the UKM, ECM, CON-XGB and CON-Mean models. However, overall improvements or degradations within the basin are slight.

Corresponding author address: Ryan M. Remondelli, Florida State University, Department of Earth, Ocean, and Atmospheric Sciences, Tallahassee FL, 32308; email: rremondelli@fsu.edu.

In the Eastern Pacific (EP; Figure 1) Basin, timing improvements are notable in the NAV, ECM, GFS, and UKM models, but a degradation in CMC points to the MLTG's model-specific effectiveness within this basin. Interestingly, this basin is the only instance where the CON-XGB underperforms the CON-MEAN.

The Western Pacific (WP; Figure 1) Basin presents another mixed scenario, where

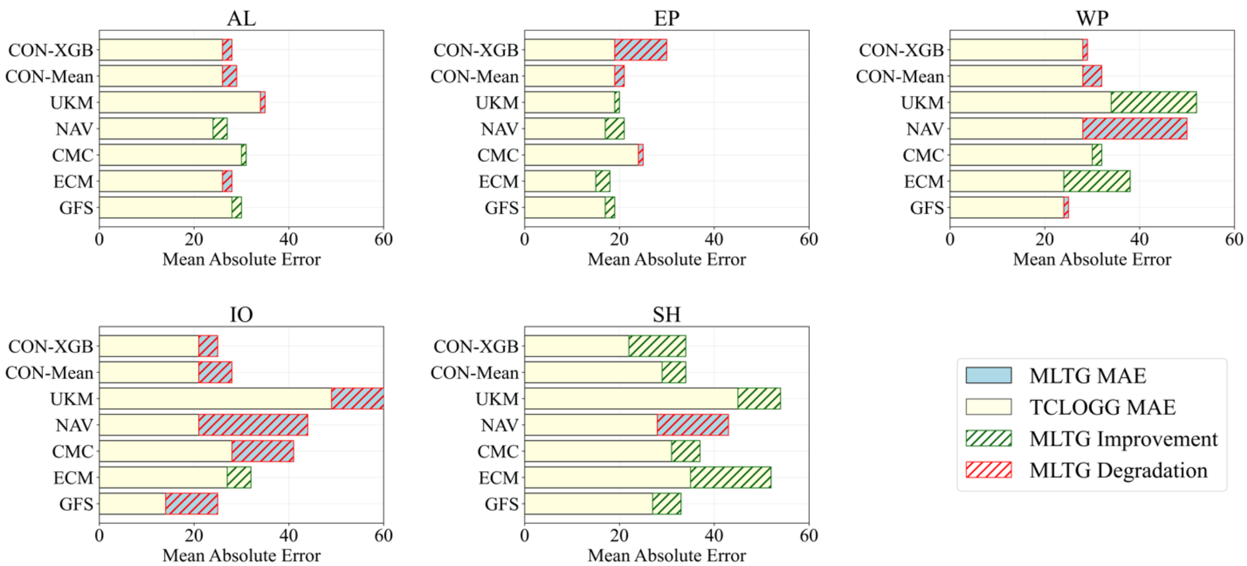

Figure 1: 2023 RFR and raw TCLOGG timing MAE for all deterministic guidance, a simple MLTG averaged consensus guidance (CON-Mean), and XGBoost Regressor consensus guidance (CON-XGB). The raw TCLOGG timing MAE is shown in yellow, while the RFR and XGB-based guidance MAE is blue. Degradations or improvements of the MAE of the machine learning guidance is shown in hatched green (improvements) or red (degradations).

significant improvements are seen in the UKM and the ECM models, but a considerable drop in performance is noted for the NAV. Compared to the CON-MEAN, CON-XGB, and GFS models, the differences between MLTG and raw timing are minimal, suggesting a nuanced impact of MLTG guidance across models.

In the Indian Ocean (IO; Figure 1) Basin, nearly all models show forecast degradation, except for ECM, which is an outlier. The CON models display small degradations, while the UKM, CMC, and NAV models are significantly degraded. This outcome is particularly concerning given the model's positive performance in the validation dataset, hinting at possible overfitting issues or the environmental conditions during the season significantly diverged from those in the developmental dataset. The smaller training dataset for the IO Basin compared to the WP and Southern Hemisphere (SH) basins may also affect the overall skill of the models. However, given the low case count for 2023 in the IO Basin, further validation in 2024 is necessary before considering adjustments to the model.

For the Southern Hemisphere (SH; Figure 1) Basin, the general trend leans towards MLTG improvements across all models, except for the NAV, which shows timing degradation. This basin stands out as the most successful in terms of overall MLTG improvement over the raw TCLOGG tracker, highlighting the potential benefits and challenges of implementing MLTG guidance across different basins.

6. SUMMARY

The development of the MLTG guidance was pursued to provide more robust and precise timing guidance for forecasters. This project aimed to improve TCLOGG, a widely used tool for effective TC genesis forecasting.

Corresponding author address: Ryan M. Remondelli, Florida State University, Department of Earth, Ocean, and Atmospheric Sciences, Tallahassee FL, 32308; email: rremondelli@fsu.edu.

Overall, the improvements offered by the MLTG guidance have been mixed. The majority of timing enhancements were observed within the JTWC Area of Responsibility, where initial timing errors tended to be larger compared to those in the NHC. This discrepancy may stem from the different methodologies employed by the NHC and JTWC in determining genesis, with the NHC's approach being closer to the model guidance used in developing the MLTG.

Nonetheless, the results prompt critical examination of whether generating a single genesis time yields the most robust guidance. This question is particularly notable given the initial CIs which are predominantly centered around zero. The relatively small CI sizes call into question the usefulness of relying on a singular value method. This skepticism is further supported by the inherent uncertainty in precise genesis timing within NWP models, resulting from the 6-hour temporal resolution of the grids. Such rounding could lead to genesis forecasts being off by approximately 5 hours, potentially explaining much of the observed CI, especially within the NHC AOR.

Therefore, it appears that resolving genesis timing with a single forecast hour (FHR) might not adequately capture the inherent uncertainties. An interval-based approach to predicting genesis timing could offer a more accurate and skillful tool for forecasters by better representing forecast uncertainty. The next generation of timing guidance should strongly consider adopting such an interval-based method to enhance the reliability and utility of tropical cyclone genesis timing predictions.

7. REFERENCES

Beven, J. L., 1999: The boguscane—A serious problem with the NCEP medium range forecast model in the Tropics. Preprints, 23rd Conf. on Hurricanes and Tropical Meteorology, Dallas, TX, American Meteorological Society, 845–848.

Brammer, A., Grogan, D., Dunion, J., 2022: Tropical Cyclone Genesis index and Global Ensemble Forecasts: Expanding to Global Basins and Evaluation of Real-Time 2021 Forecasts. Presented at American Meteorological Society Annual Conference, Houston, Texas, January 2022.

Brammer, A., Grogan, D., Dunion, J., 2023: Evaluation of a new Hybrid Global Ensemble Statistical Genesis Index for Tropical Cyclones. Presented at American Meteorological Society Annual Conference, Denver, Colorado, January 11th, 2023.

Breimen, H., 2001: Machine-mediated learning, Volume 45, 5-32.

Chen, T., and Guestrin, C., 2016: XGBoost: A scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2016. 785-794.
Weather and Forecasting, 28, 1423-1445.

Halperin, D., R. Hart, H. Fuelberg, and J. Cossuth, 2017: The development and evaluation of a statistical-dynamical tropical cyclone genesis guidance tool. Weather and Forecasting, 32, 27-46.

Halperin, D.J., A. B. Penny, and R. Hart, 2020: A comparison of tropical cyclone genesis forecast verification from three Global Forecast System (GFS) operational configurations. Weather and Forecasting, 35, 5.

Kotal, S. D., and Bhattacharya, B., 2022: Tropical cyclone Genesis Potential Parameter (GPP) and it’s application over the north Indian Sea. Environmental Science.

Schumacher, A.B., DeMaria, M. and Knaff, J.A., 2009. Objective estimation of the 24-h probability of tropical cyclone formation. Weather and Forecasting, 24(2), pp.456-471.

Corresponding author address: Ryan M. Remondelli, Florida State University, Department of Earth, Ocean, and Atmospheric Sciences, Tallahassee FL, 32308; email: rremondelli@fsu.edu.