
Machine Learning Approaches on Identifying Tropical Waves That Develop into Hurricanes

Haochang Luo

Department of Earth and Atmospheric Sciences
City College of New York
New York, NY, 10031
hluo@ccny.cuny.edu

Spencer A. Hill

Department of Earth and Atmospheric Sciences
City College of New York
New York, NY, 10031

Abstract

African Easterly Waves (AEWs) are synoptic-scale atmospheric disturbances that serve as precursors to tropical cyclones (TCs) in the North Atlantic and North Africa. As climate changes, TC activities are increasingly frequent, leading to exponentially growing socio-economic losses. So understanding the physical mechanisms governing the tropical cyclogenesis (TCG) of AEWs remains a crucial problem. Competing theoretical frameworks, including baroclinic instability, barotropic instability, and moisture-vortex instability (MVI) have been proposed, but their relative importance and temporal evolution during storm development remain unclear. In this study, machine learning algorithms are used to empirically analyze the governing mechanisms of AEW development based on 40 years of reanalysis data (1979-2018). We develop a computer vision framework utilizing convolutional neural networks (CNNs) and transformer architectures to identify developing AEWs (DAEWs) from non-developing AEWs (NDAEWs) based on wave-centered composites of key thermodynamic and dynamic variables for storm development. The model results suggest that the MVI framework is a critical factor for high classification accuracy in distinguishing developers from non-developers.

1 Introduction and motivation

African Easterly Waves (AEWs) are westward-propagating atmospheric disturbances that originate from instabilities along the African Easterly Jet over the Sahel region of North Africa [13]. These synoptic-scale waves represent one of the most important seedlings for tropical cyclone development in the North Atlantic basin, with approximately 85% of major hurricanes (Categories 3-5) having AEW origins [1]. Despite their critical importance for seasonal hurricane forecasting and climate prediction, the fundamental processes governing which AEWs develop into tropical cyclones remain poorly understood [11, 1].

The transformation of an AEW into a tropical cyclone through tropical cyclogenesis (TCG) is a complex, multi-scale process involving interactions between atmospheric dynamics and thermodynamics. Only approximately 3% of AEWs successfully undergo TCG, making this a highly imbalanced classification problem from both meteorological and machine learning perspectives. The rarity of development, combined with the chaotic nature of the atmosphere and inherent predictability limits

in atmospheric systems, has made it extremely challenging to establish robust statistical relationships between environmental conditions and development outcomes.

1.1 Theoretical Frameworks

Several competing theoretical frameworks have been proposed to explain AEW development and critical to TCG:

The baroclinic instability depicts the conversion of potential energy associated with horizontal temperature gradients into kinetic energy of the disturbance [11]. The energy source for wave amplification is the meridional temperature gradients created by land-sea contrasts [11].

The barotropic instability states that the horizontal wind shear, particularly the strong easterly jet structure over West Africa is the energy source [11, 13]. The curvature and shear of the African Easterly Jet provide the necessary conditions for barotropic wave growth, with studies showing this process dominates AEW initiation and early development phases [12].

The moisture-vortex instability (MVI) was recently proposed by Adames and Ming (2018)[14]. This framework emphasizes the role of latent heat release and moisture advection in the vortex growth [14]. The theory suggests that the meridional advection of moisture toward the wave center creates a positive feedback loop: increased moisture leads to enhanced latent heat release, which strengthens the vortex through diabatic heating and vortex stretching. Recent observational studies have provided empirical support for this mechanism in distinguishing developing from non-developing AEWs [7].

While each mechanism has theoretical support and has been observed in case studies, their relative importance and temporal evolution during the AEW lifecycle remain unclear. Traditional statistical approaches have struggled to disentangle these competing effects due to the nonlinear interactions between variables.

1.2 Machine Learning Approach

Recent advances in machine learning, particularly in computer vision and explainable AI, offer new opportunities to address this long-standing problem. By treating atmospheric fields as images and applying deep learning techniques, we can potentially derive complex nonlinear relationships that are neglected in traditional statistical models [9]. Recent studies have demonstrated the effectiveness of machine learning approaches for tropical cyclogenesis prediction, with neural networks and ensemble methods showing improved skill over traditional statistical approaches [9].

In our study, we feed computer vision models with wave-centered horizontal composites that represent the critical growth mechanisms of AEWs. Those physically meaningful variables derived from established theoretical frameworks will provide us with insights into the underlying physics.

The primary goals of this study are to: 1. Develop a robust machine learning framework for categorizing developing vs. non-developing AEWs. 2. Identify and examine which theoretical mechanisms are most predictive of development at different lifecycle stages. This paper will focus on the work in the first stage.

2 Data and Preprocessing

2.1 Datasets

Two major datasets are used in this study: storm track dataset and the fifth reanalysis product from the European Centre for Medium-Range Weather Forecasts reanalysis data (ERA5, [5]).

The AEW tracking dataset from 1979-2018 was derived using an alternative algorithm of Brammer and Thorncroft (2015) [2, 4]. This dataset identifies AEWs by tracking vorticity maxima over North Africa and the North Atlantic basin, providing 6-hourly central positions (latitude/longitude) and intensity metrics for each disturbance. The dataset includes lifecycle information, documenting transitions from tropical waves to tropical depressions (TD), tropical storms (TS), and hurricanes when applicable. The total dataset contains approximately 6,570 individual AEW cases, with roughly 3% classified as developers (DAEWs) that eventually undergo tropical cyclogenesis and 97% non-developers (NDAEWs).

ERA5 provides global atmospheric conditions at $0.25^\circ \times 0.25^\circ$ horizontal resolution with 6-hourly temporal resolution, matching the storm tracking dataset. The high spatial and temporal resolution of ERA5 is essential for capturing the mesoscale features associated with AEW development [8].

2.2 Physical Variables

Based on the theoretical frameworks outlined above, the following key variables that capture the essential physics of AEW development are selected as input channels: 1. The column-integrated moist static energy (MSE) is a thermodynamically conserved quantity that combines the effects of temperature, moisture, and geopotential height:

$$MSE = C_p T + L_v q + \Phi \quad (1)$$

where $C_p = 1004 \text{ J} \cdot (\text{K} \cdot \text{kg})^{-1}$ is the air heat capacity at constant pressure; T is temperature; $L_v = 2.5 \times 10^6 \text{ J} \cdot \text{kg}^{-1}$ is the specific latent heat through phase changing; q is specific humidity and Φ is geopotential. The column-integrated MSE is calculated as:

$$\langle MSE \rangle = \frac{1}{g} \int_{p_t}^{p_s} MSE \, dp \quad (2)$$

where $g = 9.8 \text{ m} \cdot \text{s}^{-2}$ is the gravitational acceleration; p_s is the surface pressure, p_t is the pressure at the top of the atmosphere.

2. Following the MVI framework, we calculate the meridional component of MSE advection: $\langle v \frac{\partial MSE}{\partial y} \rangle$ where v is the meridional wind component and y is the meridional coordinate. This variable captures the moisture transport toward the wave center, which is hypothesized to be critical for development.

3. Potential vorticity (PV) combines rotation and stratification effects and is conserved following fluid motion in the absence of diabatic heating and friction. The 850 hPa level captures the low-level circulation associated with AEWs while avoiding boundary layer complexities.

Other variables, such as barotropic conversion, will be incorporated and compared with MVI in the next stages of this work.

2.3 Wave-centered composites

To create a standardized comparison framework, we construct wave-centered composites that follow each AEW, throughout its lifecycle. For each 6-hourly time step, we extract $16^\circ \times 16^\circ$ snapshots of each variable, centered on the tracked wave position.

Temporal composites are created by averaging snapshots from the beginning of each wave's lifecycle. For non-developers (NDAEWs), they are averaged over the entire observed lifecycle. For developers (DAEWs), they are averaged from birth to the moment of TCG.

2.4 Data Preprocessing

The dataset is split to avoid data leakage: 60% of the cases are used for training models, 20% are used for validation and 20% are used for testing. All variables are standardized by subtracting the mean and dividing by the standard deviation of the training set.

Since the data is extremely unbalanced, the Synthetic Minority Oversampling Technique (SMOTE) was implemented as it generates synthetic examples of the minority class (DAEWs) by interpolating between existing minority class instances in feature space [3].

3 Model Architecture and Configuration

3.1 Customized Architecture

Our primary model is based on the Convolutional Neural Network (CNN) architecture. The model treats the wave-centered composites as multi-channel images, with each physical variable representing a separate channel. The model accepts inputs of shape $3 \times 64 \times 64$, representing the $16^\circ \times 16^\circ$ spatial

domain with three channels corresponding to column-integrated MSE, meridional MSE advection, and 850 hPa potential vorticity. The architecture employs multiple convolutional blocks, each containing two 2D convolutional layers, batch normalization, ReLU activation function, max-pooling layers and dropout layers. The width of convolutional layers increases with depth from 32 to 64 and 128. An adaptive average pooling layer is applied after the three blocks. Fully connected layers are added to filter the results into a single probability. A final dense layer with sigmoid activation turns the probability to a binary classification output.

3.2 Pre-defined Architectures

Other than our customized CNN model, we compared the results from re-defined models including: ResNet(resnet18), wide ResNet (wide_resnet50_2), ResNext(resnext50_32x4d), EfficientNet(efficientnet_b0) and VGG(vgg16). Their heads are modified so they can be connected to the input channels.

3.3 Hyperparameter Optimization

Model hyperparameters are optimized using Optuna, a Bayesian optimization framework. The algorithm will automatically search better hyperparameters based on a Tree-structured Parzen Estimator (TPE). The optimization process tunes: learning rate (1×10^{-5} to 1×10^{-2}), weight decay rate (1×10^{-6} to 1×10^{-2}), batch size (16, 32, 64), optimizer (Adam, SGD, AdamW), loss function (Binary Cross Entropy (BCE), Focal loss), and dropout rate (0.1-0.7). The random seed is 42. The optimization seeks the highest F1-score on the validation set. Each of the models goes through 100 trials with an automatic median pruner. Each trial runs 50 epochs.

4 Preliminary Results and Ongoing Work

The models are tuned to retain the best F1-score of DAEWs in the validation set. The results are shown in Table. 1:

model name	learning rate	batch size	optimizer	loss function	weight decay rate	F1-score
Customized CNN	0.292×10^{-4}	64	AdamW	BCE	0.903×10^{-3}	0.700
ResNet	0.250×10^{-3}	16	SGD	BCE	0.696×10^{-5}	0.643
Wide ResNet	0.320×10^{-4}	32	Adam	Focal	0.123×10^{-5}	0.585
ResNeXt	0.279×10^{-3}	64	AdamW	Focal	0.757×10^{-2}	0.628
VGG	0.102×10^{-4}	16	Adam	Focal	0.319×10^{-4}	0.625
EfficientNet	0.620×10^{-2}	32	Adam	Weighted BCE	0.354×10^{-2}	0.587

Table 1: Hyperparameters and the best F1-score

Among all architectures, the customized CNN model shows a relatively good F1-score compared to the other pre-defined models. However, its performance exhibits substantial variability across the five cross-validation folds. This variance is primarily due to the extreme class imbalance between developing and non-developing AEWs, even after applying SMOTE oversampling. Following previous work [7], we mitigate this issue by further splitting the non-developing AEWs based on their lifetime-maximum 850 hPa potential vorticity. Their intensity distribution density is estimated using a Gaussian kernel. The 95th percentile of the distribution is adopted as the threshold ($\sim 4.1 \times 10^{-5} \text{ s}^{-1}$) for strong and weak cases. The analysis proceeds with the 320 strong non-developing AEWs. Because the developing AEW precision tends to be low in trials, we use the F_β -score with $\beta = 0.75$, an inclination to precision, as the metric. We also include a Vision Transformer (ViT) for comparison.

The performance for all architectures is listed in Table. 2. Hyperparameters are optimized using Optuna for each architecture. Wide ResNet and EfficientNet achieve the best F_β -scores, with EfficientNet yielding the highest precision and the customized CNN obtaining the highest recall. Overall, the F_β -scores range between 0.69 to 0.77.

Using the customized CNN as a baseline, we evaluate different combinations of the three predictor variables (Table. 3), optimizing hyperparameters separately for each case. The full three-variable combination yields the highest F_β -score, and performance generally increases with the number of input channels. This suggests that for a given architecture, incorporating additional physically relevant features improves predictive skill.

Model	F_β -score	Precision	Recall
Customized CNN	0.69	0.64	0.80
ResNet	0.70	0.68	0.72
Wide ResNet	0.77	0.82	0.69
ResNeXt	0.75	0.81	0.67
VGG	0.75	0.83	0.64
EfficientNet	0.77	0.92	0.59
ViT	0.71	0.71	0.64

Table 2: Architecture comparison

Combination	F_β -score	Precision	Recall
MSE_adv+MSE+PV	0.69	0.64	0.80
MSE_adv	0.58	0.54	0.65
MSE	0.66	0.71	0.60
PV	0.65	0.65	0.65
MSE_adv+MSE	0.64	0.67	0.60
MSE_adv+PV	0.64	0.67	0.60
MSE+PV	0.66	0.63	0.75

Table 3: Combinations of variables based on customized CNN

To evaluate model behavior under a fixed hyperparameter configuration, we use the AutoML framework AutoGluon [10]. The experiments employ the default multimodal configuration: a Swin Transformer (swin_base_patch4_window7_224; [6]) pretrained on ImageNet-1k at 224×224 resolution, AdamW optimization, ROC-AUC validation metric, layer-wise cosine-decayed learning rate initialized at 0.0001, weight decay of 0.001, and macro F1-score as evaluation metric. Table 4 shows that the combination of MSE advection and column MSE yields the highest scores across all four metrics. Unlike the CNN experiments, these results highlight that when architecture and hyperparameters are fixed, adding more variables does not necessarily improve performance; instead, predictors must be selected and weighted according to their physical relevance.

Combination	Strong non-developing F1	developing F1	macro avg F1	weighted avg F1
MSE_adv+MSE+PV	0.77	0.53	0.65	0.68
MSE_adv	0.76	0.33	0.55	0.60
MSE	0.77	0.00	0.38	0.48
PV	0.62	0.47	0.54	0.56
MSE_adv+MSE	0.82	0.54	0.68	0.71
MSE_adv+PV	0.77	0.49	0.63	0.66
MSE+PV	0.77	0.55	0.66	0.69

Table 4: Combinations of variables using Autogluon default setup

In the ongoing and future work, we will enable the hyperparameter optimization pipeline within AutoGluon and explore additional model architectures with larger parameter counts. Alternative resampling methods will also be evaluated to address the extreme class imbalance in AEW development. Additional dynamical and thermodynamical variables, such as barotropic conversion terms, will be incorporated to systematically assess competing physical theories of AEW development. Given the limited number of samples, we plan to employ data augmentation methods and potentially generate additional storm cases using numerical weather prediction models. We will apply layer-wise relevance propagation (LRP) to understand which physical structures contribute most strongly to development predictions. The current wave-centered composites neglect potentially important temporal information. In future work we will analyze the full time evolution of AEWs to assess when moist vortex instability-like conditions emerge and how these temporal signals influence predictability.

References

- [1] Emily Bercos-Hickey and Christina M. Patricola. African Easterly Wave Strength and Observed Atlantic Tropical Cyclone Genesis and Characteristics. *Journal of Geophysical Research: Atmospheres*, 129(10):e2024JD040858, 2024. [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JD040858](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JD040858).

- [2] Alan Brammer and Chris D. Thorncroft. Variability and evolution of african easterly wave structures and their relationship with tropical cyclogenesis over the eastern atlantic. *Monthly Weather Review*, 143(12):4975–4995, December 2015.
- [3] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, June 2002.
- [4] Yuan-Ming Cheng. Variability of African easterly waves. *Legacy Theses & Dissertations (2009 - 2024)*, January 2019.
- [5] Hans Hersbach, Bill Bell, Paul Berrisford, András Horányi, Joaquín Muñoz Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Morten Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.
- [6] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. *CoRR*, abs/2103.14030, 2021.
- [7] Víctor C. Mayta, Ángel F. Adames Corraliza, Kayleen Torres Maldonado, Haochang Luo, and Kelly M. Núñez Ocasio. Thermodynamic Processes Governing the Evolution of Developing and Strong Non-developing African Easterly Waves. April 2025. Section: Journal of the Atmospheric Sciences.
- [8] Kelly M. Núñez Ocasio, Zhuo Wang, and George H. Bryan. African easterly wave evolution and tropical cyclogenesis in a pre-helene (2006) hindcast using the model for prediction across scales-atmosphere (mpas-a). *Journal of Advances in Modeling Earth Systems*, 15(3):e2022MS003181, 2023.
- [9] Yurong Qian, Kai Zhao, Jianping Li, and Zhiwei Zhu. Reduced tropical cyclone genesis in the future as predicted by a machine learning model. *Earth’s Future*, 10(2):e2021EF002455, 2022.
- [10] Zhiqiang Tang, Haoyang Fang, Su Zhou, Taojiannan Yang, Zihan Zhong, Tony Hu, Katrin Kirchhoff, and George Karypis. Autogluon-multimodal (automm): Supercharging multimodal automl with foundation models. *arXiv preprint arXiv:2404.16233*, 2024.
- [11] Chris D. Thorncroft and Brian J. Hoskins. An idealized study of african easterly waves. I: A linear view. *Quarterly Journal of the Royal Meteorological Society*, 120(518):953–982, 1994.
- [12] Joseph D. White and Anantha Aiyyer. African easterly waves in an idealized general circulation model: instability and wave packet diagnostics. *Weather and Climate Dynamics*, 2(2):311–329, 2021.
- [13] M.-L. C. Wu, O. Reale, S. D. Schubert, M. J. Suarez, and C. D. Thorncroft. African easterly jet: Barotropic instability, waves, and cyclogenesis. *Journal of Climate*, 25(5):1489–1510, 2012.
- [14] Ángel F. Adames and Yi Ming. Interactions between water vapor and potential vorticity in synoptic-scale monsoonal disturbances: Moisture vortex instability. *Journal of the Atmospheric Sciences*, 75(6):2083–2106, March 2018.