Identifying and Predicting the Lagrangian coherence of eddies in the Gulf of Mexico using machine learning and satellite observations

Luna Hiron[1*], Olmo Zavala-Romero[1,2], Eric P. Chassignet[1], Philippe Miron[1], and Bulusu Subrahmanyam[3]

[1]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida, United States
[2]Department of Scientific Computing, Florida State University, Tallahassee, Florida, United States
[3]School of the Earth, Ocean, and Environment, University of South Carolina, Columbia, South Carolina, United States

Key Points:

- Machine learning can successfully identify and predict the Lagrangian coherence of eddies in the Gulf of Mexico
- The machine learning model achieved accuracy rates of 90% for identification and 93% for lifetime prediction of Loop Current Eddies
- Incorporating chlorophyll data enhances the machine learning model's ability to predict the Lagrangian coherence of eddies.

*L. H. and O. Z.-R. contributed equally to the generation of the results.

Corresponding author: Luna Hiron, `lhiron@fsu.edu`.

Abstract

Lagrangian coherent eddies efficiently transport water properties, such as heat and salt, as well as tracers, including oil, larvae, and Sargassum, throughout the ocean. For instance, during the 2010 Deepwater Horizon oil spill, part of the oil was captured within a Loop Current Frontal Eddy (LCFE), preventing it from reaching the Florida Keys. Similarly, Loop Current Eddies (LCEs) carry warmer, saltier waters typical of the Caribbean Sea for the western Gulf of Mexico (GoM). In this study, we employ machine learning alongside various satellite observations—absolute dynamic topography (ADT), sea surface temperature (SST), and chlorophyll-a (Chl-a)—to identify Lagrangian coherent eddies in the GoM and predict their lifetime. Three durations of Lagrangian coherence are investigated: 5, 10, and 20 days. This study also investigates the contributions of Chl-a in identifying and forecasting LCEs and LCFEs' Lagrangian coherence, aiming to assess the advantages of integrating this dataset into data-assimilative Gulf ocean models, in addition to ADT and SST. The machine learning model trained with ADT successfully *identifies* and *predicts* the lifetimes of eddies, achieving accuracy rates of 90% for LCE identification and 93% for lifetime prediction, along with 71% and 61% for LCFEs, respectively. Incorporating SST and Chl-a combined enhanced eddy predictions over ADT-only or ADT and SST combined, in particular LCEs and LCFEs, highlighting the benefits of assimilating Chl-a into ocean models to improve the representation and the forecast of these eddies. This machine learning framework has the potential to advance predictions of eddy lifetimes and the advection of various tracers.

Plain Language Summary

Lagrangian coherent eddies are types of vortices in the ocean that trap water in their interior and transport it without exchange with the exterior water. These eddies play a key role in transporting water properties such as heat and salt, as well as tracers such as oil, larvae, and seaweed (e.g., Sargassum) across the ocean. For example, during the 2010 Deepwater Horizon oil spill, a type of eddy called a Loop Current Frontal Eddy (LCFE) trapped some of the oil, keeping it from reaching the Florida Keys. This study uses machine learning and satellite data—sea surface height, sea surface temperature, and chlorophyll concentration—to identify and predict the lifetimes of these eddies in the GoM. Three durations of eddy coherence (5, 10, and 20 days) are analyzed. The machine learning model trained with sea surface height successfully *identifies* and *predicts* the lifetimes of eddies, achieving accuracy rates of 90% for LCE identification and 93% for lifetime prediction, along with 71% and 61% for LCFEs, respectively. Adding chlorophyll data from satellite improved the predictions compared to using sea surface height and temperature alone. This machine learning framework can advance predictions of eddy lifetimes and tracer transport.

1 Introduction

Mesoscale eddies are important contributors to the transport of water masses, heat, salt, and passive tracers within the ocean (Dong et al., 2014). The ability of ocean vortices to trap and transport water and passive tracers without exchange with the exterior is denominated Lagrangian coherence (Haller & Beron-Vera, 2013). By definition, no mass flux occurs across the boundary of a Lagrangian coherent eddy, ensuring that water is conserved within its interior with no exchange with the surroundings. Therefore, Lagrangian coherent eddies are very efficient in transporting water properties (heat, salt, and oxygen) and tracers (oil, larva, and Sargassum algae) across the ocean.

In the Gulf of Mexico (GoM), Loop Current Eddies (LCEs) are formed by the detachment of a portion of the Loop Current (LC) (Figure 1) and they transport warmer and saltier Caribbean waters to the central and western Gulf, where they eventually mix with local colder and fresher Gulf waters (Meunier et al., 2018, 2024). These warm eddies, which have been shown to remain Lagrangian coherent for up to three months (Beron-Vera et al., 2018), present high values of tropical cyclone heat potential and are known to fuel hurricane intensification (Shay et al., 2000; Shay, 2000; Jaimes et al., 2016). Another type of eddy in the Gulf, smaller, cold-core eddies, actively modulate the local circulation by attracting, trapping, and transporting water in their interior—these cold eddies have played a particularly important role in the capture of offshore oil from the 2010 Deepwater Horizon oil spill, preventing contamination of the Florida keys (Walker et al., 2013; Hiron et al., 2022). These smaller cyclonic eddies are called Loop Current Frontal Eddies (LCFEs) and are formed by barotropic and baroclinic instability of the LC. They propagate in the vicinity of the LC (Donohue et al., 2016a, 2016b) and are Lagrangian coherent from the surface down to $\sim$600 m and can conserve water in their interior for up to a month (Hiron et al., 2022). In addition to modulating the circulation in the eastern GoM, LCFEs also contribute to the shedding of LCEs by intensifying and constricting the neck of the LC (Cochrane, 1972; Vukovich & Maul, 1985; Hiron et al., 2020).

Haller and Beron-Vera (2013) (hereinafter referred to as HBV13) developed a method to identify the boundaries of Lagrangian coherent eddies based on trajectories derived from gridded velocity fields such as model outputs or geostrophic velocities derived from altimetry. This method was used to study, among others, the Lagrangian coherence of LCEs and LCFEs (Beron-Vera et al., 2018; Hiron et al., 2022). Although efficient, mathematically exact, and objective, the HBV13 method is computationally expensive, and its usage is limited to a small community familiar with the algorithm. In this manuscript, we first test a machine learning alternative that is able to identify Lagrangian coherent eddies in the GoM using solely Absolute Dynamic Topography (ADT) maps from altimetry. Such a machine learning model would allow for much faster detection and would be more user-friendly for the oceanographic community than HBV13. A secondary goal of this study is to test as to whether such a machine learning model can be used to *predict* if a given eddy/rotational feature will become Lagrangian coherent and for how long it will remain using only ADT from the present and previous days. Finally, the third goal of this study is to evaluate whether the inclusion of satellite-derived chlorophyll-a (Chl-a) maps, alongside altimetry-derived ADT and sea surface temperature (SST), compared to ADT-only or ADT and SST models, contributes to improving the performance of the machine learning model in identifying and predicting Lagrangian coherent eddies, specifically LCEs and LCFEs. LC and LCEs exhibit lower Chl-a concentrations than surrounding Gulf waters, making these features easily detectable on Chl-a maps, in particular in spring, summer, and fall (Chassignet et al., 2005; Hiron et al., 2022; Trott et al., 2024). LC and LCEs can also be detected in SST maps, in particular in fall, winter, and spring, due to their higher temperatures compared to the surrounding colder Gulf waters (e.g., Walker et al., 2013). An interesting aspect of using these two datasets together is that they are complementary seasonally: when the LC and LCEs become indistinguishable from the background waters on the SST maps in the summer, the Chl-a gradient between the LC/LCE and Gulf waters is at its peak (e.g., Trott et al., 2024; Walker et al., 2013). In winter, the opposite occurs.

In addition to providing complementary information from multiple satellite fields, another important aspect of incorporating additional satellite data is evaluating the benefits of including Chl-a, along with ADT and SST, for identifying LCEs and LCFEs in data-assimilative ocean models. Current discussions within the oceanography community revolve around assimilating Chl-a into regional GoM models (e.g., the 1/100° HYbrid Coordinate Ocean Model using the data assimilative Tendral Statistical Interpolation package or HYCOM-TSIS; Ntaganou et al. (2024)). In terms of satellite observations, only altimetry ADT tracks ($\sim$1/16°) and SST gridded product ($\sim$1/10°) are currently being assimilated into HYCOM-TSIS (A. Bozec, *personal communication*), the primary model used for operational forecasting in the GoM with very-high resolution ($\sim$1/100°). Satellite Chl-a ($\sim$1/25°) could provide additional, higher-resolution information on the location and structure of the LC, LCE, and, potentially, LCFE fronts if assimilated into HYCOM-TSIS. Therefore, quantifying the improvements that Chl-a provides in identifying and forecasting LCEs and LCFEs is a first estimation of the benefits of assimilating this field in terms of better placing those features in the models, with potential benefits to regional forecast models, in particular for forecasting LCE detachments and their evolution.

To our knowledge, this is the first study using machine learning to detect Lagrangian coherent eddies and to *predict* the Lagrangian coherence of eddies and their lifetime, which is not possible to do with the HBV13 approach since it requires the integration of trajectories during the coherence period of the eddies. Machine learning has been a powerful tool for the prediction of climate signals (e.g., Arcodia et al., 2023) and one aim of this paper is to demonstrate that it can also effectively be used to predict the behavior of mesoscale features that significantly influence upper-ocean transport. Predicting the lifetime of ocean vortices in the Gulf can have different applications, such as forecasting the transport of water masses or tracers such as oil and Sargassum.

The structure of the paper is as follows: Section 2 describes the datasets; Section 3 describes the HBV13 method and the machine learning model; in Section 4 are the results and discussions; and we finish with conclusions in Section 5.

2 Datasets

This study uses three satellite-derived datasets to train the ML models for identifying and predicting Lagrangian coherent eddies in the Gulf of Mexico: altimetry ADT, SST, and Chl-a data.

2.1 Absolute Dynamic Topography (ADT)

ADT data was obtained from the Copernicus Marine Environment Monitoring Service (CMEMS) Sea Level Thematic Assembly Center (Copernicus Marine Service, 2022). ADT represents the sea surface height above the geoid and includes both the mean dynamic topography and the sea level anomalies. The ADT data are distributed as daily, delayed-time Level-4 gridded products, derived from multiple satellite altimetry missions such as TOPEX/Poseidon, Jason series, Sentinel-3A and 3B, CryoSat-2, and others. The dataset covers the global ocean with a horizontal grid-spacing of $1/4°$ ($\sim$ 25 km in the GoM) in both latitude and longitude. The temporal coverage used in this study spans from January 1993 to December 2022. For this study, we focus on the Gulf of Mexico region, extracting ADT data within the domain of 18°N to 32°N latitude and 99°W to 75°W longitude. The daily ADT mean over the deep Gulf waters ($\geq$ 200 m) was removed from each daily field to remove the contraction/expansion due to seasonal changes in ADT, as done in Leben (2005) and Hiron et al. (2020). Geostrophic velocities derived from the ADT fields are used to compute trajectories, which are then used to find the boundary of Lagrangian coherent eddies. Additionally, the ADT data serve as a primary input to the machine learning models for eddy detection and prediction. Link for data access: `https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_L4_MY_008_047/description`.

2.2 Sea Surface Temperature (SST)

SST data was obtained from the OSTIA system, developed by the UK Met Office and distributed through the Group for High-Resolution Sea Surface Temperature (GHRSST) and CMEMS (Donlon et al., 2012). The OSTIA product is a Level-4, high-resolution analysis that merges observations from various satellite sensors and in situ observations to provide gap-free global SST fields. The OSTIA SST data have a horizontal grid-spacing of $1/20°$ ($\sim$ 5 km in the GoM) and are available daily from October 1981 to the present. To ensure consistency with the ADT data, the SST fields are interpolated onto the same grid covering the GoM. Link for data access: `https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_011/description`.

2.3 Chlorophyll-a Concentration (Chl-a)

Chlorophyll-a concentration data were obtained from the CMEMS ocean color products, which include data from the SeaWiFS and other ocean color sensors such as MODIS-Aqua, VIIRS, and the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3A and 3B (Gohin et al., 2002). The Chl-a dataset is a Level-4, multi-sensor merged product that provides daily, gap-free coverage of global chlorophyll concentration at the ocean surface. The Chl-a data have a native horizontal grid-spacing of approximately 4 km in the GoM. For consistency with the ADT and SST datasets, the Chl-a fields are interpolated onto the same 0.25° grid covering the Gulf of Mexico. This dataset is available from September 1997 to the present. For this study, we utilized data spanning from January 1998 to December 2022. Link for data access: `https://data.marine.copernicus.eu/product/OCEANCOLOUR_GLO_BGC_L4_MY_009_104/description`.

2.4 Data Preprocessing

All datasets underwent preprocessing steps to ensure compatibility and optimal performance in the machine learning models:

- **Spatial Interpolation**: The ADT, SST, and Chl-a data were interpolated onto a common regular grid with a spatial resolution of 0.25° in both latitude and longitude, covering the Gulf of Mexico from 18°N to 31°N and 98°W to 80°W.
- **Temporal Alignment**: Daily data from all datasets were temporally aligned to ensure that observations from the same date were used together. This alignment is critical for capturing the coincident physical and biological signals associated with eddies.
- **Seasonal adjustment**: For all variables, the spatial mean was removed on a daily basis to eliminate seasonal effects. This process allows the ML models to focus on the anomalies of the fields. The removal of the spatial mean effectively creates daily anomaly fields for ADT, SST, and Chl-a.
- **Normalization**: All these input fields were normalized to have zero mean and unit variance. Normalization is important for machine learning models to ensure that all input features contribute equally to the training process.

3 Methods

3.1 Identifying the boundary of Lagrangian coherent eddies

We use the method developed by Haller and Beron-Vera (2013) to identify the boundary of Lagrangian coherent eddies. This methodology requires the evaluation of a set of trajectories distributed across the domain. The evolution of those trajectories can be represented with the flow map $F_{t_0}^{t_0+T} : \mathbf{x_0} \to \mathbf{x}(t; t_0, \mathbf{x_0})$, which maps the final position of a T-long trajectory starting at $(\mathbf{x_0}, t_0)$.

To identify the structures of a flow field, the HBV13 Lagrangian method is based on the Cauchy-Green tensor, which is formed from the derivatives of the flow map operator.

$$C_{t_0}^{t_0+T}(\mathbf{x_0}) = \nabla F_{t_0}^{t_0+T}(\mathbf{x_0})^\top \nabla F_{t_0}^{t_0+T}(\mathbf{x_0}) \tag{1}$$

The eigenvectors of $C_{t_0}^{t_0+T}(\mathbf{x_0})$ represent the stretching *direction* of the flow at $\mathbf{x_0}$ along the trajectories. Similarly, the eigenvalues λ_i of the tensor represent the stretching *magnitude*. In two dimension, the Cauchy-Green tensor is a two-by-two matrix, so it has 2 sets of eigenvectors (ξ_1, ξ_2) and eigenvalues (λ_1, λ_2) defined at each initial position $\mathbf{x_0}$.

Lagrangian coherent eddies are identified as material loops that defy the typical exponential stretching occurring in unsteady fluids. Such loops $r(s)$ are closed trajectories of

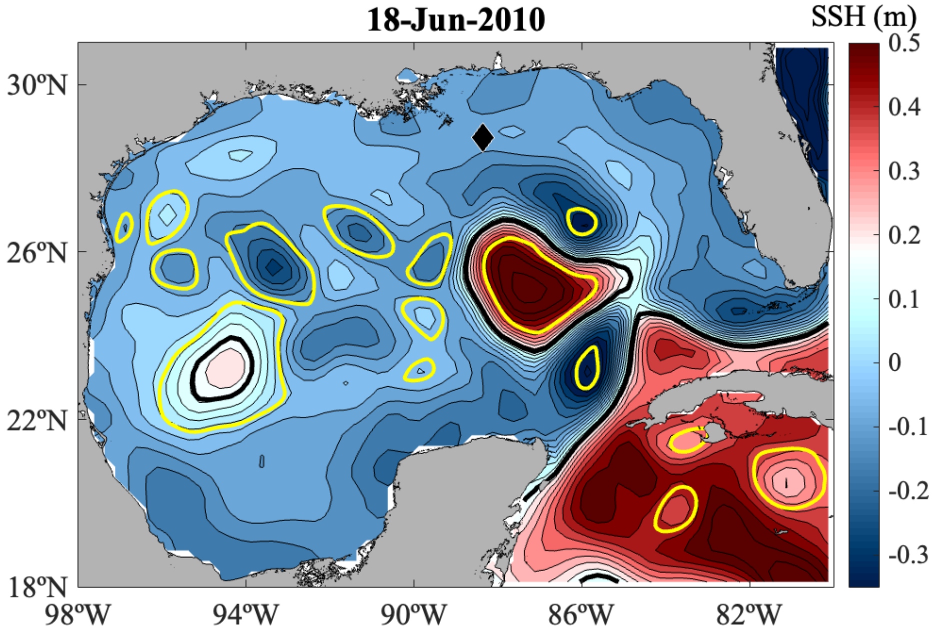

Figure 1. Altimetry Absolute Dynamics Topography (ADT; CMEMS Copernicus). The yellow lines show the vortices that remained Lagrangian coherent for 14 days (from 18 June 2010 to 2 July 2010), and the eddy on the northeast flank for the LCE is the LCFE responsible for attracting and trapping oil during the 2010 Deepwater Horizon oil spill. The black diamond indicates the location of the Deepwater Horizon oil rig. The 17 cm ADT contour is shown by the black line to indicate the LC front.

the vector field $\eta_\lambda^\pm$ and *uniformly stretch* by some amount λ. The $\eta_\lambda^\pm$ field is formed from a combination of both eigenvectors and eigenvalues of $C_{t_0}^{t_0+T}(\mathbf{x_0})$, as follows:

$$r'(s) = \eta_\lambda^\pm(r(s)), \quad \eta_\lambda^\pm = \sqrt{\frac{\lambda_2 - \lambda^2}{\lambda_2 - \lambda_1}}\xi_1 \pm \sqrt{\frac{\lambda^2 - \lambda_1}{\lambda_2 - \lambda_1}}\xi_2. \tag{2}$$

The last step of the methodology is to integrate $r'(s)$ and identify outermost limit cycles of $\eta_\lambda^\pm$ across the domain. We use a methodology described in Karrasch et al. (2015), which allows to efficiently identify locations where coherent eddies can be present, hence speeding up calculations.

The boundary of all Lagrangian coherent eddies from 1993 to 2022 were identified in the GoM using geostrophic velocities derived from altimetry ADT, and for different Lagrangian coherent advection times: 5 days, 10 days, and 20 days.

3.2 Machine learning model

The proposed machine learning models use preprocessed ADT, SST, and Chl-a datasets as inputs and are designed to detect and predict eddies identified via the HBV13 method (Figure 2). We conduct four experiments to assess the effects of incorporating additional satellite data (SST and Chl-a), the importance of the temporal extent of the training data (1993–2002 vs. 1998–2022), and the specific impact of including Chl-a into the models:

1. **ADT Only (1993–2022):** Models are trained and validated using only altimetry-derived ADT data from the full available period (1993–2022). This establishes a baseline, as ADT gradients directly reflect geostrophic currents and mesoscale dynamics. The long temporal coverage allows evaluating how the size of the training dataset influences the model's performance.
2. **ADT, SST, and Chl-a (1998–2022):** Models are trained and validated with ADT, SST, and Chl-a data over their overlapping period (1998–2022) to examine how incorporating multiple satellite products affects eddy detection and prediction.
3. **ADT and SST (1998–2022):** Models are trained and validated using ADT and SST data for 1998–2022 to evaluate whether adding Chl-a information further improves performance over using just ADT and SST.
4. **ADT Only (1998–2022):** Models are trained and validated using only ADT data for 1998–2022 to provide a direct comparison to the multi-dataset configuration in the same temporal window.

The models are based on the U-Net architecture (Ronneberger et al., 2015), a well-established framework. Although U-Net sometimes produces slightly blurred outputs in image-generation tasks and may offer less global context than attention-based variants (Oktay, 2018), these issues are less critical for our application. We do not require high-resolution outputs, and detecting each coherent vortex primarily depends on information from nearby pixels, making U-Net a suitable choice given also our moderate dataset size. All models share the same U-Net backbone, with each day of input data for each modality treated as an additional input channel. To ensure a fair comparison, we only vary the number of input channels, keeping all other parameters (e.g., number of hidden layers, number of filters per layer, filter size, batch normalization) constant. Figure 2 provides a detailed illustration of one such model.

The performance of the ML models is evaluated for the detection and prediction of eddies that remain Lagrangian coherent for 5, 10, and 20 days. For **detection**, the models incorporate data from both before and during the coherence period, similar to the HBV13 method. For **prediction**, only data collected prior to the eddies becoming Lagrangian coherent is used. Both detection and prediction scenarios are assessed under different input configurations (ADT full, ADT, ADT+SST, and ADT+SST+Chl-a). We conducted **six** tests: [-1,0], [-2,0], [0,0], [-2,+T], [-1,+T], and [0,+T], where the first number indicates the number of input days before coherence onset, and the second number represents the coherence period itself. For example, [-2,+T] uses data from two days before the eddy becomes coherent and throughout the entire coherence period T. Each test was run twice to account for variability introduced by random model weight initialization. This process is repeated for each of the three coherence durations (5, 10, and 20 days) and for each of the four input configurations.

3.2.1 Training

Eighty percent of the data was used for training, and the remaining twenty percent was reserved for validation. The proposed architecture is trained using contours identified by the HBV13 method, which relies on trajectories computed from geostrophic velocities derived from altimetric ADT data. These HBV13-generated contours, initially provided as lists of geospatial coordinates, are post-processed into binary grids at a uniform 0.25° resolution in latitude and longitude. In these binary grids, Lagrangian coherent vortices appear as closed masks with values of 1, and these masks serve as ground-truth labels for the ML models.

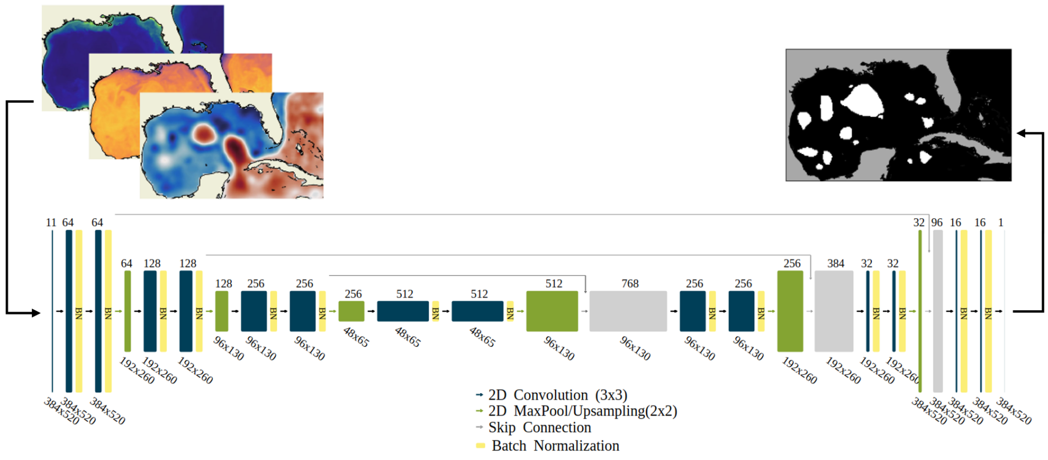

Figure 2. U-Net architecture: (upper-left) Chl-a, SST, and ADT maps serve as input for the machine learning model, which is then segmented into a series of convolutional network (lower panel), and is weighted by the Lagrangian coherent eddies detected using HBV13 (upper-right).

The models are trained using the Adam optimizer (with a learning rate of 0.001) and a learning rate scheduler. Training is terminated if the validation error does not improve for 100 consecutive epochs. The loss function is defined as:

$$\text{DSC Loss(A, B)} = 1 - \frac{2|A \cap B|}{|A| + |B|}, \tag{3}$$

where A is the predicted mask of Lagrangian coherent vortices, and B is the corresponding ground truth mask derived from the HBV13 method. This loss corresponds to $1 -$ the Dice Similarity Coefficient (DSC), and a smaller value indicates better agreement between the predictions and the true vortex masks.

A given Lagrangian coherent eddy is considered as *detected* by the machine learning model if the overlap between the HBV13 eddy and the machine learning eddy is of at least 33%. We examine the performance of the machine learning model in detecting (a) all eddies in the GoM, and only (b) LCEs, and (c) LCFEs. After removing the mean ADT, the ADT field in the GoM varies roughly between -0.35 m and 0.5 m. Since the machine learning model is trained with ADT fields, for (a), we focus on the detection of somewhat stronger eddies, which have an ADT signal on average larger than 0.3 m (anticyclonic eddies) and smaller than -0.1 m (cyclonic eddies). For the detection of LCEs (b), an eddy is considered an LCE if the maximum ADT within the eddy is larger than 0.17 m, which is the contour that has been vastly used to detect the LC and LCE fronts (Leben, 2005; Hiron et al., 2020). For LCFEs (c), a given eddy is considered an LCFE if it is located east of 90°W, the minimum distance between the given eddy and the 17 cm ADT contour (LCFE or LCE) is smaller than 100 km, similar to Hiron et al. (2020), the averaged ADT is smaller than -0.1 m, and the minimum ADT is smaller than -0.2 m. The thresholds used to detect LCEs and LCFEs were validated visually.

4 Results and discussion

4.1 Validation loss

The validation loss (VL) is a metric used during the training of artificial neural networks to assess the model's performance on a validation set. The VL provides an estimate of how

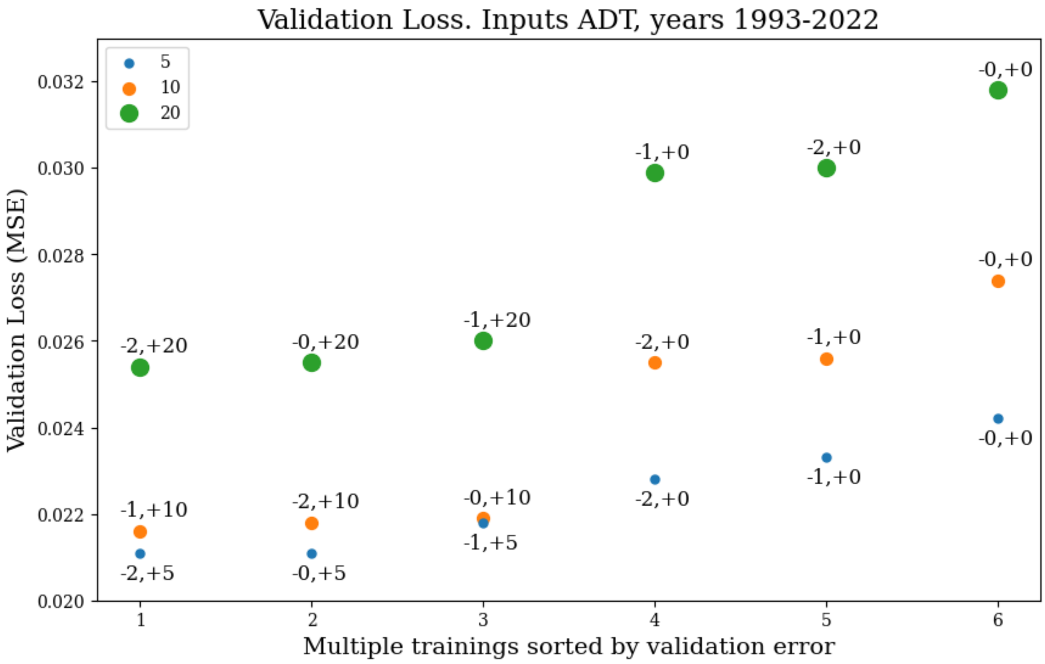

Figure 3. Validation loss using ADT maps for training sequences performed with various input periods to detect Lagrangian coherent eddies of different coherent times (5, 10, and 20 days; colors). The number before (after) the negative (plus) sign indicates the number of daily ADT maps inputted before (during) the eddies' Lagrangian coherence.

well the model will generalize to unseen examples; when the validation set is large and properly representative of the true data distribution, this estimate is typically reliable. In other words, VL indicates how effectively the model has learned patterns during training, with lower values signifying better performance. In Figure 3, we show the performance of all machine learning models for the Lagrangian coherent advection times of 5, 10, and 20 days (distinct colors), and trained using ADT from different input periods: (i) during the time of eddy Lagrangian coherence ([-0,+T], where T is the time of Lagrangian coherence of the eddies), (ii) only days before the eddies become Lagrangian coherent ([-2,+0] and [-1,+0]), and (iii) all combined ([-1,+T] and [-2,+T]). Evaluating a machine learning model trained solely on satellite data before the eddies become Lagrangian coherent gives insights into the ability of the model to *predict* the coherence of eddies.

The trained models exhibit three distinct levels of performance (Figure 3). The best performance (lower values) corresponds to *detection*, which leverages information from days when the eddy is coherent ($[-2,+T]$, $[-1,+T]$, and $[-0,+T]$). Next, we see the models used for *prediction*, only incorporating data from earlier days ($[-2,+0]$ and $[-1,+0]$). Finally, performance decreases when only the current day is provided as input ($[-0,+0]$). The VLs for the trainings using ADT and SST, and ADT, SST, and Chl-a presented a similar pattern.

4.2 Identification and prediction of Lagrangian coherent eddies using machine learning and ADT

The machine learning models trained using ADT fields successfully detect the Lagrangian coherent eddies present in the GoM with different lifetimes, spanning from 5 to 20 days (Figure 4a,d for a 10-day lifetime). When trained and validated with all ADT data available (1993–2022), we find that machine learning can identify ([-2,+T]) 65% of the ed-

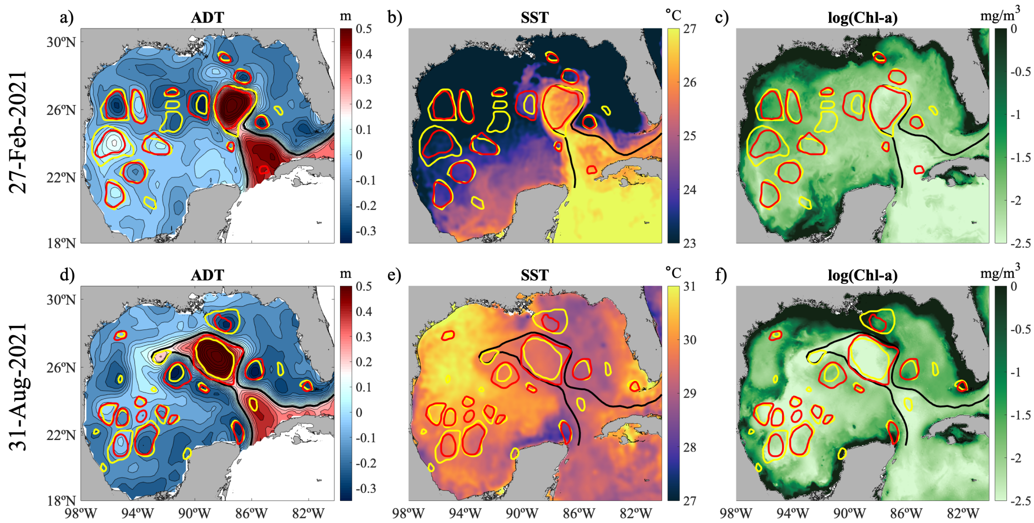

Figure 4. (a,d) ADT, (b,e) SST, and (c,f) Chl-a maps for the winter (a-c, 27 February 2021) and summer (d-f, 31 August 2021) superposed with the Lagrangian coherent eddy boundaries (10 days of coherence) detected with the HBV13 method (yellow contours) and the machine learning model (red contours) trained and validated with ADT data ([-2, 10]) for the 1993–2022 period. Note that the temperature colorbar ranges differ between the two dates, and it was specifically chosen to emphasize the absence of LC/LCE signatures during the summer. The black line is the 17-cm ADT contour used to track the LC and LCEs.

dies that remain Lagrangian coherent for 5 days, 65% of the 10-day lived eddies, and 61% of the eddies with a 20-day lifetime. For the prediction ([-2,0]) of Lagrangian coherent eddies, the machine learning models were trained using solely two days of ADT data and HBV13 eddy contours prior to the coherence of the eddies. We find that machine learning trained with ADT fields can predict 54% of the 5-day, 54% of the 10-day, and 50% of the 20-day lived eddies.

The performance of the machine learning models is even more effective in identifying and predicting the Lagrangian coherence of just the LCEs and LCFEs. For the LCEs, the models accurately identified 87% of the 5-day lived eddies, 87% of the 10-day lived, and 90% of the 20-day ones, and accurately predicted 86%, 87%, and 93% of the 5-, 10-, and 20-day coherent LCEs, respectively. For LCFEs, the models accurately identified 71% of the 5-day lived eddies, 71% of the 10-day, and 51% of the 20-day ones, and accurately predicted 60%, 61%, and 43% of the 5-, 10-, and 20-day coherent LCFEs, respectively.

The decrease in the percentage of detection of eddies with a lifetime of 20 days, in particular for all eddies and the LCFEs, is likely due to the lower number of eddies that live 20 days in the GoM (Tables 1–3), which decreases the number of data available to train the model and, therefore, impacts its performance. LCEs, on the other hand, can remain Lagrangian coherent for much longer (up to 200 days; Beron-Vera et al., 2018), and 20-day Lagrangian coherent LCEs tend to be more organized, and thus have a more detectable ADT signal, explaining the increase in the percentage of detection for both identification and prediction for 20-days lived eddies in comparison with 5- and 10-day lived ones.

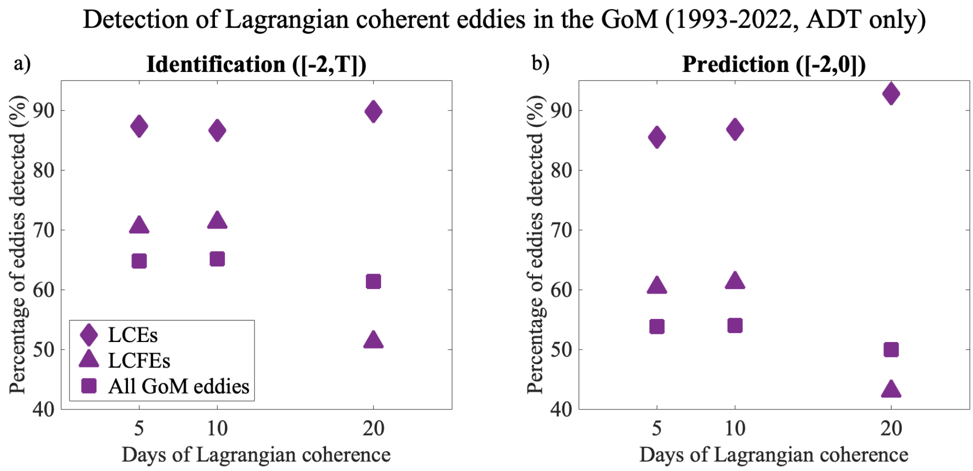

Figure 5. Percentage of Lagrangian coherent eddies detected for different coherent times (5, 10, and 20 days) for the machine learning model trained and validated only with ADT (1993–2022) to (a) identify [-2,+T] and (b) predict [-2,+0] all Lagrangian coherent eddies in the Gulf of Mexico (square), LCEs (diamonds), and LCFEs (triangles). The number before (after) the negative (plus) sign indicates the number of daily ADT maps inputted before (after) the eddies became Lagrangian coherent.

4.3 Prediction and identification of Lagrangian coherent eddies using machine learning and ADT, SST, and Chl-a

4.3.1 Combined SST and Chl-a increases the prediction of eddies

The addition of combined SST and Chl-a data to the training of the machine learning models, along with ADT data, enhanced the forecast (input [-2,0]) of all the eddies in the GoM, including the forecast of LCFEs and LCEs, compared to models trained with ADT-only, or ADT and SST combined without Chl-a (Figure 6). When comparing with models trained with ADT alone, the inclusion of combined SST and Chl-a data increased the detection of predicted Lagrangian coherent eddies for all eddies in the GoM from 57% to 58% for 5-day lived eddies, from 55% to 56% for 10-day, and from 55% to 57% for 20-day lived ones (Figure 6a). For LCFEs, the detection of predicted eddies increased from 62% to 67% for 5-day coherence, from 63% to 68% for 10-day coherence, and from 46% to 47% for 20 days (Figure 6b). For LCEs, the increase in the predicted eddies occurred for 10-day coherence (from 89% to 92%) and 20 days (from 89% to 95%) lived eddies, but not for 5-day lived LCEs, in which the detection decreased from 89% to 84% (Figure 6c). Note that the models trained with ADT and SST, without Chl-a, underperformed, in some cases, the models trained with ADT alone.

The contribution of combined SST and Chl-a in the improvement of the machine learning models detecting eddies in the GoM is visible in Figure 4b,c,e,f, in particular for LCFEs and LCEs. In fall, winter, and spring, when the Gulf is still colder than Caribbean waters, the LC and LCE have a thermal signature at the surface, which is not visible in the summer when the Gulf temperature rises. In terms of Chl-a, the nutrient-poorer LC and LCEs are distinguishable from the background Gulf waters in spring, summer, and fall (e.g., Chassignet et al., 2005; Trott et al., 2024). Since LCEs are Lagrangian coherent vortices (Beron-Vera et al., 2018), the warmer and nutrient-poorer waters remain inside the LCEs with minimum exchange with surrounding waters. Additionally, the strong flow associated with the LC and LCEs fronts facilitates the advection of Mississippi River waters rich in

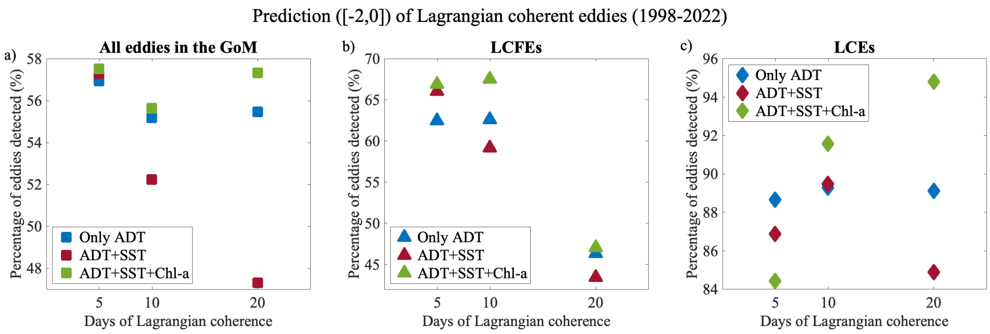

Figure 6. Percentage of Lagrangian coherent vortices detected for different coherent times (5, 10, and 20 days) for the machine learning model trained and validated with only ADT (blue), ADT and SST (red), and a combination of ADT, SST, and Chl-a (green) to predict [-2,+0] Lagrangian coherent eddies: (a) all eddies in the Gulf of Mexico, (b) LCFEs, and (c) LCEs.

Chl-a from the shelf along the front, enhancing the Chl-a gradient across the boundary of the LC and LCEs (Figure 4f). In winter, due to higher winds and winter convective mixing, the signature of LC/LCEs in the Chl-a maps decreases (Damien et al., 2021). The lack of eddy signature in the SST fields in the summer could be the (or one of the) reason(s) why the models with ADT-only outperformed, in some cases, the models with combined ADT and SST.

For the LCFEs, the signature in SST is mostly due to the horizontal advection of warmer water from the LC/LCE around the cold-core eddies (Figure 4b for both LCFEs on the east and north flank of the LC). Another SST signature associated with LCFEs can be a cool signature associated with the upward lifting of deeper isotherms to the upper ocean due to geostrophic adjustment, especially when these LCFEs are larger and stronger, which occurs typically in the northern and eastern flanks of the LC in the last stages before and during LCE shedding (Hiron et al., 2020). Similarly to LCEs, large LCFEs are also Lagrangian coherent structures (Hiron et al., 2022), which means that they have well-defined boundaries and preserve water in their interior during the time of coherence. However, an important distinction from LCEs is that LCFEs are formed by Gulf waters (Hiron et al., 2022). Thus, their signature in Chl-a maps is not due directly to their coherence, as with LCEs, but is instead likely associated with the strong velocities in the LC-LCFE (LCE-LCFE) fronts that attract Mississippi River, nutrient-rich waters along the LC-LCFE front and around the LCFEs (e.g., Androulidakis et al., 2014; Hiron et al., 2022). This configuration occurs particularly when the LCFEs are in the northern and eastern flanks of an extended LC or in the vicinity of an LCE (e.g., LCFE on the northern flank of the LC in Figure 4f). For more information on LC-LCFE fronts we recommend consulting Olascoaga and Haller (2012), Hiron et al. (2020), and Hiron et al. (2022). Some cases of very strong LCFEs can cause vertical advection of deeper, rich-in-nutrient waters to the surface (e.g., LCFE on the eastern flank of the LC in Figure 4f, and Hiron et al. (2020)).

4.3.2 The impact of SST and Chl-a in the identification of eddies

Contrary to the prediction ([-2,0]) of eddies, the inclusion of SST and Chl-a maps decreased the number of eddies identified ([-2,+T]) using the machine learning models (see Tables 1-3). We believe this is due to the discontinuity presented in SST and Chl-a maps caused by cloud coverage, which can impact the performance of the machine learning model. For the prediction of eddies, only data for two days (prior to the coherence) is used, whereas for the identification, a total of 7 days ([-2,+5]) of data is used for the 5-day of Lagrangian coherence and a total of 12 and 22 days of data is used for 10 and 20 days of Lagrangian coherence, respectively. Therefore, using multiple days as input increases the chances of having maps with missing data, impacting the performance of the machine-learning models.

5 Conclusions

This study explores the ability of a machine learning model to identify and predict the Lagrangian coherence of eddies in the GoM using only satellite-derived observations, including ADT, SST, and Chl-a datasets. Three different eddy lifetimes are tested: 5, 10, and 20 days. The identification of the eddies is done by using, for both training and validation, data from two days prior to Lagrangian coherence and days for the whole period of the eddies' lifetime (or period when the eddy is Lagrangian coherent). For the prediction runs, only data from the two days prior to coherence were used. Eight sets of training were performed: one using solely ADT data for the whole period of data available (1993–2022) and one using ADT, SST, and Chl-a for the period of time these datasets overlap (1998–2022). A third and fourth sets of trainings were conducted using only ADT, and combined ADT and SST data for the period from 1998 to 2022, aiming to compare with the run that combined ADT, SST, and Chl-a data. Each of these configurations was used to train two models: one to identify and another to predict Lagrangian coherent eddies. The machine learning approach identifies vortices significantly faster than HBV13, achieving speedups of up to two orders of magnitude (1s with machine learning vs. 180s with a Julia implementation of HBV13), in addition to being a much more user-friendly way to detect Lagrangian coherent eddies.

We find that machine learning can identify and predict Lagrangian coherent eddies in the GoM for different eddy's lifetimes relying solely on current ADT information, which is not possible with the HBV13 method. This is the first study to use machine learning for detecting Lagrangian coherent eddies and predicting their Lagrangian coherence and lifetime, which is also not possible to do with the HBV13 method. On average, the machine learning models trained and validated with ADT (1993–2022) identified 65% of the eddies with lifetimes of 5 and 10 days and predicted 54% of the eddies with Lagrangian coherence of 5 and 10 days. The ability to detect eddies decreased for eddies with longer lifetimes. The performance of the machine learning models increased when testing only for the detection of two important types of eddies in the GoM: LCE and LCFEs. The models identified 87% of the LCEs with a lifetime of 5 and 10 days, and 90% for 20 days. For the prediction of LCEs, the machine learning model detected 86% of LCEs for a lifetime of 5 days, 87% for 10 days, and 93% for 20 days. For LCFEs, the percentage of detection for identified LCFEs was 71% for 5-day and 10-day lived eddies, and 51% for 20-day lived eddies. For the forecasted LCFEs, the percentage of detected eddies was 60% for 5-day lived eddies, 61% for 10-day ones, and 43% for the 20-day ones.

We also find that the inclusion of SST and Chl-a combined in the training and validation of the machine learning models, in addition to ADT, for the 1998–2022 period, increases the prediction of the Lagrangian coherence of eddies, including the LCEs and LCFEs. This approach outperforms models trained with ADT alone or with ADT and SST without Chl-a. This finding is particularly important given the current discussions within the GoM modeling community related to the assimilation of SST and Chl-a into hindcast and forecast models for the GoM to improve the representation of the LCEs and LCFEs. We find that, for the forecast of LCEs, the inclusion of SST and Chl-a combined increased the detected eddies from 89% to 92% for eddies with 10-day lifetime, and from 89% to 95% for those with 20-day lifetime, compared to the models using ADT-only. Additionally, when comparing with the models using ADT-only, the prediction of LCFEs increased from 62% to 67% for 5 days of coherence, from 63% to 68% for 10 days, and from 46% to 47% for 20 days when including SST and Chl-a combined in the training and validation.

In summary, we demonstrate that (1) machine learning coupled with satellite observations can effectively be used to identify and predict the lifetime of Lagrangian coherent eddies, which have a significant influence on upper-ocean transport, and that (2) Chl-a provides additional information on the Lagrangian coherence of eddies in the GoM, on particular LCEs and LCFEs for both identification and prediction, highlighting the benefits of assimilating this dataset in GoM ocean models. This machine learning framework has the potential to enhance predictions of eddy lifetimes and the advection of various tracers while also making these methods more accessible to the community compared to traditional dynamical system-based eddy extraction techniques.

Acknowledgments

This research is funded by the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine under award number 2000013149. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Gulf Research Program or the National Academies of Sciences, Engineering, and Medicine.

Open Research Section

All satellite data used in this story is publicly available. The ADT fields can be found at `https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_L4_MY_008_047/description`, the SST data at `https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_011/description`, and the Chlorophyll-a concentration data can be downloaded at `https://data.marine.copernicus.eu/product/OCEANCOLOUR_GLO_BGC_L4_MY_009_104/description`.

References

Androulidakis, Y. S., Kourafalou, V. H., & Le Hénaff, M. (2014). Influence of frontal cyclone evolution on the 2009 (ekman) and 2010 (franklin) loop current eddy detachment events. *Ocean Science*, *10*(6), 947–965. Retrieved from https://www.ocean-sci.net/10/947/2014/ doi: 10.5194/os-10-947-2014

Arcodia, M. C., Barnes, E. A., Mayer, K. J., Lee, J., Ordonez, A., & Ahn, M.-S. (2023, sep). Assessing decadal variability of subseasonal forecasts of opportunity using explainable ai. *Environmental Research: Climate*, *2*(4), 045002. Retrieved from https://dx.doi.org/10.1088/2752-5295/aced60 doi: 10.1088/2752-5295/aced60

Beron-Vera, F. J., Olascoaga, M. J., Wang, Y., Triñanes, J., & Pérez-Brunius, P. (2018). Enduring lagrangian coherence of a loop current ring assessed using independent observations. *Scientific Reports*, *8*(1), 11275. doi: https://doi.org/10.1038/s41598-018-29582-5

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Barron, C. N., Ko, D. S., Rhodes, R. C., ... Arnone, R. A. (2005). Assessment of data assimilative ocean models in the gulf of mexico using ocean color. In *Circulation in the gulf of mexico: Observations and models* (p. 87-100). American Geophysical Union (AGU). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/161GM07 doi: https://doi.org/10.1029/161GM07

Cochrane, J. (1972). Separation of an anticyclone and subsequent developments in the loop current. In *Contributions on the physical oceanography of the gulf of mexico* (p. 91-106). Gulf Publishing Co.

Copernicus Marine Service. (2022). *Global Ocean Gridded L4 Sea Surface Heights and Derived Variables Reprocessed.* https://doi.org/10.48670/moi-00149. (Dataset)

Damien, P., Sheinbaum, J., Pasqueron de Fommervault, O., Jouanno, J., Linacre, L., & Duteil, O. (2021). Do loop current eddies stimulate productivity in the gulf of mexico? *Biogeosciences*, *18*(14), 4281–4303. Retrieved from https://bg.copernicus.org/articles/18/4281/2021/ doi: 10.5194/bg-18-4281-2021

Dong, C., McWilliams, J. C., Liu, Y., & Chen, D. (2014). Global heat and salt transports by eddy movement. *Nature Communications*, *5*, 3294. doi: https://doi.org/10.1038/ncomms4294

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (ostia) system. *Remote Sensing of Environment*, *116*, 140–158.

Donohue, K., Watts, D., Hamilton, P., Leben, R., & Kennelly, M. (2016b). Loop current eddy formation and baroclinic instability. *Dynamics of Atmospheres and Oceans*, *76*, 195 - 216. Retrieved from http://www.sciencedirect.com/science/article/pii/S0377026516300057 (The Loop Current Dynamics Experiment) doi: https://doi.org/10.1016/j.dynatmoce.2016.01.004

Donohue, K., Watts, D., Hamilton, P., Leben, R., Kennelly, M., & Lugo-Fernández, A. (2016a). Gulf of mexico loop current path variability. *Dynamics of Atmospheres and Oceans*, *76*, 174 - 194. Retrieved from http://www.sciencedirect.com/science/article/pii/S0377026515300130 (The Loop Current Dynamics Experiment) doi: https://doi.org/10.1016/j.dynatmoce.2015.12.003

Gohin, F., Druon, J., & Lampert, L. (2002). A five channel chlorophyll concentration algorithm applied to seawifs data processed by seadas in coastal waters. *International journal of remote sensing*, *23*(8), 1639–1661.

Haller, G., & Beron-Vera, F. J. (2013). Coherent lagrangian vortices: the black holes of turbulence. *Journal of Fluid Mechanics*, *731*, R4. doi: 10.1017/jfm.2013.391

Hiron, L., de la Cruz, B. J., & Shay, L. K. (2020). Evidence of loop current frontal eddy intensification through local linear and nonlinear interactions with the loop current. *Journal of Geophysical Research: Oceans*, *125*(4), e2019JC015533. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JC015533 (e2019JC015533 10.1029/2019JC015533) doi: https://doi.org/10.1029/2019JC015533

Hiron, L., Miron, P., Shay, L. K., Johns, W. E., Chassignet, E. P., & Bozec, A. (2022). Lagrangian coherence and source of water of loop current frontal eddies in the gulf of mexico. *Progress in Oceanography*, *208*, 102876. Retrieved from https://www.sciencedirect.com/science/article/pii/S0079661122001355 doi: https://doi.org/10.1016/j.pocean.2022.102876

Jaimes, B., Shay, L. K., & Brewster, J. K. (2016). Observed air-sea interactions in tropical cyclone isaac over loop current mesoscale eddy features. *Dynamics of Atmospheres and Oceans*, *76*, 306-324. Retrieved from https://www.sciencedirect.com/science/article/pii/S0377026516300203 (The Loop Current Dynamics Experiment) doi: https://doi.org/10.1016/j.dynatmoce.2016.03.001

Karrasch, D., Huhn, F., & Haller, G. (2015). Automated detection of coherent lagrangian vortices in two-dimensional unsteady flows. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *471*(2173), 20140639. Retrieved from https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2014.0639 doi: 10.1098/rspa.2014.0639

Leben, R. R. (2005). Altimeter-derived loop current metrics. In *Circulation in the gulf of mexico: Observations and models* (p. 181-201). American Geophysical Union (AGU). doi: 10.1029/161GM15

Meunier, T., Bower, A., Pérez-Brunius, P., Graef, F., & Mahadevan, A. (2024). The energy decay of warm-core eddies in the gulf of mexico. *Geophysical Research Letters*, *51*(1), e2023GL106246. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2023GL106246 (e2023GL106246 2023GL106246) doi: https://doi.org/10.1029/2023GL106246

Meunier, T., Pallás-Sanz, E., Tenreiro, M., Portela, E., Ochoa, J., Ruiz-Angulo, A., & Cusí, S. (2018). The vertical structure of a loop current eddy. *Journal of Geophysical Research: Oceans*, *123*(9), 6070-6090. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JC013801 doi: https://doi.org/10.1029/2018JC013801

Ntaganou, N., Chassignet, E. P., & Bozec, A. (2024). Impact of horizontal model resolution on mixing and dispersion in the northeastern gulf of mexico. *Journal of Geophysical Research: Oceans*, *129*(11), e2024JC021315. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2024JC021315` (e2024JC021315 2024JC021315) doi: https://doi.org/10.1029/2024JC021315

Oktay, O. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.

Olascoaga, M. J., & Haller, G. (2012). Forecasting sudden changes in environmental pollution patterns. *Proceedings of the National Academy of Sciences*, *109*(13), 4738–4743. Retrieved from `https://www.pnas.org/content/109/13/4738` doi: 10.1073/pnas.1118574109

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted intervention – miccai 2015* (pp. 234–241). Cham: Springer International Publishing.

Shay, L. K. (2000). Air-sea interactions in tropical cyclones. In *Global perspectives on tropical cyclones* (p. 93-131). Retrieved from `https://www.worldscientific.com/doi/abs/10.1142/9789814293488_0003` doi: 10.1142/9789814293488_0003

Shay, L. K., Goni, G. J., & Black, P. G. (2000). Effects of a warm oceanic feature on hurricane opal. *Monthly Weather Review*, *128*(5), 1366 - 1383. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/128/5/1520-0493_2000_128_1366_eoawof_2.0.co_2.xml` doi: 10.1175/1520-0493(2000)128⟨1366:EOAWOF⟩2.0.CO;2

Trott, C. B., Subrahmanyam, B., Hiron, L., & Zavala-Romero, O. (2024). Tracking loop current eddies in the gulf of mexico using satellite-derived chlorophyll-a. *Remote Sensing*, *16*(12). Retrieved from `https://www.mdpi.com/2072-4292/16/12/2234` doi: 10.3390/rs16122234

Vukovich, F. M., & Maul, G. A. (1985). Cyclonic eddies in the eastern gulf of mexico. *Journal of Physical Oceanography*, *15*(1), 105-117. doi: 10.1175/1520-0485(1985)015⟨0105:CEITEG⟩2.0.CO;2

Walker, N. D. N. D., Pilley, C. T. C. T., Raghunathan, V. V. V. V., D'Sa, E. J. E. J., Leben, R. R. R. R., Hoffmann, N. G. N. G., … Turner, R. E. R. E. (2013). Impacts of loop current frontal cyclonic eddies and wind forcing on the 2010 gulf of mexico oil spill. monitoring and modeling the deepwater horizon oil spill: A record-breaking enterprise. In *Monitoring and modeling the deepwater horizon oil spill: A record-breaking enterprise* (p. 103-116). (eds Y. Liu, A. Macfadyen, Z.-G. Ji and R.H. Weisberg), American Geophysical Union (AGU). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GM001120` doi: 10.1029/2011GM001120

Table 1. Number of eddies detected by the HBV13 method and the machine learning for both identification and prediction of coherent eddies for all eddies in the Gulf of Mexico.

All eddies in the Gulf of Mexico								
ADT 1993-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	7329	4749	64.80	5 days	7329	3946	53.84
	10 days	7998	5208	65.12	10 days	7998	4319	54.00
	20 days	5076	3115	61.37	20 days	5076	2536	49.96
ADT 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	6176	3681	59.60	5 days	6176	3516	56.93
	10 days	6589	4136	62.77	10 days	6589	3636	55.18
	20 days	3970	2373	59.77	20 days	3970	2202	55.47
ADT + SST 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	6176	3745	60.64	5 days	6176	3538	57.29
	10 days	6589	3914	59.40	10 days	6589	3442	52.24
	20 days	3970	2196	55.31	20 days	3970	1878	47.30
ADT + SST + Chl-a 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	5969	3422	57.33	5 days	5972	3435	57.52
	10 days	6585	3832	58.19	10 days	6585	3664	55.64
	20 days	3967	2436	61.41	20 days	3967	2274	57.32

Table 2. Number of eddies detected by the HBV13 method and the machine learning for both identification and prediction of coherent eddies for Loop Current Frontal Eddies.

Loop Current Frontal Eddies								
ADT 1993-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	1319	930	70.51	5 days	1319	797	60.42
	10 days	1286	917	71.31	10 days	1286	787	61.20
	20 days	677	347	51.26	20 days	677	291	42.98
ADT 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	1140	747	65.53	5 days	1140	712	62.46
	10 days	1105	759	68.69	10 days	1105	692	62.62
	20 days	559	267	47.76	20 days	559	259	46.33
ADT + SST 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	1140	800	70.18	5 days	1140	753	66.05
	10 days	1105	756	68.42	10 days	1105	654	59.19
	20 days	559	257	45.97	20 days	559	243	43.47
ADT + SST + Chl-a 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	1114	745	66.88	5 days	1114	745	66.88
	10 days	1105	772	69.86	10 days	1105	746	67.51
	20 days	559	302	54.03	20 days	559	263	47.05

Table 3. Number of eddies detected by the HBV13 method and the machine learning for both identification and prediction of coherent eddies for Loop Current Eddies.

Loop Current Eddies								
ADT 1993-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	822	718	87.35	5 days	822	703	85.52
	10 days	910	788	86.59	10 days	910	790	86.81
	20 days	764	686	89.79	20 days	764	709	92.80
ADT 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	644	582	90.37	5 days	644	571	88.66
	10 days	700	609	87.00	10 days	700	625	89.29
	20 days	597	576	96.48	20 days	597	532	89.11
ADT + SST 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	610	535	87.70	5 days	610	530	86.89
	10 days	685	571	83.36	10 days	685	613	89.49
	20 days	596	543	91.11	20 days	596	506	84.90
ADT + SST + Chl-a 1998-2022	**Identification ([-2,+T])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%	**Prediction ([-2,0])**	Lagrangian coherent eddies (HBV13)	Detected by machine learning	%
	5 days	589	507	86.08	5 days	591	499	84.43
	10 days	700	617	88.14	10 days	700	641	91.57
	20 days	596	563	94.46	20 days	596	565	94.80