EddyScan: A Physically Consistent Ocean Eddy Monitoring Application

James H. Faghmous[†], Luke Styles[†], Varun Mithal[†], Shyam Boriah[†], Stefan Liess[‡], and Vipin Kumar[†] [†]Department of Computer Science and Engineering [‡]Department of Soil, Water and Climate University of Minnesota, Minneapolis, MN, USA {jfagh, styles, mithal, sboriah, liess, kumar} @cs.umn.edu

Abstract-Rotating coherent structures of water known as ocean eddies are the oceanic analog of storms in the atmosphere and a crucial component of ocean dynamics. In addition to dominating the ocean's kinetic energy, eddies play a significant role in the transport of water, salt, heat, and nutrients. Therefore, understanding current and future eddy activity is a central challenge to address future sustainability of marine ecosystems. The emergence of sea surface height observations from satellite radar altimeter has recently enabled researchers to track eddies at a global scale. The majority of studies that identify eddies from observational data employ highly parametrized connected component algorithms using expert filtered data, effectively making reproducibility and scalability challenging. In this paper, we improve upon the state-of-theart connected component eddy monitoring algorithms to track eddies globally. This work makes three main contributions: first, we do not pre-process the data therefore minimizing the risk of wiping out important signals within the data. Second, we employ a physically-consistent convexity requirement on eddies based on theoretical and empirical studies to improve the accuracy and computational complexity of our method from quadratic to linear time in the size of each eddy. Finally, we accurately separate eddies that are in close spatial proximity, something existing methods cannot accomplish. We compare our results to those of the state of the art and discuss the impact of our improvements on the difference in results.

I. INTRODUCTION

Very much like the atmosphere, our planet's oceans experience their own storms and internal variability. The ocean's kinetic energy is dominated by mesoscale variability: scales of tens to hundreds of kilometers over tens to hundreds of days [20, 19, 5]. Mesoscale variability is generally comprised of linear Rossby waves and as nonlinear ocean eddies (coherent rotating structures much like cyclones in the atmosphere; hereby eddies). Unlike atmospheric storms, eddies are a source of intense physical and biological activity (see Figure 1). In contrast to linear Rossby waves, the rotation of nonlinear eddies transports momentum, mass, heat, nutrients, as well as salt and other seawater chemical elements, effectively impacting the ocean's circulation, large-scale water distribution, and biology. Therefore, understanding eddy variability and change over time is of Frode Vikebø° Michel dos Santos Mesquita* °Institute of Marine Research *Bjerknes Centre for Climate Research Bergen, Norway frovik@imr.no michel.mesquita@uni.no



Figure 1. Image from the NASA TERRA satellite showing an anti-cyclonic (counter-clockwise in the Southern Hemisphere) eddy that likely peeled off from the Agulhas Current, which flows along the southeastern coast of Africa and around the tip of South Africa. This eddy (roughly 200 km wide) is an example of eddies transporting warm, salty water from the Indian Ocean to the South Atlantic. We are able to see the eddy, which is submerged *under* the surface because of the enhanced phytoplankton activity (reflected in the bright blue color). This anti-cyclonic eddy would cause a depression in subsurface density surfaces in sea surface height (SSH) data. Image courtesy of the NASA Earth Observatory. Best seen in color.

critical importance for projected marine biodiversity as well as atmospheric and land phenomena.

Until recently, ocean eddies were tracked using sea surface temperatures (SST) and ocean color. Now, sea surface height (SSH) observations from satellite radar altimeters have emerged as a better-suited alternative for studying eddy dynamics on a global scale. Eddies are generally classified as either cyclonic if they rotate counter-clockwise (in the Northern Hemisphere) or anticyclonic otherwise. Cyclonic



Figure 2. **Top:** A schematic cross section of an anti-cyclonic eddy (in the Northern Hemisphere) density surfaces are depressed within the eddy causing an increase in SSH. The elevation of subsurface density surfaces replenishes the upper part of the ocean with nutrients needed for primary production. **Bottom:** A cyclonic eddy causes an decrease in SSH. Bottom image by Robert Simmons of NASA. Best seen in color.

eddies, like the one in Figure 2 (bottom panel), cause a decrease in SSH and elevations in subsurface density surfaces. Anti-cyclonic eddies, such as the one depicted in Figure 2 (top panel), cause an increase in SSH and depressions in subsurface density surfaces. These characteristics allow us to identify ocean eddies in SSH satellite data. In Figure 3, anti-cyclonic eddies can be seen in patches of positive (dark red) SSH anomalies, while cyclonic eddies are reflected in closed contoured negative (dark blue) SSH anomalies.

We present a global eddy monitoring algorithm, EddyScan, that leverages the physical properties of eddies to increase accuracy and scalability compared to existing methods. Our method has three main contributions: first, we do not pre-process the data, effectively increasing the reproducibility of our results. Second, we employ theoretical and empirical findings on global eddy size distributions to reduce the algorithm's computational complexity. Finally, we improve accuracy by separating merged eddies better than existing methods.

In the next section, we will briefly review existing eddy tracking algorithms. In section 3, we introduce EddyScan and the challenges associated with tracking eddies globally. After that, we present our results and compare them to the eddies identified by Chelton et al. [7] (CH11 hereafter). We conclude the paper with a discussion of the study's contributions and future research directions.

II. PREVIOUS WORK

The earliest methods for automatic identification and tracking of ocean eddies relied primarily on proxy variables such as ocean color or SST. The main challenge when using such proxies to track eddies is that they are influenced by a variety of factors in addition to eddies. Thus it is difficult to link changes in those variables to eddy activity alone. Some of the earliest works based on image processing techniques used an edge detection algorithm to detect eddies along the Gulf Stream [14]. Similar image-based algorithms included a neural-network model trained to identify eddies from SST images [3] and an edge detection scheme to isolate eddies between two consecutive SST images [12]. D'Alimonte [8] used the isothermal lines of the SST field to automatically detect eddies. Finally, Dong et al. [9] transformed SST observations into a thermal-wind-velocity field and subsequently tracked eddies in the transformed space.

The recent introduction of SSH satellite observations provided researchers with data that are directly related to ocean eddies. The majority of eddy tracking algorithms define eddies as closed contoured (positive or negative) SSH anomalies (see Figure 3). Initial studies built upon techniques developed previously for turbulence simulations [15]. Since then, numerous variations of the approach used by Isern-Fontanet et al. [15] were introduced, e.g. [11, 4]. Chelton et al. [5] tracked eddies globally using a unified set of parameters. They also introduced the notion of eddy non-linearity (the ratio of rotational and transitional speeds) to differentiate between eddies and Rossby waves. In the most comprehensive SSH-based eddy tracking study to date, CH11 identified eddies globally as closed contoured smoothed SSH anomalies using a nearest neighbor search. Recently, Faghmous et al. [10] proposed a spatio-temporal approach to monitoring eddies as an alternative to existing image-based approaches. The authors used the eddies' physical properties to sparsify the search space in time and subsequently searched for eddies spatially. A more detailed review of SSH-based eddy detection methods can be found in Appendix B of CH11.

Despite a large body of work, eddy detection algorithms continue to suffer from several limitations. First, watersurface property signatures such as surface temperature or color do not convey much information on the dynamic process of eddies [13]. Second, certain SSH-based methods such as those introduced by Chaigneau et al. [4] use derivatives of the SSH field, which amplify the noise in the SSH signal [7]. Connected component algorithms, such as CH11, tend to be highly parameterized and apply scale-dependent filters to the data to remove features larger (smaller) than a threshold as well as to remove seasonal and internal variability (see Appendix A in CH11 or online supplementary material of [6] for full details on data filtering.) Additionally, CH11 was



Figure 3. Global sea surface height (SSH) anomaly for the week of October 10 1997 from the AVISO dataset. Eddies can be observed globally as closed contoured negative (dark blue; for cyclonic) or positive (dark red; for anti-cyclonic) anomalies. Best seen in color.

unable to systematically separate groups of eddies that were merged together because of their close spatial proximity. The spatio-temporal method proposed by Faghmous et al. [10] still had several hard-coded parameters and did not track many eddy properties (such as radius and amplitude).



Figure 4. Schematic of an anti-cyclonic eddy that is embedded in a large scale background with a larger amplitude than the eddy. If we were to apply a threshold at h = 0 the eddy would be missed. This is motivation to use multiple threshold from h = -100cm to 100cm as suggested by CH11. Figure adapted from [7]

III. METHODS

To monitor global eddy activity we used the Version 3 dataset of the Archiving, Validation, and Interpretation of Satellite Oceanographic (AVISO) which contains 7-day averages of SSH on a 0.25° grid from October 1992 through January 2011¹. We tracked eddies globally as closed contour of SSH anomalies. This was done in two steps: first we identified features that displayed the spatial properties of an ocean eddy. This was accomplished by assigning binary values to the SSH data based on whether or not a varying threshold was exceeded, and subsequently identifying mesoscale connected component features. We then pruned the identified connected components based on other criteria that are physically consistent with eddies at a given latitude.

Given the large variations in SSH on a global scale, tracking eddies globally presents several challenges: first, SSH data is noisy. Second, until recently, nonlinear eddies were commonly confused with linear Rossby waves in satellite data [6]. Third, eddies can manifest themselves as local minima (maxima) embedded in a large-scale background of negative (positive) anomalies [5] (see Figure 4). Therefore, applying a single global threshold would wipe out many relevant features. Fourth, although eddies generally have an ellipse-like shape, the shape's manifestation in gridded SSH data differs based on latitude. This is because of the stretch deformation of projecting spherical coordinates into a twodimensional plane. As a result, one cannot restrict eddies by shape (e.g. circle, ellipse, etc.) Finally, eddy sizes vary by latitude, which makes having a global "acceptable" eddy size unfeasible [13].

Our algorithm addresses some of these challenges, while maintaining physical relevance:

• To address the challenge of identifying mesoscale eddies superimposed on features with larger amplitudes, we repeatedly threshold the data at regular 1cm intervals from -100cm to +100cm. At each threshold tr_i , we identify all connected components that have an SSH anomaly of at least tr_i . The algorithm then removes from consideration all pixels belonging to the identified connected component and tr_i is incremented. For identification of anticyclonic eddies, tr_i is initialized at -100cm and increased in 1cm steps to +100cm. Conversely, detection of cyclonic eddies is accomplished by decreasing tr_i from +100cm to -100cm. In this way, we identify the largest possible closed contour of an eddy. The gradual thresholding method was proposed by CH11 who also tested subcentimeter threshold increments but did not observe increased accuracy.

¹Available at http://www.aviso.oceanobs.com/es/data/products/ sea-surface-height-products/



Figure 5. Mean weekly cyclonic (top) and anticyclonic (bottom) eddies by latitude as detected by EddyScan for the October 1992 - January 2011 period.

- To address the eddies' varying size by latitude, we use a quadratic function based on theoretical [2, 18, 17] and empirical studies [5, 13, 7] to restrict a reasonable eddy radius based on latitude.
- If an eddy is larger than expected at the latitude (see previous point), there is a chance that two or more eddies were mistakenly merged together. For these "larger than normal" eddies, we apply a convex hull function to determine the size of the smallest convex set that contains all pixels comprising the eddy. If the area of the convex hull is much larger than that of the connected component, it is likely multiple merged eddies and the connected component is not labeled as an eddy. By discarding the connected component, it will remain in the group of pixels to be examined at later thresholds effectively increasing the chance that a higher threshold would eventually break the larger connected component into smaller features.

At a high level, our algorithm extracts candidate connected components from SSH data by gradually thresholding the data and finding connected component features at each threshold. For each connected component, we apply five criteria to determine that it is an eddy: (i) A minimum eddy size of 9 pixels; (ii) a maximum eddy size of 1000 pixels; (iii) a minimum amplitude of 1 cm; (iv) the connected component must contain at least a minimum/maximum and (v) each connected component must have a predefined convex hull ratio as a function of the latitude of the eddy. The first four conditions are similar to those proposed by CH11. The convexity criterion is to ensure that we select the minimal set of points that can form a coherent eddy, and thus avoid mistakenly grouping multiple eddies together. Once the eddies are detected, the pixels representing the eddy are removed from consideration for the next threshold level. Doing so ensures that the algorithm does not overcount eddies. Removing the pixels will not compromise the accuracy of the algorithm given that the first instance an eddy is detected will be at its most likely largest size as a function of the threshold. The main distinction between our implementation, EddyScan, and CH11 are two-fold: First, we use unfiltered data while CH11 pre-process the data. Second, to ensure the selection of compact rotating vortices, CH11 required that the maximum distance between any pairs of points within an eddy interior be less than a specified threshold, while EddyScan uses the convexity criterion to ensure compactness. The primary motivation to use convexity is to reduce the run time complexity of the algorithm from $O(N^2)$ to O(N). An examination of the advantages of using convexity over a predefined maximum distance can be found in the discussion section.

EddyScan's pseudo-code is listed in Algorithm 1. An open-source implementation in MATLAB is available at https://github.com/jfaghm/ClimateCode.git

IV. RESULTS

We tracked eddies globally in weekly unfiltered SSH data from October 1992 to January 2011 using the procedure described in the previous section. On a weekly average, there were 2100 cyclonic and 2077 anticyclonic eddies with a larger number of eddies in the Southern Hemisphere. The slight preference for cyclonic eddies is consistent with the findings in CH11², although our results are not fully comparable since they only report eddies with lifetimes of at least 4 weeks, while we do not track eddies across timeframes.

²Available at: http://cioss.coas.oregonstate.edu/eddies/nc_data.html



Figure 6. Aggregate counts for eddy centroids that were observed through each $1^{\circ} \times 1^{\circ}$ region over the October 1992 - January 2011 period as detected by CH11 (left) and EddyScan (right). These results show high eddy activity along the major currents such as the Gulf Stream (North Atlantic) and Kuroshio Current (North Pacific). The high eddy counts along continental and map edges is an artifact of edge effects in the data that we will address in future work. Note the difference in color scale between the the left ([0 - 200]) and right panel ([0 - 270]). This figure emphasizes the similarity in the spatial distribution between the two methods since their exact eddy counts are not comparable (see text). Best seen in color.

Algorithm 1: EddyScan: An automatic global eddy tracking algorithm

Input: SSH Output: E : global eddy list with corresponding pixels belonging to each eddy; A : the amplitude of each eddy; S : the surface area of each eddy For each timestep t_i : For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel $i \ (p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end end	
belonging to each eddy; A : the amplitude of each eddy; S : the surface area of each eddy For each timestep t_i : For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel i $(p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	Input: SSH
eddy; S : the surface area of each eddy For each timestep t_i : For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel $i (p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	Output : E : global eddy list with corresponding pixels
For each timestep t_i : For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel $i (p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	belonging to each eddy; A : the amplitude of each
For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel $i (p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	eddy; S : the surface area of each eddy
For each threshold $tr_i \in \{-100 : 100\}$: if SSH value at pixel $i (p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	For each timestep t_i :
if SSH value at pixel $i(p_i) < tr_i$: $p_i = 0$; else $p_i = 1$; Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	For each threshold $tr_i \in \{-100: 100\}$:
$p_i = 0;$ else $p_i = 1;$ Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	if SSH value at pixel $i(p_i) < tr_i$:
$p_i = 1;$ Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	
Identify all connected component objects left; For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	else
For each connected component CC_i : if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	$p_i = 1;$
if CC_i meets criteria listed in text label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	Identify all connected component objects left;
label all pixels $p_j \in CC_i$ as an eddy; remove p_j from data; end end	For each connected component CC_i :
remove p_j from data; end end	if CC_i meets criteria listed in text
remove p_j from data; end end	label all pixels $p_i \in CC_i$ as an eddy;
end	
	- 5
end	end
	end

Figure 5 shows the average latitude-based distribution of cyclonic and anticyclonic eddies. There is a significant difference in the total number of eddies between EddyScan and CH11 (not shown; but they report approximately 3000 eddies per frame to our 4200) most notably at high latitudes and along the equator. The most likely explanation is that the high-pass filtering, while it enhances certain features it also removes others especially at higher latitudes (near the equator) where eddies tend to be small (large). Also CH11's spatial domain spanned 80° N to 80° S while we searched for eddies from 90° N to 90° S.

Figure 6 shows the aggregated spatial distribution of eddies on a $1^{\circ} \times 1^{\circ}$ grid. Given that our algorithm detects more eddies globally than CH11, due mostly to the fact that CH11 only reports eddies that last four weeks or longer,

we used different color scales between the left ([0 - 200])and right panel ([0 - 270]). The eddy centroid distribution is similar to CH11 (right panel) where high density regions tend to be along currents (*i.e.* Gulf Stream (North Atlantic) and Kuroshio Current (North West Pacific)) and in open oceans.

V. DISCUSSION

Our method provides a significant advance to the state of the art by reducing the computational complexity of common connected component algorithms. However, our results depend on a few parameters - most notably data filtering (or lack thereof) and our convexity ratio parameter. Below we analyze the advantages and sensitivity of our results to such parameterazation.

A. Sensitivity of results to pre-processing

Spatial filtering is a commonly used technique in eddy detection [7]. Numerous studies employ expertly-designed filters to remove signals larger (smaller) than certain scales in addition to filtering seasonality and noise. While filtering has its benefits, the risk of removing important features in the signal is always present. The results presented in the previous section were produced without any preprocessing. To test the sensitivity of the results to filtering, we applied a high-pass filter to remove signals larger than 10° by latitude and 20° by longitude similar to CH11. Figure 7 shows the difference in global cyclonic eddies identified by our algorithm using unfiltered (top panel) and high-pass filtered data (bottom panel) for a single time-step. In the filtered data case, we find significantly more eddies at the equator and they tend to be much smaller than expected (eddies have a radius of 200km near the equator [13]). We suspect that these "ghost eddies" are the result of residue noise left after filtering out large features at the equator. Moreover, filtering changes the contours of the data, which means the connected components that result from thresholding a filtered dataset will be geometrically different than those based on the raw data. Because eddy measurements are made on the geometry of the connected component (e.g. surface area) as well as the underlying physical data (amplitude), filtering becomes a source of measurement error.



Figure 7. An example of the effect high-pass filtering has on EddyScan's output for single time-step. **Top:** Global cyclonic eddies as detected by EddyScan using unfiltered SSH data. The data are on a grayscale for easier visualization of the detected eddies. **Bottom:** Same as top except for high-pass filtered data. High-pass filtering might introduce ghost eddies along the equator as well as noise in eddy characteristics (radius and amplitude).

B. Advantage of using a convexity metric

As described earlier, binarization of the SSH data during connected component analysis often leads to merging multiple eddies into a single connected component. It is thus necessary to discern those coherent structures representing a single eddy from those comprised of multiple features merged together. Although the vortical character of eddies makes their SSH contours theoretically circular, multiple externalities prohibit us from simply imposing a circularity criterion on connected components. The presence of noise and SSH variability are examples of such factors. In addition, the projection of the SSH data into a two-dimensional grid induces substantial geometric distortions, and re-projecting connected components back onto a spherical surface to analyze their geometry is computationally expensive. CH11 addresses this problem by restricting any two pixels of a connected component to be within a maximum distance

101

based on latitude. This criterion successfully removes connected components that are particularly eccentric or large, but does not address the issue of merging smaller eddies together (see Appendix B in [7]). More importantly, the maximum distance criterion is extremely inefficient as it requires comparing every pair of pixels within a connected component – a function that grows quadratically with the number of pixels in the connected component. Instead of restricting the maximum distance between any two pixels, we proposed to monitor a feature's convexity to ensure multiple small eddies are not labeled as a single larger eddy. Our convexity criterion addresses the two issues of long incoherent features and breaking merged features while being computationally inexpensive. Figure 8 shows how without the convexity criterion (top panel) our algorithm groups several smaller eddies into a single eddy, whereas CH11's maximum distance criterion successfully breaks the large eddy into smaller ones. As expected, when we apply our convexity criterion (bottom panel) we are able to identify similar eddies to CH11.



Figure 8. An example of the effect our convex hull criterion has on breaking large blobs. **Top**: Without the convexity requirement, EddyScan (red circles) labels a very large blob as a single eddy (large purple ellipse), whereas CH11 (yellow crosses) labels it as four different eddies. **Bottom**: EddyScan with the additional convex hull constraint successfully breaks the large blob into 4 distinct eddies similar to CH11. The convex hull criterion is more efficient than the maximal distance criterion used by CH11. Best seen in color.

There are other instances, however, when the maximum distance criterion is unable to avoid merging several smaller eddies together. Figure 9 shows an example where the minimal distance between any pair of pixels in the blob is met despite there being several eddies. As a result CH11 (yellow cross) labels the entire feature as a single eddy. EddyScan, however, is able to break the large blob into coherent small eddies.



Figure 9. An example of when CH11's maximum distance criterion is met, yet the large feature is in fact several eddies merged together. **Top:** a zoomed-in view on SSH anomalies in the Southern Hemisphere showing at least four coherent structures with positive SSH anomalies. **Bottom:** CH11 (yellow cross) identifies a single eddy in the region, while our convexity parameter allows EddyScan to successfully break the larger blob into four smaller eddies. Best seen in grayscale to improve visibility of the identified eddies. Best seen in color.

C. Sensitivity of results to convexity parameter

Our results depend on the convexity ratio parameter (the ratio of the feature's area to its convex hull area). A ratio of one indicates that the identified feature is perfectly convex. A low ratio indicates a less coherent feature. We tested EddyScan's performance using a variety of convexity ratios and settled on 0.85 because it gave the best balance between cohesiveness and accuracy. Figure 10 shows the difference in EddyScan's output based on the choice of convexity ratio. If the convexity ratio is set too low (top panel), large blobs are labeled as eddies throughout the globe dramatically reducing the global eddy count (e.g. see lower right corner of the top panel in Figure 10). If the convexity ratio is set too high (bottom panel), the global eddy count is not severely affected (the count increased by less than 1% globally) but the mean amplitude and radius are. In the bottom panel of Figure 10 the SSH anomalies are in grayscale for clarity, it easy to see in its attempt to finding the most compact features possible, the contours of many cyclonic eddies are much smaller than expected as shown by the white contours around many eddies (the more accurate labeling would encompass all positive anomalies or white pixels within the eddy's perimeter).

D. Complexity Analysis

To ensure that only compact rotating vortices are labeled as eddies, CH11 imposes a restriction that the distance between any two pixels in a connected component must be



Figure 10. EddyScan's sensitivity to the choice of convexity parameter. **Top:** When the minimal convexity ratio is set too low (0.5), large incoherent blobs are labeled as eddies significantly affecting the global eddy count. **Bottom:** When the minimum convexity ratio is set to one, identified eddies are much smaller than their actual size because the algorithm picks the most compact features possible. Blue area characterize land. Best seen in color.

less than a latitude-specific maximum distance. Computing the distance between every pair of pixels in a set of size Nwould result in a run-time complexity of $O(N^2)$. This operation is prohibitively expensive especially with the additional overhead to convert the distances from pixel/Euclidean to Great Circle distance. In contrast, our convex hull criterion has a runtime of O(N) given that the pixels are already sorted lexicographically (by row) [1], resulting in significant speedup and increased accuracy (by successfully separating merged eddies).

VI. CONCLUSION AND FUTURE WORK

We presented EddyScan: an automated, accurate, and scalable eddy detection algorithm for SSH data. EddyScan improves on the state of the art by not preprocessing the data, effectively breaking up merged eddies, and running in a fraction of the time required by traditional eddy detection methods – a significant improvement given the expected dramatic increase in earth science data.

Our method suffers from a few limitations: First, it does not track eddies across time, making it susceptible to noise and linear Rossby waves. Second, we do not account for edge cases and coastal regions where over-counting tends to be high. Finally, imposing a minimal eddy size of 9 pixels makes it impossible to identify smaller eddies that are common at higher latitudes.

Our methodology can be further improved by applying a spatio-temporal context to eddy monitoring. As shown in Faghmous et al. [10] monitoring the temporal profile of SSH data can decrease false discovery rates by ensuring eddies persist over time. CH11 tracks eddies across time and only keeps eddies that persist for at least 4 weeks. This approach is only feasible by restricting the minimal eddy size to 9 pixels; otherwise one cannot resolve eddies across frames due to the numerous small eddies in a neighborhood. An alternative approach would be to relax the minimal eddy size restriction and use the temporal profile of each "eddy" to ensure that it is not spurious. Another improvement would be to incorporate space and time information simultaneously to label eddies, in contrast to Faghmous et al. [10] where the spatio-temporal analysis was done orthogonally (first in time, then in space). Network-based approaches such as those presented in [16] may be useful in this context. Finally, this approach could be generalized to employ a systematic spatio-temporal analysis to monitor persistent features within a continuous and noisy field.

ACKNOWLEDGMENTS

This work was funded by an NSF Graduate Research Fellowship, an NSF Nordic Research Opportunity Grant, the Norwegian National Research Council, the Planetary Skin Institute, NSF Grant IIS-0905581, and an NSF Expeditions in Computing Grant (IIS-1029711). Access to computing resources was provided by the University of Minnesota Supercomputing Institute.

REFERENCES

- [1] A. Andrew. Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219, 1979.
- [2] X. Capet, P. Klein, B. Hua, G. Lapeyre, and J. McWilliams. Surface kinetic energy transfer in surface quasi-geostrophic flows. *Journal of Fluid Mechanics*, 604(1):165–174, 2008.
- [3] M. Castellani. Identification of eddies from sea surface temperature maps with neural networks. *International journal of remote sensing*, 27(8):1601–1618, 2006.
- [4] A. Chaigneau, A. Gizolme, and C. Grados. Mesoscale eddies off peru in altimeter records: Identification algorithms and eddy spatio-temporal patterns. *Progress* in Oceanography, 79(2-4):106–119, 2008.
- [5] D. Chelton, M. Schlax, R. Samelson, and R. de Szoeke. Global observations of large oceanic eddies. *Geophysical Research Letters*, 34:L15606, 2007.
- [6] D. Chelton, P. Gaube, M. Schlax, J. Early, and R. Samelson. The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. *Science*, 334(6054):328–332, 2011.

- [7] D. Chelton, M. Schlax, and R. Samelson. Global observations of nonlinear mesoscale eddies. *Progress* in Oceanography, 2011.
- [8] D. D'Alimonte. Detection of mesoscale eddy-related structures through iso-sst patterns. *Geoscience and Remote Sensing Letters*, *IEEE*, 6(2):189–193, 2009.
- [9] C. Dong, F. Nencioli, Y. Liu, and J. McWilliams. An automated approach to detect oceanic eddies from satellite remotely sensed sea surface temperature data. *Geoscience and Remote Sensing Letters, IEEE*, (99): 1–5, 2011.
- [10] J. Faghmous, Y. Chamber, F. Vikebø, S. Boriah, S. Liess, M. d.S. Mesquita, and V. Kumar. A novel and scalable spatio-temporal technique for ocean eddy monitoring. In *Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*, 2012.
- [11] F. Fang and R. Morrow. Evolution, movement and decay of warm-core leeuwin current eddies. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50 (12-13):2245–2261, 2003.
- [12] A. Fernandes and S. Nascimento. Automatic water eddy detection in sst maps using random ellipse fitting and vectorial fields for image segmentation. In *Discovery Science*, pages 77–88. Springer, 2006.
- [13] L. Fu, D. Chelton, P. Le Traon, and R. Morrow. Eddy dynamics from satellite altimetry. *Oceanography*, 23 (4):14–25, 2010.
- [14] R. Holyer and S. Peckinpaugh. Edge detection applied to satellite imagery of the oceans. *Geoscience and Remote Sensing, IEEE Transactions on*, 27(1):46–56, 1989.
- [15] J. Isern-Fontanet, E. García-Ladona, and J. Font. Identification of marine eddies from altimetric maps. *Journal of Atmospheric and Oceanic Technology*, 20(5): 772–778, 2003.
- [16] M. Kim and J. Han. A particle-and-density based evolutionary clustering method for dynamic networks. *Proceedings of the VLDB Endowment*, 2(1):622–633, 2009.
- [17] A. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 434 (1890):9–13, 1991.
- [18] A. N. Kolmogorov. On degeneration of isotropic turbulence in an incompressible viscous liquid. In *Dokl. Akad. Nauk SSSR*, volume 31, pages 538–540, 1941.
- [19] P. Richardson. Eddy kinetic energy in the north atlantic from surface drifters. *Journal of Geophysical Research*, 88(C7):4355–4367, 1983.
- [20] K. Wyrtki, L. Magaard, and J. Hager. Eddy energy in the oceans. *Journal of Geophysical Research*, 81(15): 2641–2646, 1976.