Adaptive Extraction and Quantification of Geophysical Vortices

Sean Williams, Mark Petersen, Peer-Timo Bremer, Matthew Hecht, Valerio Pascucci,
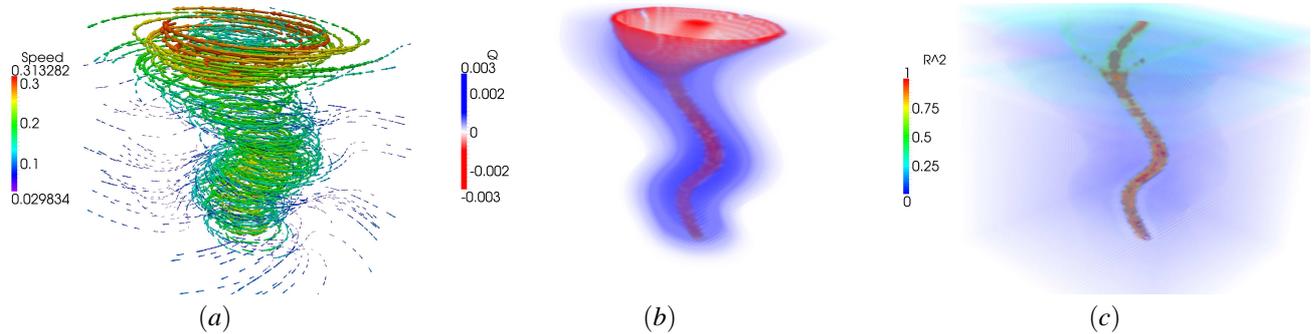James Ahrens, Mario Hlawitschka, Bernd Hamann

Fig. 1. Three views of the tornado data set. (a) Arrows pointing along streamlines of velocity reveal the basic structure of the tornado. However, we are primarily interested in its vortex core, and specifically, where the vortex core resembles a Gaussian vortex. (b) The *Q-criterion* for the tornado data set, as expected, shows a vortical interior surrounded by a strain cell, where the spinning air of the tornado shears against the calmer air outside. The vortex core is brought out in the strong negative domain of Q, but an additional vortical funnel appears in the upper half of the tornado. (c) For our method, we compute the similarity of different Q thresholds to an idealized Gaussian vortex. The similarity holds well on the interior, where the flow of air resembles an idealized vortex, but quickly decays outside the core. Additionally, the funnel on top is not modeled well by Gaussian vorticity (except near the interface between the funnel and the main core), so it also has a low fit value and is excluded.

Abstract—We consider the problem of extracting discrete two-dimensional vortices from a turbulent flow. In our approach we use a reference model describing the expected physics and geometry of an idealized vortex. The model allows us to derive a novel correlation between the size of the vortex and its strength, measured as the square of its strain minus the square of its vorticity. For vortex detection in real models we use the strength parameter to locate potential vortex cores, then measure the similarity of our ideal analytical vortex and the real vortex core for different strength thresholds. This approach provides a metric for how well a vortex core is modeled by an ideal vortex. Moreover, this provides insight into the problem of choosing the thresholds that identify a vortex. By selecting a target coefficient of determination (i.e., statistical confidence), we determine on a per-vortex basis what threshold of the strength parameter would be required to extract that vortex at the chosen confidence. We validate our approach on real data from a global ocean simulation and derive from it a map of expected vortex strengths over the global ocean.

Index Terms—Vortex extraction, feature extraction, statistical data analysis

---------------------------◆---------------------------

1 INTRODUCTION

Vorticity is defined as the magnitude and rotational axis of vortical behavior in a turbulent flow. In real flow, however, noise creates small

- *Sean Williams is with the Institute for Data Analysis and Visualization at University of California, Davis and Los Alamos National Laboratory, E-mail: sjwill@ucdavis.edu.*
- *Mark Petersen is with Los Alamos National Laboratory, E-mail: mpetersen@lanl.gov.*
- *Peer-Timo Bremer is with Lawrence-Livermore National Laboratory, E-mail: bremer5@llnl.gov.*
- *Matthew Hecht is with Los Alamos National Laboratory, E-mail: mhecht@lanl.gov.*
- *Valerio Pascucci is with the Scientific Computing and Imaging Institute at University of Utah, E-mail: pascucci@sci.utah.edu.*
- *James Ahrens is with Los Alamos National Laboratory, E-mail: ahrens@lanl.gov.*
- *Mario Hlawitschka is with the Universität Leipzig, Germany, E-mail: hlawitschka@informatik.uni-leipzig.de*
- *Bernd Hamann is with the Institute for Data Analysis and Visualization at University of California, Davis, E-mail: hamann@cs.ucdavis.edu.*

perturbations in the vorticity and other derived fields that must be separated from actual vortices. The standard solution to this problem is to only consider vorticity with magnitude above a specified threshold. However, it is not obvious how to choose such a threshold automatically nor whether a globally valid threshold even exits. In practice, one typically tries multiple thresholds until one produces "reasonable" results. Nevertheless, more sophisticated methods do exist, such as choosing maximum contours of λ_2 [14], but theoretical justifications for maximum contours decisions are still lacking.

For this research, we begin instead from a fluid dynamics definition of an ideal vortex in a two-dimensional domain. This analysis also allows axis-aligned three-dimensional vortices, i.e., vortices in which the vorticity vectors inside the vortex core are fairly well-aligned. Notably, this includes most vortices of interest in the oceanic and atmospheric sciences. The fluid dynamics definition we use models vorticity as a Gaussian, from which we derive several other properties, such as the strain tensor. We describe a nearly-linear relationship between *Q-criterion* thresholds and the surface area of the feature entirely below each threshold. Furthermore, we show that this relationship only holds while the vortex is dominated by Gaussian vorticity. Once a threshold begins including the background flow, the relationship no longer holds—providing information about what threshold of Q is appropriate for that vortex.

We specify the similarity of an actual vortex to an ideal vortex as a coefficient of determination of a linear fit, so rather than thresholds
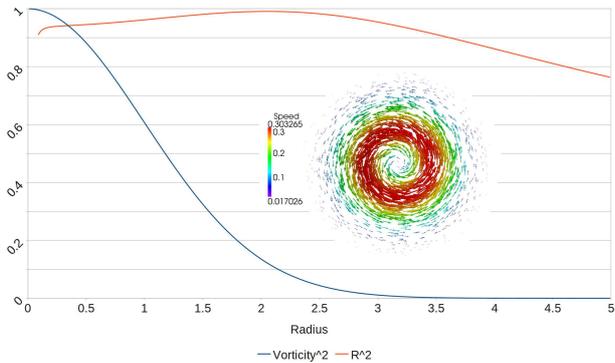
Fig. 2. The Gaussian function used for idealized vorticity, $\omega(r) = \exp\left(\frac{-r^2}{2}\right)$, plotted as a function of r, the distance from the center of the vortex. A corresponding Gaussian vortex is inset, with arrows pointing along streamlines of velocity and colored by speed. Despite being a Gaussian, for the regions where vorticity dominates this idealized vortex—out to a radius of about 1.5—the vorticity function is very close to linear. To quantify this, for each value of radius, we plot the coefficient of determination (R^2) of vorticity to a least-squares linear fit of vorticity out to that radius. This fit indicates that the Gaussian is at least 90% similar to a linear function to a radius of about 3.5, which is well past the edge of this vortex. Thus, we assume for this research that, for vortex interiors, vorticity can be approximated as a linear function of radius.
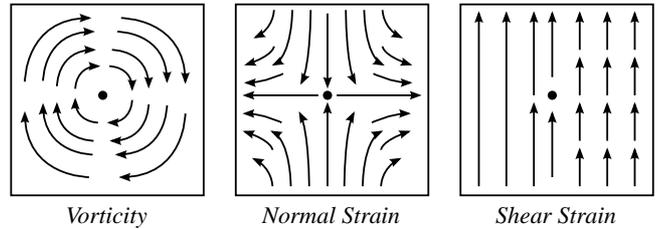

Fig. 3. Schematic views of the three components that make up the *Q-criterion*. Vorticity is caused by flow circling an axis, normal strain is caused by two flows colliding and pulling apart, and shear strain is caused by a lateral change in velocity.

being chosen as semi-arbitrary values of vorticity or Q, we instead use statistical confidence levels. Another way of looking at this result is to compare the Q thresholds that produce vortices meeting a particular confidence level. Since the vortices are rated based on how well they structurally conform to an ideal vortex, deriving Q thresholds from a confidence level produces an approximation of the stability of vortices throughout a data set.

To take advantage of this additional information, we apply our method to a simulated global ocean data set to begin quantifying the structural differences between vortices in a variety of currents around the world. Comparing these results against what is observationally known about the real ocean lends further validity to our method, and opens a new avenue of study for vortices in applied scientific domains.

The primary contributions of this paper are:

- An algorithm for detecting vortices based on a new physics-based characterization

- A local criterion to find a Q threshold for each vortex based on statistical confidence

- An empirical study to validate the procedure both on classical visualization data, and on global ocean simulation data showing different characteristics in different oceans

2 RELATED WORK

Extracting and visualizing turbulence in vector fields has long been of interest to the visualization community [8]. Methods generally focus on either extracting specific structures (e.g., vortices) and drawing a bounding volume [15], or extracting the overall topology and visualizing it through glyphs or other proxies [18]. For extracting vortex-like structures in particular, popular methods include finding regions of high vorticity [19], streamline geometry [12] or, if the data are available, by looking for regions of low pressure at the center of a vortex [1].

For two-dimensional or axis-aligned vortices, one can also look for circular behavior in the velocity field directly. Jiang et al analyze the problem topologically [5], by looking for kernels in which each vector in the kernel points in a unique direction range. Similarly, Sood et al [17] identify vortex centers in the ocean by passing a 5×5 kernel of the angles between an east-pointing vector and the tangents of a circle

centered in the kernel to find circular flows, then fitting an ellipse over the entire vortex.

Several techniques have also been developed to further refine analysis of the vorticity field. We begin from the *Q-criterion* [4], which highlights areas where vorticity dominates strain. The *Q-criterion* will be explained in greater detail in the next section. Our choice of this criterion was initially motivated by its use in oceanography, where the *Q-criterion* is known as the *Okubo-Weiss criterion* [11]. Also of note is the λ_2-*criterion* [6], that uses the second eigenvalue of the strain tensor times the sum of the squares of strain and vorticity to determine whether vorticity or strain are dominant in a region. We chose the *Q-criterion* over λ_2 because it is the standard vortex criterion within the oceanographic community.

3 VORTEX DETECTION

We begin with the assumption that vorticity in an ideal vortex can be modeled as a Gaussian in the radial direction. This is an established model in the fluid dynamics community [7]. Further, in the ocean it has been found to be a reasonable model both analytically [3] and observationally [13], and in the atmosphere, this idealization has been used to approximate the vorticity profile of hurricanes [9]. The general form of this Gaussian, for vertical and horizontal scaling coefficients c_1 and c_2, is:

$$\omega(r) = c_1 \exp\left(\frac{-r^2}{2c_2^2}\right) \tag{1}$$

To simplify this analysis, we claim, based on the graph in Figure 2 that, for vortex interiors, vorticity can be approximated as a linear function of radius. If we consider the vortex interior to be the negative domain of Q, then by figure 4a, this extends to a radius of about 1.5. At this distance, the linear fit is still extremely high. Now consider the definition of the *Q-criterion*, $Q = \varepsilon^2 - \omega^2$. Here, ε is the magnitude of the strain tensor, and ω is the magnitude of vorticity. Schematics of flows typical of high vorticity and the two components that contribute to strain are shown in Figure 3. The chief assumption of the *Q-criterion* is that, inside a vortex, strain is low and vorticity is high, while the boundary of a vortex is indicated by high strain and low vorticity. Since vorticity dominates strain on vortex interiors, we conclude that $Q \approx \omega^2 \approx r^2$. Furthermore, we can see from Figure 4b that the dominance of vorticity inside the vortex (again, for the negative domain of Q) is such that this does in fact create a linear relationship between Q and r^2 (since r^2 is proportional to area).

We also assume that an ideal vortex is circular. As a result, the area of the vortex will grow quadratically with radius. If the vortex is elliptical, it will grow as the product of half its major and minor axes. This still results in quadratic growth: given an ellipse with major and minor axis lengths a and b, respectively, and constant eccentricity e, then $b = ae$, so the area grows as $\frac{1}{4}ea^2$.

Hence, for a vortex that is governed by Gaussian vorticity in its core and has a core that is either circular or a well-behaved ellipse, the *Q-criterion* and its area will grow linearly with respect to each other. This result is important for three main reasons.

First, how well a vortex conforms to these rules can be quantified as a statistical measure. We are concerned with how closely Q and area, both as functions of radius, grow linearly with respect to each other, which can be done by computing the coefficient of determination (R^2) between Q versus A and a linear least-squares fit of the same. This changes the art of choosing a threshold from one based on *Q-criterion* values that are difficult to interpret to a straightforward choice of how confidently the data fit an ideal model.

Second, a statistical confidence level still implies a threshold on Q, but on a per-vortex basis. This provides a consistent way of comparing the relative stabilities of different vortices, since the higher the allowable value of Q, the longer Gaussian vorticity remains the dominant effect in the flow. We specifically apply this to data from a global ocean simulation to show the stabilities of vortices in different major currents around the world.

Finally, this relationship provides an easy way to separate Gaussian-like vortices from other high-vorticity features. In the Figure 6a, sharp turns in strong currents like the Gulf Stream generate comma-shaped spikes in vorticity that, on visual inspection, are clearly not vortices. As a result of these high-vorticity meanders in the strong currents being neither Gaussian vortices nor represented by a circle or ellipse, almost all of them fail to hold this linear relationship. The few that hold this relationship are extremely special cases, and will be discussed in greater detail later.

For the next three sections, we will demonstrate this relationship empirically, first with an idealized analytical vortex, then a simple tornado data set, and finally with data from a complex global ocean simulation.

4 IDEAL VORTICES

According to [7], a planar slice of an idealized axis-aligned vortex can be described in polar coordinates in terms of vorticity ω and strain ε:

$$\omega = \left[0, 0, \left(\frac{1}{r}\frac{\partial r v_\theta}{\partial r} - \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right)\right]^T \qquad (2)$$

$$\varepsilon = \begin{bmatrix} \frac{\partial v_r}{\partial r} & \frac{1}{2}\left(r\frac{\partial (v_\theta/r)}{\partial r} + \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right) \\ \frac{1}{2}\left(r\frac{\partial (v_\theta/r)}{\partial r} + \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right) & \frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}. \end{bmatrix} \qquad (3)$$

Here, v_r and v_θ are the radial and azimuthal velocities, respectively. If this vortex is placed at the origin, and taking the idealized case that $v_r = 0$ and $\frac{\partial}{\partial \theta} = 0$, the *Q-criterion* is then:

$$Q = \|\varepsilon\|^2 - \|\omega\|^2 \qquad (4)$$

$$= \left(r\frac{\partial (v_\theta/r)}{\partial r}\right)^2 - \left(\frac{1}{r}\frac{\partial r v_\theta}{\partial r}\right)^2 \qquad (5)$$

$$= -4\frac{v_\theta}{r}\frac{\partial v_\theta}{\partial r}. \qquad (6)$$

One may compute a profile of velocity, strain, and the *Q-criterion* for this vortex as a function of radius, as shown in Figure 4a. In the vortex core, the vortex is dominated by vorticity, so it initially increases with radius as a particle must travel faster to make a full revolution of the vortex center in the same amount of time. Further from the center, however, flow begins to slow down as vorticity is less dominant and, to compensate, the flow begins shearing. This causes strain to increase while vorticity decreases. The combination of these effects causes the characteristic behavior of the *Q-criterion*: Q is negative inside the vortex core, where vorticity dominates, but increases until becoming positive, where strain dominates. Finally, all parameters converge to zero as the vortex loses all influence over the flow.

Figure 4b then demonstrates the near-linear relationship between Q and area. Q is plotted with respect to area, then broken into three regions, colored red, green, and blue. For each region, a linear fit is taken between Q and area for the points in that region and all points in all previous regions, so the green fit is for data in both the red and
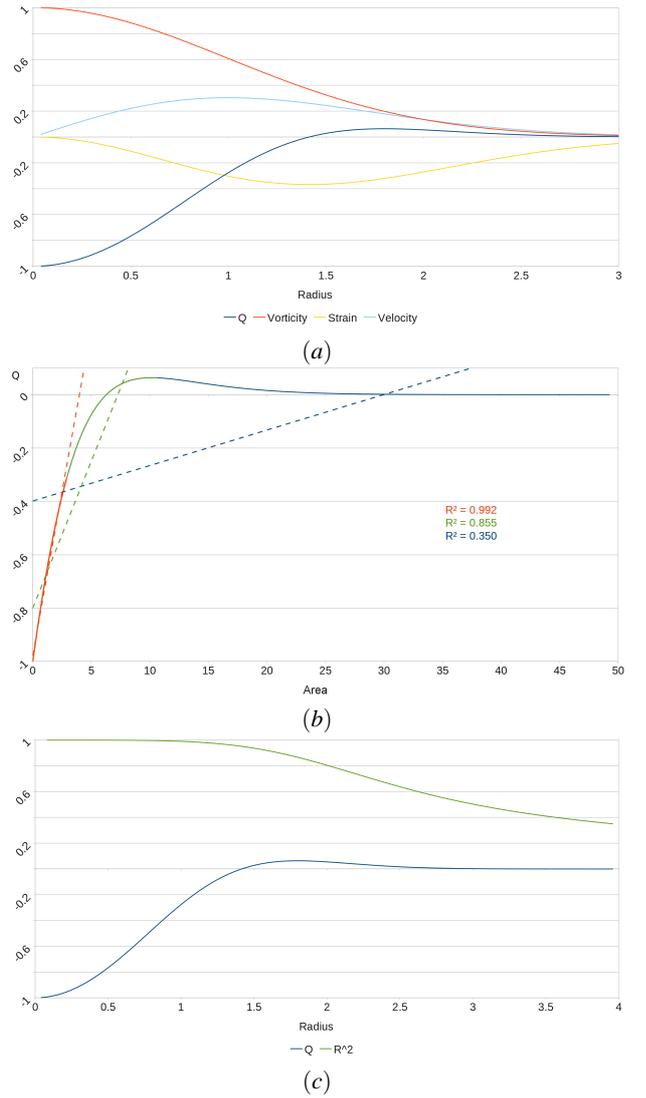

(a)

(b)

(c)

Fig. 4. The properties of an idealized vortex, with vorticity modeled as a Gaussian in radius. (*a*) Azimuthal velocity, strain, vorticity, and *Q-criterion* as functions of radius. Vorticity is modeled as a Gaussian, from which the other properties were derived. The interaction of vorticity and strain causes the characteristic behavior of the *Q-criterion*: Q is negative inside the vortex core, where vorticity dominates, but increases until becoming positive, where strain dominates. (*b*) The *Q-criterion* as a function of area, along with stippled lines showing a least squares linear fit of each component of the curve from 0 until the end of that color. Inside the vortex core (red), the two grow nearly linearly. Taking a linear fit of the red section yields a coefficient of determination of 0.99. If the area under consideration is expanded outside the vortex core, this nearly-linear relationship breaks down (green), and gets worse as more of the background flow is added (blue). (*c*) The *Q-criterion* and the coefficient of determination (R^2) as a function of radius. The R^2 value plotted in this graph is the quality of a linear fit of area versus Q if the corresponding Q value were used as a threshold, i.e., an R^2 value here is the quality of fit for all points to its left. As expected, in the vorticity-dominated region, area and Q conform well to a linear relationship, while outside, the coefficient of determination drops precipitously.

green regions. The fit is initially quite good, at 99% in the red region (to a radius of about 0.96), but after Q passes 0 and ceases being dominated by vorticity, the relationship breaks down. By the time one considers both the red and green regions (at a radius of about 1.8) the fit is already down to 88%, while the fit over the entire data set is down at 63%. Taking instead a "rolling fit"—computing the fit at each resolved radius—we get Figure 4c, in which the fit is shown to be quite high for radii still within the vortex core, but dropping off quickly once the radius extends outside the core.

We now present an algorithm to make use of this relationship, assuming a data set with a single vortex and no local minima other than the global minimum (i.e., the data decrease monotonically), as a list of Q values for all points, here just called Q. We take an input r, the resolution of Q values to be checked, and an input c, the desired confidence level. The algorithm returns the maximum Q threshold that results in the desired confidence, or $\perp$ if that confidence is never reached. The assumptions (one vortex and monotonic values of Q) can be discarded by making the algorithm instead breadth-first search from minima of Q, and preserving search visitation marks between vortices, so that a point is never visited again, even on future searches. This way if multiple minima are contained in a single vortex, they will all be visited (and marked to not be considered in the future), so that the vortex will not be counted multiple times.

Input: $r > 0, c \in [0, 1]$
  $Q_h, A_h \leftarrow \emptyset$ {Lists to store a history of (Q, A) pairs}
  $q \leftarrow \min(Q)$ {Iterated value, from min(Q) to 0}
  $s \leftarrow -q/r$ {How much to step Q to get r iterations}
  $c_q \leftarrow \perp$ {Highest Q value meeting c% confidence}
  while $q \leq 0$ **do**
    $a \leftarrow |\{q_c \in Q, q_c \leq q\}|$ {Area: #points below current threshold}
    $Q_h \leftarrow Q_h \bigcup \{q\}$
    $A_h \leftarrow A_h \bigcup \{a\}$
    $r \leftarrow \text{cod}(Q_h, A_h)$ {Coefficient of determination of a linear fit}
    if $r \geq c$ **then**
      $c_q \leftarrow q$
    end if
    $q \leftarrow q + s$
  end while
  return c_q

The function "cod" returns the coefficient of determination of a linear fit between the two input sets. This is computed first by finding a linear least squares fit of the data, which returns a fitting function, i.e., it returns coefficients c_0 and c_1 such that $A_a = c_0 + c_1 Q_a$. Thus, for any i, the predicted area is $f_i = c_0 + c_1 q_i$. Additionally, we need $\bar{A}$, the average of all $a_i \in A_h$. Then the coefficient of determination is:

$$R^2 = 1 - \frac{\sum_i (a_i - f_i)^2}{\sum_i (a_i - \bar{A})^2} \qquad (7)$$

5 Simple Vortices: The Tornado

In order to demonstrate the effectiveness of this method on real data, we begin with a simulation of a tornado [2]. The data are on a 64^3 regular grid, containing a three-vector of velocity at every point. The data contain only a single tornado, with surrounding air, and the core of the tornado roughly follows the z-axis. For a first look at the tornado data set, we show arrows pointing along streamlines of velocity, colored and sized based on speed (Figure 1a). Notably, the tornado is faster and wider at the top than near the bottom, but primarily only in an outer envelope.

Computing the *Q-criterion* for the tornado data set, in Figure 1b, shows a strong vortex core up the center surrounded by a large strain cell. However, the Q field also shows a strong cone-shaped vortical funnel near the top that is not part of the main core of the tornado. Additionally, as before, it is unclear which threshold of Q to use to extract the relevant extent of the vortex core: Q ranges from -0.2848 to 0.003329, while the color scale in that visualization (from -0.003 to 0.003) was chosen to elucidate the structure of the data without creating color scale discretization problems, i.e., from putting less than
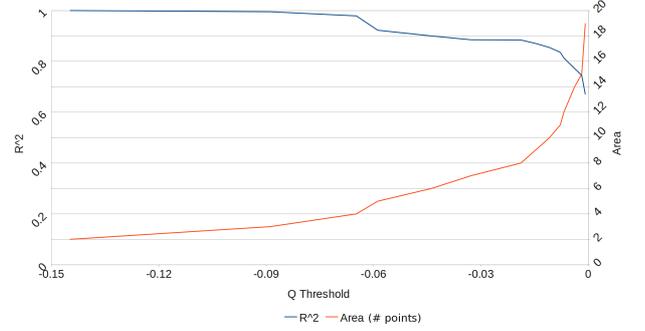

Fig. 5. Area and R^2 plotted against Q for a middle z-slice of the tornado data. Area (computed as the number of points in the slice below a given Q threshold) initially grows quite linearly. The growth of the vortex begins speeding up after Q thresholds of -0.06, while after -0.02 the growth is dramatically faster. For this slice, then, the linearity of this vortex crosses a confidence of 90% with a Q threshold of -0.04. If this graph continued into the positive domain, at a threshold of 0.01 there would be almost 4000 points in the vortex and a confidence of just 8%.

2% of the data range (the positive domain of Q) into half the color scale.

To analyze the tornado data set for Gaussian vorticity behavior, we treat the data as stacks of two-dimensional slices, though allowing them to still have a $\hat{k}$ (depth) velocity component. To visualize the effect of the fitting criterion, for each slice we take discrete iterative thresholds of Q and compute R^2 for that threshold's linear fit against area, computed as the number of points in that slice below the threshold. This is repeated on intervals of Q to a maximum of 0.002, versus the data set maximum of about 0.003. The resulting R^2 value for each interval is stored at all points in that interval. In other words, this R^2 value is assigned to all points that are at or below the current threshold, but above the previous threshold. This R^2 field is shown volume-rendered in Figure 1c, with a transfer function specifically highlighting points with high R^2. For most of the tornado data set, this extracts roughly the same vortex core as the raw Q parameter, with two notable exceptions. First, the conic funnel near the top is nearly gone, showing that only the center of the tornado behaves like a Gaussian vortex. Second, the transfer function can now be set much more meaningfully and without any tuning: the transfer function of R^2 ramps opacity linearly from a low level at 0.75 to opaque 1.0, while all values below 0.75 are kept at very low opacity.

To address the question of what threshold to choose for Q, we also plot area and R^2 against Q threshold for slice 31 in Figure 5. For very low values of Q, the linear fit matches the data quite well, with the R^2 value staying above 95% up to Q thresholds of about -0.06. R^2 crosses 90% confidence at about -0.04, and begins falling rapidly after thresholds of about -0.02. Consistent with this, area as a function of Q threshold rises in a near-linear fashion until thresholds of about -0.06, then bumps up and climbs at a slightly higher rate. After thresholds of -0.02, area climbs quite rapidly. Once the Q thresholds near 0, the background flow enters the threshold, and the area jumps into the thousands of points.

6 Complex Vortices: Global Ocean Simulation

Now that our method has been established for a reference data set, we use this technique to extract vortices from a global ocean simulation while gathering statistical information about the behavior of the vortices over time. To this end, we employ Los Alamos National Laboratory's Parallel Ocean Program (POP) [16]. The runs of POP we use employ a $3600 \times 2400 \times 42$ rectilinear grid, providing a horizontal resolution of about $\frac{1}{10}°$, or approximately 10 km at the equator. Depths span a range of 5000 m, with vertical grid cell thickness ranging from 10 m near the surface to 250 m in the deep ocean. We only require velocity information to compute Q. In the ocean, horizontal
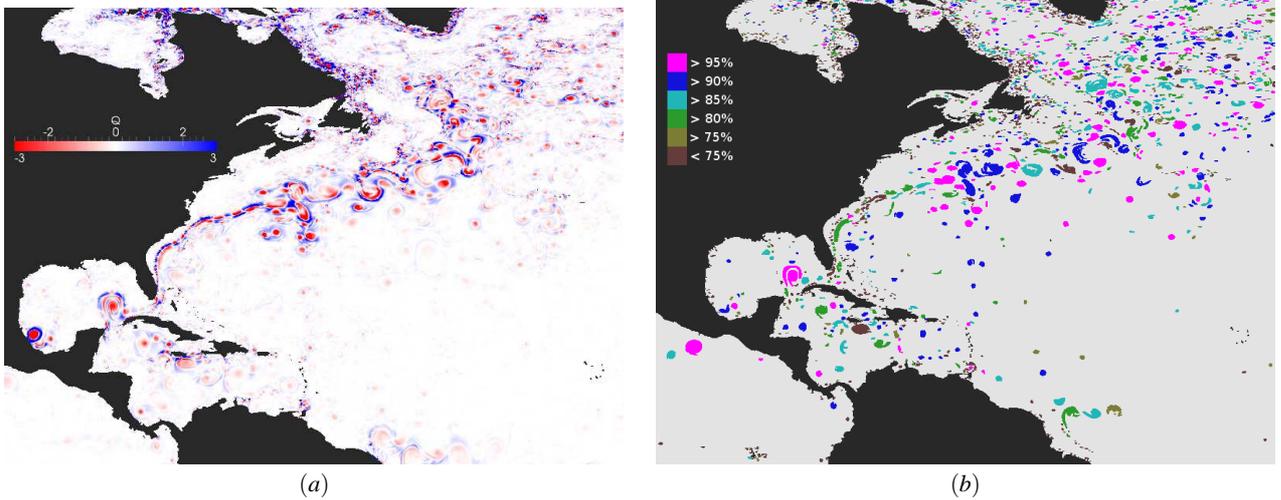
Fig. 6. Here we compare how different confidence levels affect vortex identification in the region of the Gulf Stream. (*a*) The *Q-criterion* at the surface of the Atlantic Ocean shows both a rich population of vortices, and regions of high vorticity caused by the meandering path of the Gulf Stream. For the purpose of visualization, this field is extremely dense and difficult to understand, and for the purpose of analysis, the presence of so many false positives (i.e., meanders and boundary shear) will corrupt any data about the vortices. (*b*) The confidence levels associated with each high-vorticity feature, discretized to 5%. Features above 95% (in bright magenta) very closely resemble what we would expect a vortex to look like, while those below 85% (in dark red, yellow, and green) almost all look dubious. The quality of features between 85% and 95% (blue and cyan) is mixed, with several good features in that range, but some dubious features as well. We choose a 90% threshold for future analysis based on a qualitative trade-off between false positives and false negatives, and to ensure that vortices could survive minor distortions for the purposes of vortex tracking.

velocities (i.e., in the *xy*-plane) are about three orders of magnitude greater than vertical velocities, so we only consider longitudinal and latitudinal components of velocity.

Oceanic vortices, also known as eddies, isolate heat and nutrients from the surrounding waters, and transport these properties across great distances. This in turn influences biological productivity and carbon uptake from the atmosphere, and ultimately affects the rate of carbon accumulation that forces atmospheric climate change. In studying oceanic vortices, researchers would like to categorize them into well-formed vortices that tend to transport water properties and be long-lived, while rejecting areas of vorticity caused by shear layers near boundaries and jets. The *Q-criterion* is the standard oceanographic statistic to find vortices, but it does not differentiate between these behaviors, while the method we propose here does.

The first issue to consider with the ocean simulation data is vortex identification, followed by non-vortex removal. One of the analytical problems with the *Q-criterion* in the oceanographic community is one of reconciliation. Different techniques yield different numerical ranges and scaling, but oceanographers prefer to communicate and share data in an internally consistent way. To address these problems, oceanographers typically normalize their Q values to the standard deviation of Q over the domain they consider. This creates more problems, since the different current structures in various regions lead to different normalizations based on the domain being considered. For example, a regional study of the North Atlantic, with the strong shearing and eddies of the Gulf Stream, could have a more drastic *Q-criterion* normalization than the relatively quiescent Mediterranean Sea. Nonetheless, the Q values shown in this section are normalized to the standard deviation at the surface over the entire domain to maintain the standard used by oceanographers. Since our study is concerned with confidence rather than a particular Q threshold (or range), and the confidences are computed per-vortex, this normalization does not affect our results in any meaningful way.

We demonstrated in the previous section that our method can be used to identify suitable values of Q for extracting a vortex at a chosen confidence. For the ocean data, we do this extraction on a per-vortex basis. Beginning with minima of Q, we iteratively add adjacent points, always selecting the neighboring point with lowest Q value. As the value of Q in a vortex rises, we keep track of pairs of Q and area (point

count) and compute R^2 for the running list of (Q,A) pairs against their linear fit. Once the R^2 value falls below a chosen threshold, the method stops and the maximum Q found on this search is associated with that vortex.

Figure 6a shows the Q field off the eastern coast of the United States, where the powerful Gulf Stream carries warm water from the Gulf of Mexico north along the coast, before departing the continental shelf at Cape Hatteras and eventually flowing into the North Atlantic Current. In a standard oceanographic analysis, everything below -0.2 in the *Q-criterion* would be considered a vortex; in our visualizations, this corresponds to all the red and pink regions in Figure 6a. Our new method shows that two distinct types of structures are not Guassian-like vortices and are filtered out: small scale boundary shear and meanders of jets.

Vorticity due to boundary shear is clearly seen in Figure 6a at the edge of the Greenland coast, Labrador Sea, and Hudson Bay, and is simply the result of fluid rolling up between ocean currents and a no-slip land boundary. Almost all of this is filtered out at the 95% confidence level in Figure 6b (in pink), showing that the boundary vorticity does not closely fit the description of a Gaussian vortex. At lower confidence levels, a small number of vortices in the roll-up regions are admitted.

On the other hand, as the Gulf Stream slowly meanders, large vortices are pinched off into the Atlantic Ocean. The meanders themselves, however, also generate very high vorticity, and appear in the Q plot as quarter- and half-turns in the negative (red) domain. Again, these meanders do not generally match the description of a Gaussian vortex, so are rejected. When our method is applied, as in Figure 6b, features that appear to be true vortices are shown to be highly preferred (i.e., in blue or pink) over meanders. For many meanders, simply failing to match one of the two assumptions of our method (Gaussian vorticity, and near-circular or elliptic shape) causes them to be excluded.

There are however some corner cases where a meander's vorticity profile can resemble that of a Gaussian vortex, and where its quarter- or half-turn shape is similar enough to a deformed ellipse. In most cases such meanders will still be ruled out at high confidence levels, as the meander causes water inside the area of the arc to turn as well, creating a lower-magnitude patch of vorticity that the meander turns about.
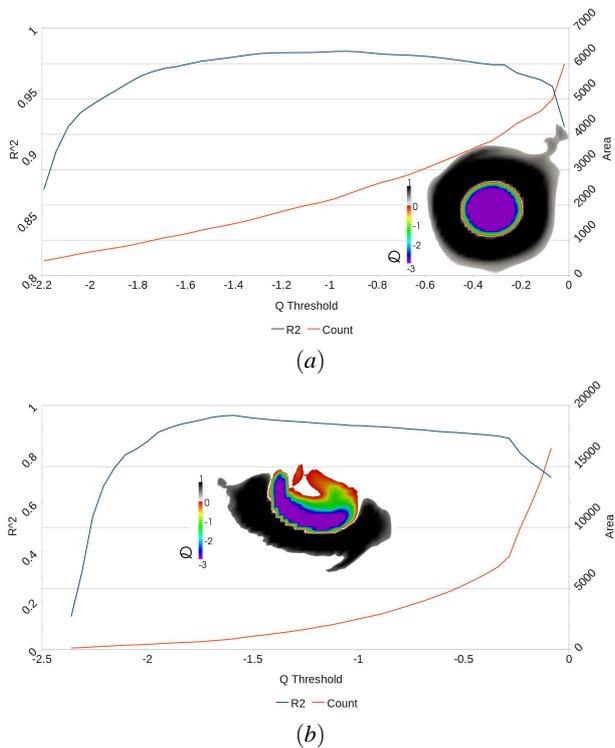
(a)

(b)

Fig. 7. Graphs of Q threshold versus area, and Q threshold versus R^2, for a well-behaved vortex (a) and a well-behaved meander (b) in the global ocean simulation data, with the features themselves inset. (a): Q grows linearly with respect to area for almost the entire span of the vortex: the graph ends just before the Q threshold is high enough to begin including the background ocean. After some initial noise (our detection algorithm accounts for this by allowing a vortex to grow initially without R^2 checks, to ensure it contains enough samples to be numerically stable), the R^2 of the vortex is above 90% for all the span checked here, and most of it above 95%. Note in the inset that the near-quadratic growth of Q with respect to radius causes each colored ring to be thinner than the last. (The color scale clips all values below -3 to the same purple, hence the large purple center.) (b): The core of the meander also resembles a vortex, and has fairly good R^2 for low thresholds of Q. However, this meander induces rotation inside its arc, causing a larger secondary vorticity patch. Once this patch begins being added at Q thresholds of about -1.5, the rate of growth of the vortex increases, causing its confidence to fall below 90% past thresholds of about -0.7. In the inset, the purple and blue form bands in a similar fashion to the vortex in (a), while the green region (representing values near -1.25) is much larger than the blue. This violates the expectation that Q grows quadratically with radius, and corresponds to the Q value in the plot where R^2 starts dropping.

The minimum Q for a meander will lie within the meander itself, so propagation will begin within the meander, and could very well grow with area at a fairly linear rate. Once the propagation reaches this inner vorticity zone, however, the amount of area added with each change of the Q threshold will greatly out-pace that inside the meander, as shown in Figure 7b, and R^2 will fall off rapidly.

Finally, we present two more in-depth analyses of the additional data generated by this method in Figures 8 and 9. These analyses rely on tracking to see whether a high confidence is related to long-term stability. We wish to summarize the data, which in their raw form take up almost a terabyte. To do this, after computing the confidence of a vortex, we write out a summary that approximates the vortex as a circle with center of the vortex's centroid at the depth where the vortex is strongest (i.e., with minimum Q value), and radius of $\sqrt{\pi A}$, where A is the number of points at the same depth. (That is, the summary is generated from the intersection of the vortex with a plane parallel to the surface of the Earth.) Additionally, we approximate the translational velocity of the vortex as the average velocity of all points in the vortex. This follows from the idea of a spinning hockey puck translating across a rink: points on opposite sides of the puck will cancel each other's rotational velocity, leaving the translational velocity to emerge from the sum.

To compute the tracking, each vortex is translated into the next time step by converting its velocity in degrees per day, since each time step is one day long, and adding that velocity to its centroid. Two vortices are considered the same (i.e., given the same id number) if their centroids are close together, and their radii are similar. Centroid closeness is taken as a ratio between centroid separation and the smaller radius, which we require to be below 0.25, and we require that the ratio between the smaller and larger radii be above 0.7. We plot the courses followed by the 5000 longest-lived vortices in Figure 8a, with colors assigned randomly in order to distinguish between nearby vortices. This plot agrees with the standard observations about ocean vortices. First, they primarily occur near three major currents: the Gulf Stream off the eastern coast of the United States, the Kuroshio Current off the coast of Japan, and the Antarctic Circumpolar Current, ringing Antarctica in the Southern Ocean. Additionally, large vortices appear in this plot at several well known regions of eddy activity: the Agulhas Rings following a course from the southern tip of Africa heading towards Brazil, and a Gulf Ring, created as a meander as the Loop Current passes through the Yucatan Channel, appearing as a coherent vortex downstream of the Channel and then migrating west across the Gulf of Mexico. Several large and long-lived vortices are seen in the Mediterranean Sea, most of which are nearly stationary.

In order to address the complementary question, of whether the vortices being excluded by this criterion should be excluded, we next analyze the relationship between our criterion and lifetime. We apply the tracking algorithm to the full set of high-vorticity features regardless of confidence. The algorithm assigns an id number to each vortex, such that vortices considered to be the same across different time steps receive the same id number. Each id is then given a lifetime, computed as the date of the last time step that id appears in minus the first time step it appears in plus one. We then divide the high-vorticity features into those that have at least 90% confidence, and those that do not. For each of these sets, we compute the probability distribution function of lifetime. In other words, for a given vortex, what is the probability that it will live for n days, for all n? These two PDFs are shown in Figure 8b, with logarithmic scaling on probability. This plot shows that, indeed, high-confidence vortices are more likely to live longer, and low-confidence vortices are more likely to have a shorter life.

For our second analysis, shown in Figure 9, we focus on five relatively long-lived vortices in the North Atlantic. The vortices are drawn in time step order, so that later time steps occlude earlier time steps. Thus we can see that all six vortices are traveling roughly west. Below the map, the confidence in each vortex is plotted over its lifetime. All plots fall between confidences of 0.8 and 1.0, with the center line at 0.9. The plots show a large amount of small-scale jitter (on a scale of about 0.02), and occasionally the vortex will jump abruptly down and back up. This is most likely due to either the vortex interacting with another, causing each to distort, or possibly due to numerical issues. Interestingly, only the blue vortex has a confidence drop-off before the tracking is lost. This could either indicate that the time scale of vortex death is quite fast, or that the algorithm excludes the dying vortex, e.g., because a fast change in radius meant the tracking criteria were not met. The specific mechanics of vortex death in the ocean, anyway, is a generally open question that we hope to investigate in future work.

7 CONCLUSIONS

We introduced a new, statistically-driven technique for specifying and extracting vortices from the *Q-criterion*. Our method derives from the math and physics of Gaussian vortices a novel near-linear relationship that can easily be evaluated using a coefficient of determination. This produces, on a per-vortex basis, a level of confidence of a comparison between a real vortex and a model vortex. We can then backtrack from a statistical confidence to the maximum threshold of Q associ-
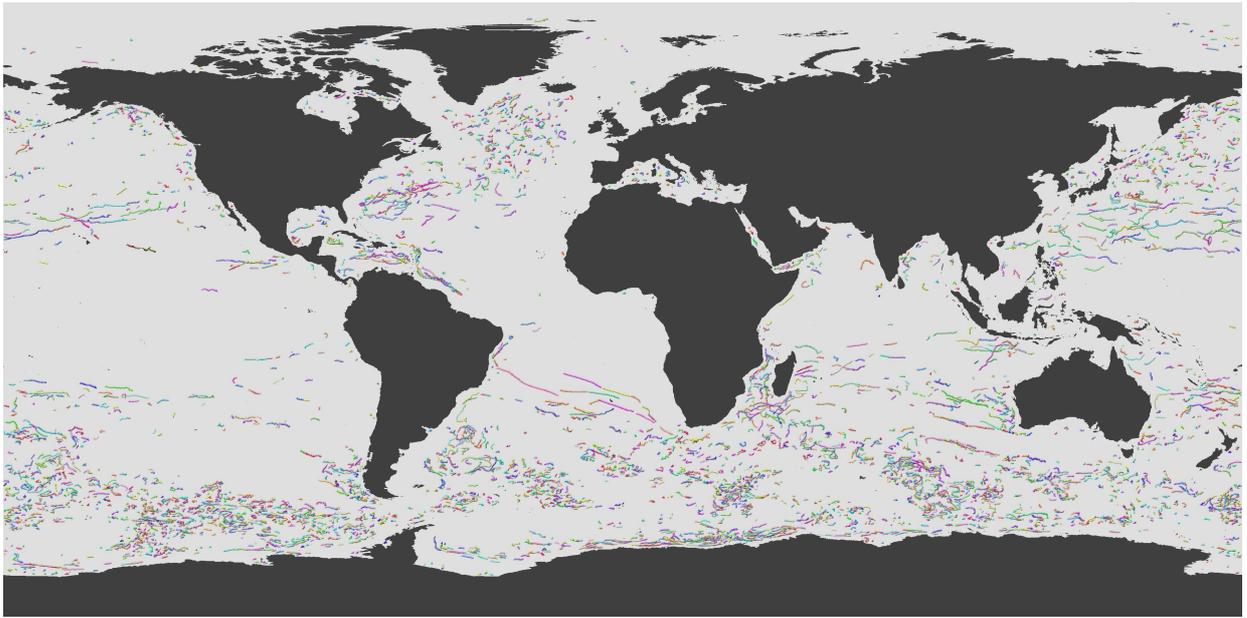
(a)
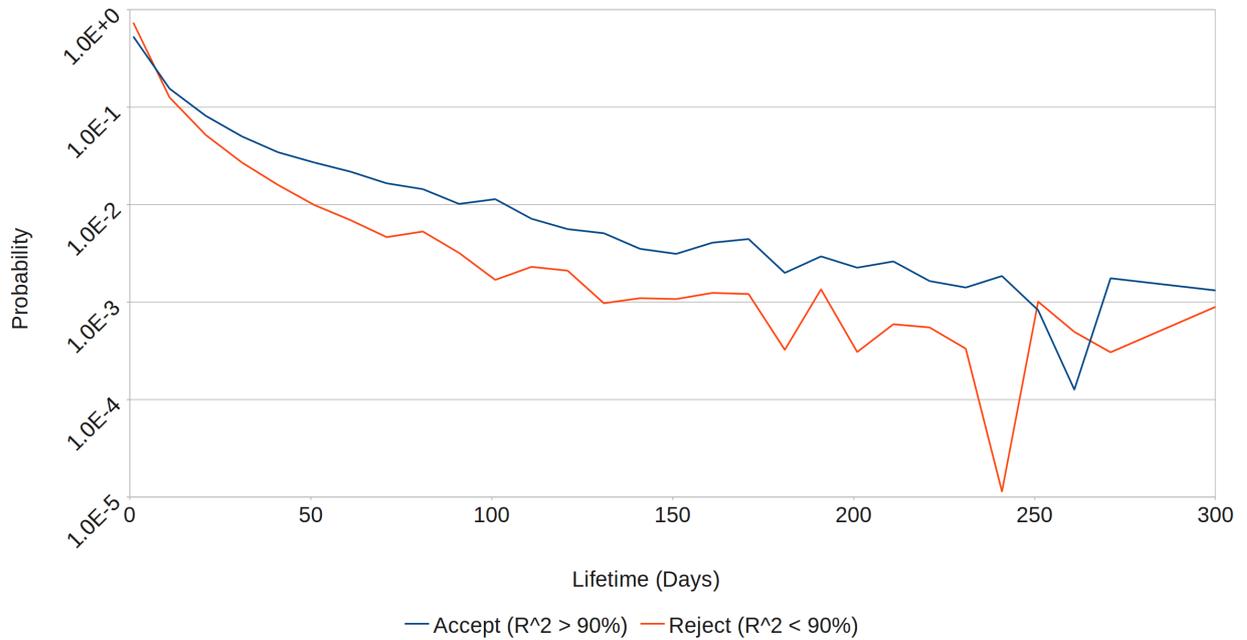
(b)

Fig. 8. (a) After applying feature tracking to the data, we take the 5000 longest-lived vortices and plot the courses they take, with random coloration to distinguish individual vortices. As expected from observation, vortices appear primarily associated with the Gulf Stream off the eastern coast of the United States, the Kuroshio Current off the coast of Japan, and the Antarctic Circumpolar Current around the Southern Ocean near Antarctica. Also as expected, vortices are relatively rare near the equator, where Coriolis forces are weak. (b) To determine whether high-confidence vortices are also longer-lived, we take all high-vorticity features (regardless of confidence) and apply tracking. Vortices are given id numbers based on the tracking (so if two vortices in adjacent time steps are identified as being the same vortex at different times, they're given the same id), then each id number is given a lifetime of the date of the last time step it appears in minus the date of the first time step it appears in plus one. We then compute a probability distribution function of vortex lifetime, but splitting vortices into those with confidence above 90% and those below. A vortex can over time go above and below the threshold, so each vortex of each id number is counted separately. We see from this that low-confidence vortices are indeed more likely to have shorter lives, while high-confidence vortices are more likely to live longer.
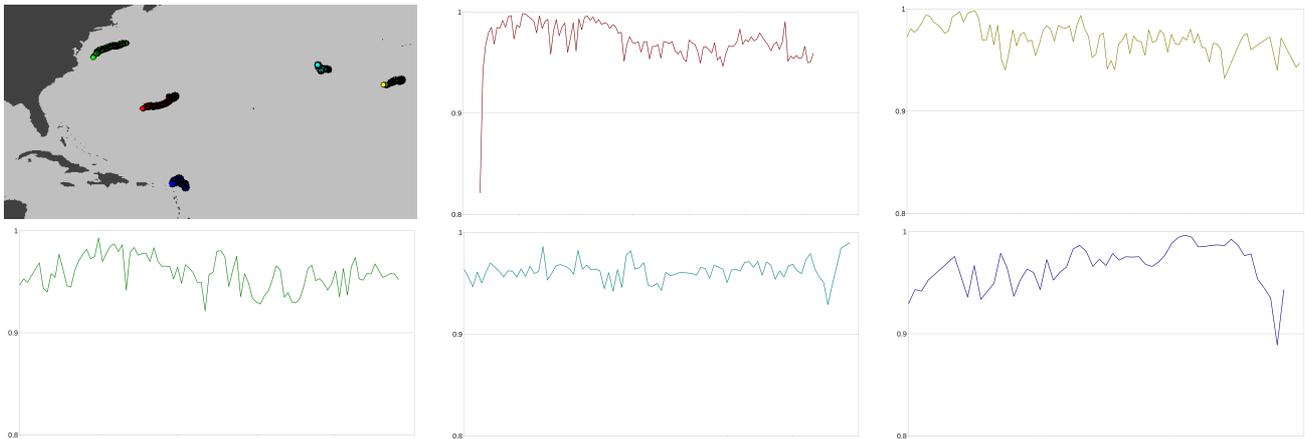
Fig. 9. We extracted five long-lived vortices from the North Atlantic and plotted them in different colors. All five vortices are tending west: the circles representing the vortices were drawn in time step order, so later time steps occlude earlier time steps. Below, the confidence of each vortex is plotted over its lifetime. All confidences range between 0.8 and 1.0, and each graph is associated with the vortex with the same color as the graph's line. The x-axis of each graph is time, scaled over each vortex's lifetime, so the start, end, and horizontal scaling of the graphs are not directly comparable. Despite some small-scale jitter and occasional sharp jumps (most likely caused by temporary distortion from interacting with another vortex), the confidences are fairly stable up until the vortex is lost.

ated with that confidence, allowing us to further describe how well, within a single confidence level, different vortices adhere to the definition of Gaussian vortices. Doing so removes the ambiguity from choosing transfer functions for rendering, and from choosing thresholds to define a vortex for data analysis.

Using data from a global ocean simulation, we show that our new method is able to pick out well-formed, Gaussian-like vortices, while rejecting vorticity in boundary shear layers and meanders of jets. This distinction is scientifically important because well-formed vortices isolate source waters and transport heat and nutrients that influence the carbon cycle and climate change [10]. One may track eddies across the ocean, and observe that its maximum confidence value remain fairly consistent for isolated eddies.

The next step is to apply this Guassian vortex method to other velocity data sets, such as satellite observations of the ocean and atmosphere, and to expand our vortex tracking to collect time-dependant statistics of other fields like temperature and salinity for each vortex. We could then investigate what component of heat and salt transport is due to discrete eddies, and whether Gaussian-like eddies transport water properties more efficiently than other eddies. Applying this method to three dimensional flows is another potential direction, which would require expanding the analytical foundations of this work and testing it on non-geophysical data sets.

REFERENCES

[1] D. C. Banks and B. A. Singer. Vortex tubes in turbulent flows: identification, representation, reconstruction. In *VIS '94: Proceedings of the conference on Visualization '94*, pages 132–139, Los Alamitos, CA, USA, 1994. IEEE Computer Society Press.

[2] R. A. Crawfis and N. Max. Texture splats for 3d scalar and vector field visualization. In *Proceedings of the 4th conference on Visualization '93*, VIS '93, pages 261–266, Washington, DC, USA, 1993. IEEE Computer Society.

[3] W. K. Dewar and P. D. Killworth. On the Stability of Oceanic Rings. *Journal of Physical Oceanography*, 25:1467–1487, June 1995.

[4] J. C. R. Hunt, A. A. Wray, and P. Moin. Eddies, streams, and convergence zones in turbulent flows. In *Studying Turbulence Using Numerical Simulation Databases, 2*, pages 193–208, Dec. 1988.

[5] M. Jiang, R. Machiraju, and D. Thompson. A novel approach to vortex core region detection. In *Proceedings of the symposium on Data Visualisation 2002*, VISSYM '02, pages 217–ff, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[6] Jinhee Jeong and F. Hussain. On the identification of a vortex. *Journal of Fluid Mechanics*, 285:69–94, 25 February 1995.

[7] P. K. Kundu. *Fluid Mechanics*. Academic Press, 1990.

[8] R. Laramee, H. Hauser, L. Zhao, and F. Post. Topology-based flow visualization, the state of the art. In H. Hauser, H. Hagen, and H. Theisel, editors, *Topology-based Methods in Visualization*, Mathematics and Visualization, pages 1–19. Springer Berlin Heidelberg, 2007. 10.1007/978-3-540-70823-0_1.

[9] K. J. Mallen, M. T. Montgomery, and B. Wang. Reexamining the Near-Core Radial Structure of the Tropical Cyclone Primary Circulation: Implications for Vortex Resiliency. *Journal of Atmospheric Sciences*, 62:408–425, Feb. 2005.

[10] D. J. McGillicuddy, Jr., L. A. Anderson, S. C. Doney, and M. E. Maltrud. Eddy-driven sources and sinks of nutrients in the upper ocean: Results from a 0.1 degrees resolution model of the North Atlantic. *Global Biogeochemical Cycles*, 17(2):1035–, 2003.

[11] A. Okubo. Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. *Deep Sea Research and Oceanographic Abstracts*, 17(3):445 – 454, 1970.

[12] C. Petz, J. Kasten, S. Prohaska, and H.-C. Hege. Hierarchical vortex regions in swirling flow. *Computer Graphics Forum*, 28(3):863–870, 2009.

[13] S. C. Riser, W. B. Owens, H. T. Rossby, and C. C. Ebbesmeyer. The Structure, Dynamics, and Origin of a Small-Scale Lens of Water in the Western North Atlantic Thermocline. *Journal of Physical Oceanography*, 16:572–590, Mar. 1986.

[14] D. Schneider, A. Wiebel, H. Carr, M. Hlawitschka, and G. Scheuermann. Interactive comparison of scalar fields based on largest contours with applications to flow visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1475 –1482, nov.-dec. 2008.

[15] D. Silver and X. Wang. Tracking and visualizing turbulent 3d features. *Visualization and Computer Graphics, IEEE Transactions on*, 3(2):129 –141, apr. 1997.

[16] R. D. Smith and P. Gent. Reference manual of the Parallel Ocean Program (POP). Los Alamos National Laboratory report LA-UR-02-2484, Los Alamos National Laboratory, Los Alamos, NM, 2002.

[17] V. Sood, B. John, R. Balasubramanian, and A. Tandon. Segmentation and tracking of mesoscale eddies in numeric ocean models. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III – 469–72, September 2005.

[18] X. Tricoche and C. Garth. Topological methods for visualizing vortical flows. In G. Farin, H.-C. Hege, D. Hoffman, C. R. Johnson, K. Polthier, and M. Rumpf, editors, *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, Mathematics and Visualization, pages 89–107. Springer Berlin Heidelberg, 2009. 10.1007/b106657_5.

[19] J. Villasenor and A. Vincent. An algorithm for space recognition and time tracking of vorticity tubes in turbulence. *CVGIP: Image Understanding*, 55(1):27 – 35, 1992.