Visualizing Ensemble Data in Scale Space

Fangyan Zhang[1], Song Zhang[1], and Andrew Mercer[2]
[1]Department of Computer Science and Engineering, Mississippi State University, Mississippi State, MS, USA
[2]Department of Geosciences, Mississippi State University, Mississippi State, MS, USA

Abstract— Ensemble data are becoming increasingly more common in scientific research, for example, numerical weather simulations or ocean flow simulations. Consequently, ensemble visualization has become an important means for analyzing and inferences of the simulation results. One challenge in ensemble visualization is the overwhelming amount of details that disguise important large-scale features. This paper approaches the problem by applying the scale space methods to ensemble data visualization. This helps remove small features at a high scale in visualization. Further, visualizing two or more ensemble members within scale space facilitates the comparison between ensemble members and the identification of differences at varying scales, making it easier to focus on features of a certain size and the key differences among separate ensemble members. We applied our approach to an ensemble numerical weather data and an ocean data, and conducted a qualitative evaluation of the approach. The results show that our method is useful in a research environment, in particular, for feature tracking and identification of large-scale phenomena within ensemble data sets.

Index Terms— Scale space, Weather Ensemble, Feature Tracking.

1 INTRODUCTION

Scientific simulations in geospatial domains (e.g., weather or ocean simulations) generate large amounts of data that are often multidimensional and multivariate. In addition, a data ensemble arises when multiple simulations are run with different parameters or initial conditions. To make sense of the data, domain experts need to track features effectively and efficiently through this sea of data, and compare ensemble members. This paper examines the usage of scale space for multidimensional and multivariate ensemble data.

Numerical weather simulation data is a prime example of such data. Visualization has proven to be an effective approach for weather data analysis [1][2][3]. Visualization helps meteorologists extract and analyze features from the data. However, how to visualize weather data effectively and how to apply visualization to weather analysis remain challenging for both visualization experts and meteorologists.

Currently, numerical weather simulations are a common part of most weather related forecasts and have become a standard component in people's daily lives. Using meteorological understanding, researchers construct numerical weather prediction models, which are utilized to make forecast inference. These simulations are often run as ensembles by varying simulation parameters and initial conditions to capture more information about the underlying distribution of possible solutions. However, the ensemble simulations also increase the amount of data, and in turn the difficulty in visualization and analysis.

One challenging aspect of geospatial data visualizations like weather data visualization is the volume of small scale features present in the data that might disguise the large and important features. For example, in a synoptic-scale extratropical cyclone simulation, meteorologists are interested in seeing the large low pressure region responsible for the major impacts. This low pressure region can be visualized by a contour on the pressure field. However, with a set isovalue, many small low pressure regions will be created as smaller contours, and there might be many highly irregular segments on the large contours that distract the meteorologists' attention from identifying the main shape of the contours. These problems become more pronounced with the increased amount of data in ensemble data visualization.

A second challenge in geospatial data visualization is how to compare ensemble data effectively at multiple scales. Comparisons of data both within and between ensemble members is a prerequisite for scientists to make meaningful inferences when using ensemble data[4].

A third challenging aspect in geospatial data visualization is how to track big features over time with many small features present. Meteorologists often demand a perception of how features evolve over time. A large amount of small features will most likely detrimentally affect the accuracy of inferences made upon a weather event. In essence, the more small features present, the more difficult it becomes to identify the main feature.

Scale space theory is a framework used in early visual operation approaches by the computer vision community [5][6][7]. It was originally designed to deal with the multi-scale nature of image data. The basic idea is to embed the original signal into a one parameter family of gradually smoothed signals, in which the signal details are successively suppressed [8]. This paper applies scale space principles to visualizing geospatial data, removing minor features, and leaving large features in a higher scale. Even though big features are slightly changed with increasing scale, it is more important that small features disappear dramatically. When visualizing weather data in scale space, features can be obtained and analyzed for each ensemble member, and they can be compared across ensemble members at multiple scales. This presentation of weather features at multiple scales can help meteorologists identify their usually large target features more efficiently and more consistently.

In scale space, a Gaussian filter is used to convolve with the original image to generate a series of derived images at multiple scales. A. Kuijper [9] has proven that the Gaussian kernel is a non-spurious detail generating filter. Z. Yang et al. [10] provided normal forms of Gaussian scale space that can keep singularities during the evolution of Gaussian scale space.

In this paper, we employ Gaussian scale space to visualize ensemble geospatial data including numerical weather simulation and ocean simulation. We apply our approach to both 2D (two-dimensional array of longitude and latitude coordinates) and 3D (three-dimensional array of longitude, latitude, and time-steps) data. We use features like extrema values and contours to visualize data and construct intuitive tree or dot representations of these features in scale space, which are discussed further in the Technique section. The contributions of this paper include the following:

- We present a framework to visualize geospatial ensemble data in scale space.
- We demonstrate the effectiveness of scale space in automated feature tracking, noise removal, and ensemble comparison.
- We provide an approach to display ensemble data across different scales at the same time.
- We applied the techniques to two kinds of ensemble data and conducted a user evaluation.

The following section provides an overview of the related work. A detailed explanation of our proposed approach is given in section 4. We apply the techniques to ensemble simulations of a significant extratropical cyclone weather event (herein referred to as the cyclone dataset), and explain how to use the techniques to interpret various weather phenomena in section 5. Additionally, a user study conducted with a meteorologist was described in section 6. Finally, we conclude our work in section 7.

2 Background and Related Work

An ensemble is a set of similar and potentially multivariate and multidimensional data from numerical simulations (e.g., weather or ocean simulations). Many previous researchers have shown that visualization is an effective method for exploring an ensemble [4][11][12]. Traditional ensemble data visualization approaches can be classified into two categories [13]. The first category is feature-based visualization, which obtains features from individual members and compares the features among ensembles. Alabi et al. [14] propose a method to identify differences and similarities between surfaces from different data sets. Smith et al. detect features using clustering based techniques on time-varying data [15]. The second category is location-based visualization that compares ensembles at a fixed location. These techniques [13][11][16] utilize commonly-used statistical measures, for example, means and variances of scalar quantities, to show the difference between ensembles at fixed locations.

In addition, ensemble visualization approaches can also be categorized as side-by-side and coincidental [4]. The first method display the ensemble members with small multiples [17]. The second approach is to integrate the ensemble members in a single display (e.g. spaghetti plot).

Although several of those approaches have been presented in ensemble visualization and existing visualization techniques are able to present ensemble members in their original scale, it can be challenging scientists to discern the main differences between multiple ensembles with lots of small features present. Scale space helps address this challenge.

Level sets such as contours and isosurfaces are prevalent tools for tracking features in the data [18][19]. They have proven to be effective approaches for uncertainty and ensemble data visualization [20] [21] [22] [18]. In weather forecasting, contours and isosurfaces have the advantage of visualizing continuous features in the atmosphere, e.g., troughs and ridges[23], which are important for weather analysis and weather forecasting. Based on those reasons, we apply both of them in our visualization.

The word “scale space” was first proposed by Witkin [5] in 1983 and referred to the properties of a 1D signal at different scales. In 1984, Koenderink [6] extended the application of scale space to 2D images. Scale space utilizes a smoothing operator to construct a series of smoothed data from the original.

A Gaussian scale space is constructed by convolving the spatial data with a Gaussian kernel. The results are a series of smoothed version of the original data. There exists nonlinear scale space such as the one defined in [24].

In addition, another concept, multi-scale, has similar denotation as scale space. The idea of multi-scale representation of data evolved from computer vision. Rosenfeld and Thurston [25] represented a signal at multi-scales, and employed it in edge detection. SiZer, a scale space visualization tool, was developed by Chaudhuri and Marron in 1999 [26] to discover the significant structures in the data.

At present, scale space has applications to many fields in visualization, such as topology simplification in vector fields and vortex detection [27][28], saliency in natural images[29][30][31], and medical image processing[32][33]. Additionally, the scale space notion is consistent with the numerous scales of atmospheric motion that exist (synoptic-scale, mesoscale, microscale [23]), and such an approach will allow users to discern features that exist at each of these scales. Despite the logical potential for application, there is little research on scale space visualization in ensemble weather data visualization research. This paper introduces scale space into geospatial data visualization and employs the technique to analyze ensemble members. As we will show, by accounting for multi-scale elements, this work will identify meaningful features based on contours and isosurfaces at various scales. We also leverage the principle of scale space to compare ensembles: make simple and effective comparison within or between ensembles at multiple scales and present these differences to domain experts.

Topology methods such as contour trees [34] have also been used to represent contours in the data. While a contour tree records the change of topology in contours according to the change of isovalues, it does not try to simplify contour at any level. Hence, the contour tree method is complementary to the scale space method.

Figure 1 illustrates the advantage of showing the weather data at multiple scales. The figure shows the absolute vorticity of a cyclone data. The left image shows contours in the original data while the right image shows the contours at a higher scale. Many details present in the original image are removed at a higher scale while the general shapes of the large contours are preserved well.

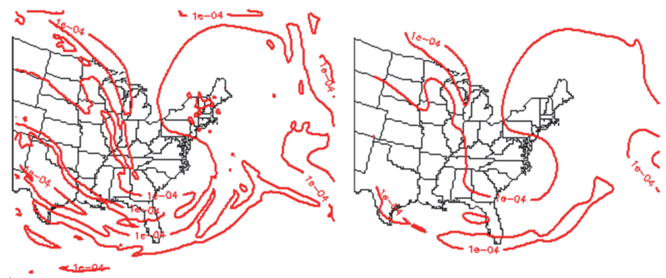


Figure 1: Comparison of contours at the original (a) and at a higher scale (b).

3 Extratropical Cyclone Data and Ocean Data

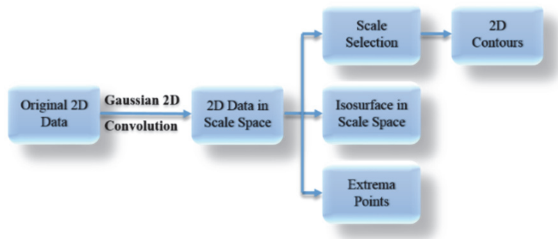
Using the National Weather Service Science Operations Officers/Science & Training Resource Center Weather Research and Forecasting (WRF) [35], we created a 30 member ensemble simulation of a winter East Coast extratropical cyclone. The cyclone originates off the coast of the Carolinas at 1800 UTC on 1 March 1999 and follows the coastline northeastward until the end of the simulation (1800 UTC 4 March 1999). Ensemble members were generated by varying the planetary boundary layer physics (2 schemes), the cloud microphysics (5 schemes), and the cumulus parameterizations (3 schemes) prior to running the model, yielding 30 possible combinations. We refer to 2D data as data at a single vertical level for a single simulation hour (latitude-longitude) and 3D data as 2D data that include all simulation hours (the full 72-hour simulation). In this application, three variables are chosen by a meteorologist: absolute vorticity (Figure 6, 7, 8, 10), mean sea level pressure (Figure 3, 4, 5, 9). We select the 300 mb isobaric level for absolute vorticity.

The ocean data created by the Naval Research Lab is a simulation of ocean flow near the coast of Japan from 2011-05-01 00:00 UTC to 2011-05-2 00:00 UTC, following the Tohoku earthquake and tsunami. The simulation consists of 32 ensemble members. Each ensemble member has 17 time steps with an interval of 3 hours. In the application, sound speed is chosen, (Figure 11).

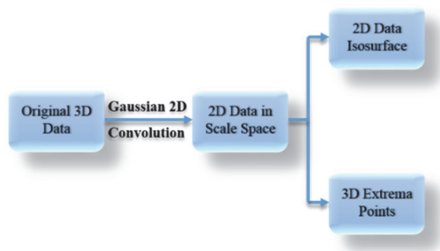
4 TECHNIQUE

This section discusses our approach to visualize ensemble 2D and 3D data sets in scale space. First, we establish scale space using the Gaussian kernel in section 4.1. Then, we present a set of techniques for ensemble comparison and tracking features in the data and their adaption to scale space.

These approaches are illustrated in Figure 2, 2D data methods (a), and 3D data methods (b).



(a)



(b)

Figure 2: Overview of approaches for (a) 2D data and (b) 3D data.

4.1 Scale Space Construction

Scale space is one type of multi-scale representation with a continuous scale parameter [8]. There are several space scale representation methods, such as Laplacian and Gaussian. Laplacian filters are commonly used to find abrupt changes in an image. A Gaussian filter has the property that there is no new spurious structures being created from the original scale to higher scale [5][8]. Also, Lindeberg [7] and Iijima [36] showed that the Gaussian kernel is the unique kernel for generating linear scale space. Linearity, shift invariance, semi-group structure, non-enhancement of local extrema, scale invariance and rotational invariance are some of the properties possessed by Gaussian kernel [37]. For these reasons, we chose the Gaussian filter to construct our scale space.

If original data is two dimensional, then two dimensional Gaussian is used to create scale space from the original data. The filter weights can be calculated as:

$$f(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ is the standard deviation of the Gaussian filter, which is also called the scale space parameter and decides the degree of coarseness of the resulting image. At scale s, the coarse image can be represented as [Reference]:

$$f_s(x,y,\sigma) = \frac{1}{2\pi s\sigma^2} e^{-\frac{x^2+y^2}{2s\sigma^2}}$$

If original data is three dimensional, then three dimensional Gaussian is applied to create scale space from the original data. The filter weights can be calculated as:

$$f(x,y,z,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}$$

likewise, σ is the standard deviation of the Gaussian filter, also called the scale space parameter, which decides the degree of coarseness of isosurface in 3D scale space. At scale s, the coarse isosurface can be computed as:

$$f_s(x,y,z,\sigma) = \frac{1}{2\pi s\sigma^2} e^{-\frac{x^2+y^2+z^2}{2s\sigma^2}}$$

Gaussian convolution creates a series of smoothed data with different value of σ. The scale space adds an additional scale dimension to the data. For example, if the original data is 2D, the outcome is 3D. If the original data is 3D, the result is 4D.

4.2 Extrema Points Evolution in 2D Scale Space

Key points, especially minima and maxima in a data set are important features that often have important physical meanings. For example, the cyclone center is identified by the location of the lowest pressure value in the region. By visualizing minima or maxima values in scale space, we allow users to track the evolution of these potentially dominant features across scales and explore how these features persist or disappear with the increasing scale, hence easily tracking persistent minima and maxima such as a cyclone center and disregarding small and unimportant ones.

4.2.1 Extrema Points of One Ensemble Member in 2D Scale Space

Extrema points are widespread in two-dimensional data. For example, Figure 3 shows the mean sea level pressure of extratropical cyclone in a heat map. Those black dots represent minima or maxima values. From the heat map, we can easily observe that there are several low pressure and high pressure regions. However, if we want to automatically identify those features, it is not an easy task as we thought because they are disguised by a lot noises (lots of noise around the low pressure region and high pressure region). An algorithm is needed to differentiate the real dominant features (low pressure region and high pressure region) from noises.

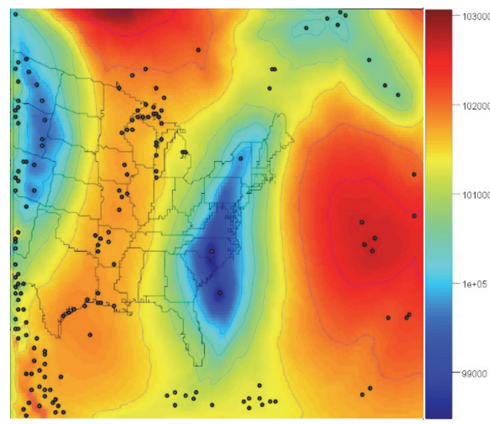


Figure 3: Minima-maxima for mean sea level pressure in extratropical cyclone data. The background is part of U.S. map. Colors in the heat map specify pressure values. Black dots are the minima and maxima points.

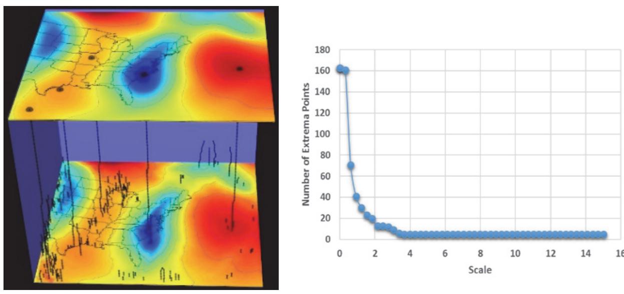
In scale space, pervious research has proved that minima and maxima points at each scale can be identified, and noises can be suppressed. Shafii [38] developed auxiliary join tree or auxiliary split tree to represent the evolution of extrema points. These key points grow continuously in the scale space, forming a tree structure. In evolution, arrows are used to indicate gradient vector field. However, auxiliary join tree or auxiliary split tree are not suitable to represent the evolution in geospatial data because it is impossible to identify the location in extrema evolution. The location of extrema points in geospatial data is significant to interpret features. In our methods, we not only identify those extrema points, but also preserve their track in evolution. In addition, our approach identifies extrema by comparing current element with its eight neighbors. Extrema points are represented in 3D during evolution procedure. The additional dimension stands for scale, which precisely denotes the changes of extrema points in evolution.

One two-dimensional data can be considered an image. Our algorithm can be described as follows.

1) Identify extrema p_0 in the original image0.
2) The original image are smoothed into image1 by choosing σ_1.
3) From image1, identify extrema p_1 who are neighbors of points in p_0. If the new extrema points are not neighbors, they are artificial extrema, remove them.
4) Repeat step 2 to step 3 until noises of extrema are merged into dominant features (only dominant extrema points left).

In evolution, in order to reduce the influence on boundary topology, we choose closest point condition by using boundary value in pad layers because it has been tested [38] that this strategy is relative best choice. For better describing the extrema evolution, we also record number of extrema.

By applying the method to extratropical cyclone data in Figure 3, we can get the results in Figure 4. Noises are gradually removed by increasing scales. At top scale, there are only five dominant features (big dots at top scale) are preserved. Apparently, extrema evolution is effective to track dominant features.

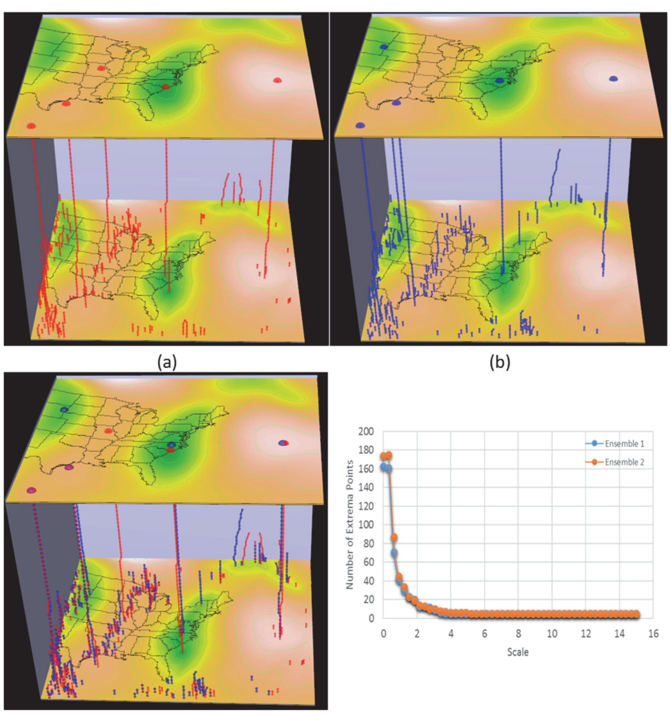


(a) (b)

Figure 4: (a) 2D (latitude, longitude) extrema points evolution for mean sea level pressure in extratropical cyclone data. The scale increases from bottom to top. The original data and highest-scale smoothed data are visualized in heat map. Black dots are the minima and maxima points. (b) The relation between number of extrema and scales.

4.2.2 Extrema Points of Two Ensemble Members in 2D Scale Space

Extrema points in ensemble data not only can track dominant features, but also compare ensemble members. The comparisons between ensemble members are facilitated by using extrema points. For example, Figure 5 shows comparison between two ensemble members through extrema points. Figure 5 (a) and (b) shows extrema evolution of the two ensemble members respectively. Figure 5 (c) is the comparison result. Figure 5 (d) illustrates how the extrema points change across scales.



(c) (d)

Figure 5: (a) Extrema points evolution for the first ensemble in extratropical cyclone data. (b) Extrema points evolution for the second ensemble in extratropical cyclone data. (c) Comparison of extrema points between the two ensemble members (in (a) and (b)). (d) The relation between number of extrema and scales.

4.3 Contours and Isosurfaces in Scale Space

Level sets, including 2D contours and 3D isosurfaces, are very commonly used for feature tracking in scientific domains for their versatility, rich details, and easy interpretation. Spaghetti plots (an overlay of contours from multiple ensemble members) are the de facto standard for visualizing ensemble weather data among meteorologists. However, the spaghetti plot easily becomes cluttered when the number of ensembles or the amount of details in the data increases. The visual clutter becomes much worse if multiple isosurfaces are overlaid in 3D. Representing contours and isosurfaces in scale space alleviate these issues.

For 2D data, the outcome of Gaussian convolution is a series of smoothed 2D data values with different scales. Contours at each scale are produced from the smoothed 2D data with the marching square method [39][40][41]. The contours can then be displayed to explore the 2D data in scale space, or overlaid among ensemble members for 2D ensemble data comparison at multiple scales.

For 3D data, the application of Gaussian convolution on 3D data will produce a series of smoothed 3D data, ranging from low scale to high scale. Isosurfaces can be generated from 3D data at any given scale. Isosurfaces can be displayed to examine one scale at a time, or overlaid to compare between ensemble members. Figure 6 is an example of 3D data in scale space. Figure 6 (a), (b), and (c) shows isosurface (red) of one ensemble member from low scale to high scale. Figure 6 (d), (e), and (f) exhibit isosurfaces (red and green) of two ensemble members from low scale to high scale.

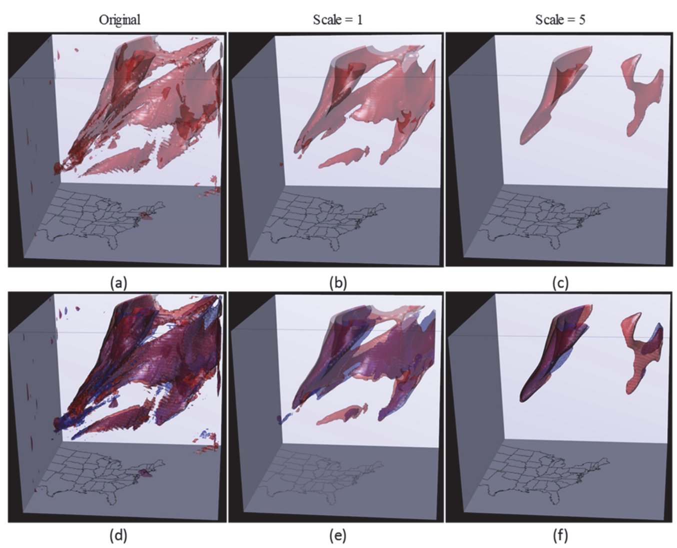


Figure 6: isosurface for one ensemble member (red in (a)(b)(c)) and two ensemble members (red and green in (d)(e)(f)) in scale space. Scale increases from left image to right image.

4.4 Interactive Scale Selection

The visualization of data in scale space shows the data and the features at multiple levels of detail. The user selection of scale is a key feature of our approach that allows the user to explore the data at various levels-of-detail, ideally within the same view. To provide maximum flexibility in scale selection, we allow users to interactively select different scales at different locations of the same data. In a case where a user is interested in features of different scales at different locations, e. g. a large extratropical cyclone centered over the Southeast United States and a squall line located over the Ohio Valley, this user-selected scale map becomes convenient. This is realized by creating an arbitrary scale surface, produced by cubic interpolation from scale values of several key points selected by users on the 2D data. Once users choose a set of positions on the map by mouse and input the scale values they want to observe at those positions, a scale surface would be created through cubic interpolation from the positions. The surface may cut across multiple scales, resulting in desired level-of-detail in all regions. Figure 7 shows the contours at low scale (a), high scale (b), and users-selected scale (d). Note that the user is more interested in the weather patterns of the east region but less interested in the west region. Hence, the user lifts the scale surface on the west while keeping the low scale in the east, resulting in detailed contours in the east while suppressing small features elsewhere. The resulting contours are the cross sections of the user-selected scale surface and the isosurface for 2D data in scale space (Figure 7 c).

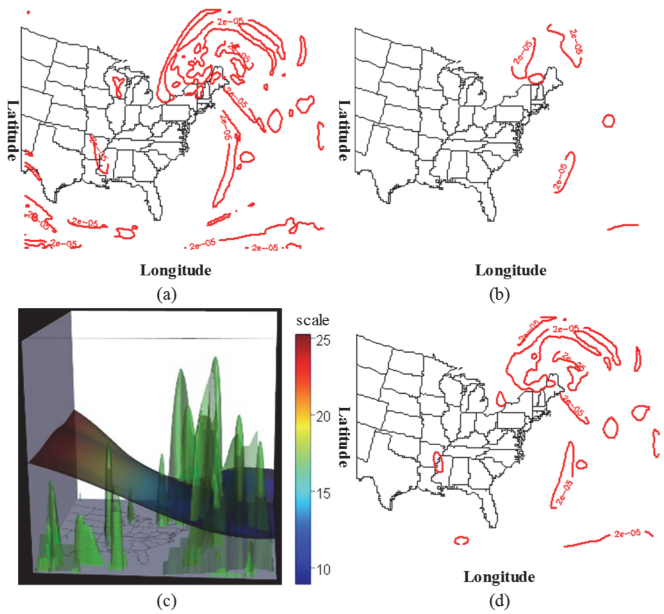


Figure 7: Interactive scale selection. (a) Contours when only low scale is selected. (b) Contours when only high scale is selected. (c): Interactive scale selection allows the user to select an arbitrary scale surface, which intersects the 2D scale space isosurface (in green) of absolute vorticity from the cyclone data at multiple scales. The bottom is part of US map, and vertical axis represents scale. (d): Resulting contours with the user-selected scale surface.

4.5 Curvature Extrema Evolution in 2D Scale Space

While contours are effective tools for tracking features, they still contain many details, especially at a lower scale. Comparison between ensemble members requires multiple contours to be drawn in the same space, often resulting in too many overlaps between them. Therefore, it is necessary to simplify contours representation.

Using key points on the contour can help reduce the complex shape to a few points, therefore facilitating comparison between multiple contours. Further, tracing these key points in the scale space may help users identify the evolution of the contours across scale more efficiently.

In our method, we compute the curvatures of all points on contours in scale space. Then we use curvature extrema to track contour key points evolution.

If a curve is defined as: $x = \varphi(t)$, $y = \psi(t)$, where x and y are 2D coordinates, $\varphi(t)$ and $\psi(t)$ are the parameterizations of the curve, t is the arc length of the curve, then the curvature function can be calculated as

$$K(t,\sigma) = \frac{\varphi'(t)\psi'' - \varphi''(t)\psi'}{[\varphi'^2(t) + \psi'^2(t)]^{3/2}}$$

We apply the 2D Gaussian convolution on whole images to produce curves or contours at different scales of the scale space instead of applying 1D convolution on curves directly. [44] and [45] employed 1D Gaussian on the curves to construct a curvature scale space. In their work, the positions of the points on a contour are only decided by its neighboring points on the contour instead of neighboring areas in the data. Since meteorologists are interested in the contours from 2D weather data at different scales, it is reasonable to examine the curvature of the contours independently generated at different scales of the 2D data instead of the contours sequentially smoothed by a 1D Gaussian.

Since data locations are discrete in weather data, we need to calculate the curvature in the discrete form. The curvature at vertex is expressed by

$$K_i = \frac{< v_i, v_{i-1}>}{|\frac{v_i - v_{i-1}}{2}| + |\frac{v_{i+1} - v_i}{2}|}$$

If V_{i-1} is (x_{i-1}, y_{i-1}), V_i is (x_i, y_i), V_{i+1} is (x_{i+1}, y_{i+1}), then

$$K_i = \frac{arctan\frac{y_{i+1} - y_i}{x_{i+1} - x_i} - arctan\frac{y_i - y_{i-1}}{x_i - x_{i-1}}}{\frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{2} + \frac{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{2}}$$

Figure 8 shows the evolution of contour curvature extrema points in scale space. We chose the variable of absolute vorticity for one ensemble of the cyclone data. The isovalue was set to 0.0001s-1. The points on the contours with minimal and maximal curvature values were selected. A threshold on the curvature value was also set to get rid of small perturbations (0.5). Figure 8 shows (a) the contours of two ensemble members with curvature extrema points in original 2D data, (b) the contours of two ensemble members with curvature extrema points in high scale of 2D scale space, and (c) Contour curvature extrema evolution of the two ensemble members in scale space. From figure 8 (c), curvature extrema points gradually reduce with increasing scales. The complex shape of contours are reduced to a few points. The comparison between contours or ensembles are transformed into the comparison between several key points. Apparently, comparison between ensemble members are facilitated, in particular for multiple contours drawn in the same space.

5 Experiment on Cyclone Data and Ocean Data

In this section, we apply the aforementioned scale space methods to the cyclone data and ocean data. We present the evolution of data in scale space and the differences between ensemble members through contours and extrema points in 2D scale space, which aids in tracking the main features across time steps and identification of the differences between ensemble members. The computation and visualization of the results were implemented in R. The data and the code will be made available to public access upon the publication of this work.

For the 2D data, we experimented with the visualization of a single and multiple ensembles in scale space. We investigated whether our approach can remove an adequate amount of spurious small features in the contours, for the relevant large features to become visible. In addition, we employed our approach to all ensemble members and examined whether the comparison became clearer with increasing scale.

Figure 9 shows the results of the extrema points of the mean sea level pressure for all ensemble members of the cyclone data. The dotted lines in the cubes trace the evolution of maxima and minima values with increasing scales.

Figure 10 shows the evolution of contours for all ensemble members in scale space. We chose the variable of absolute vorticity for one ensemble of the cyclone data. The isovalue was set to 0.0001s^{-1}, and maximum scale (σ) to 25. For generality, three time steps were chosen to depict the changes over time. In comparison, spaghetti plots at three different scales are also shown to the right. Different colors are used to distinguish between ensemble members.

Using a similar approach to Figure 9 but with a different dataset, Figure 11 shows the model sound speed of all ensemble members with value 1515 m/s at three time steps of the ocean data in scale space.

From figure 9, the dotted-lines represent the local extrema values of mean sea level pressure in scale space. If a dotted-line extends continuously across the whole scale range, it implies that the key point is persistent and likely important in many applications. From those three time steps, we can see this approach can track the lowest pressure point. This is helpful for solving the challenges stated in the introduction.

From figure 10 and 11, lots of small features vanish, and dominant features become evident with scale increasing. This implies that our approach can track the main tendency of absolute vorticity at a high scale with controlled removal of many small features. This is useful for solving the first challenge: tracking important features and disguising lots of trivial features.

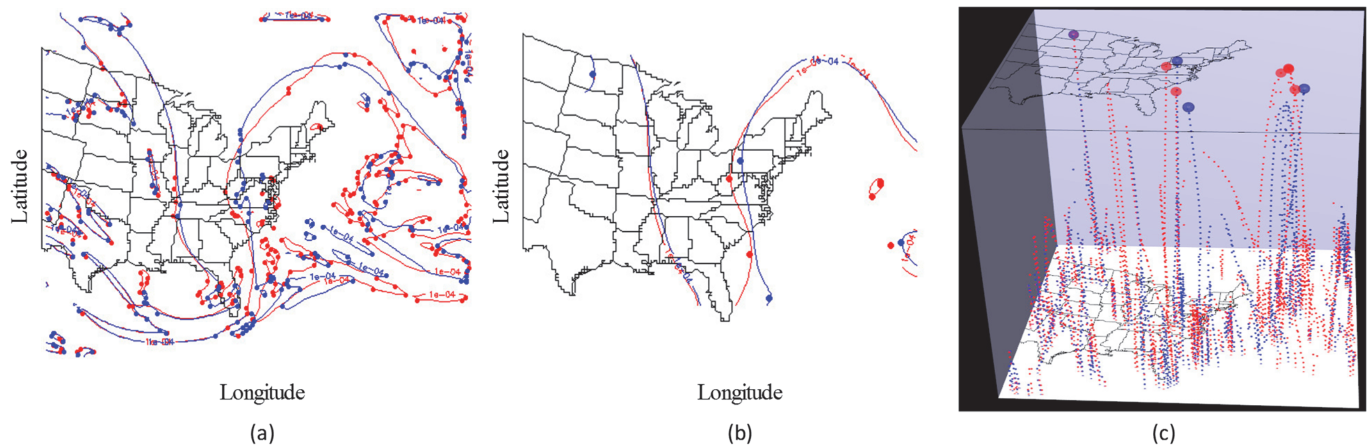


Figure 8: Contour curvature extrema evolution for two ensemble members of the cyclone data in scale space, Absolute vorticity value was used with the isovalue set to 0.0001s-1. (a) Contour curvature extrema points in low scale (b) Contour curvature extrema points in high scale (c) Contour curvature extrema evolution for the two ensemble members in scale space. Big red and blue dots represent contour curvature extrema in the highest scale. The scale increases from bottom to top.

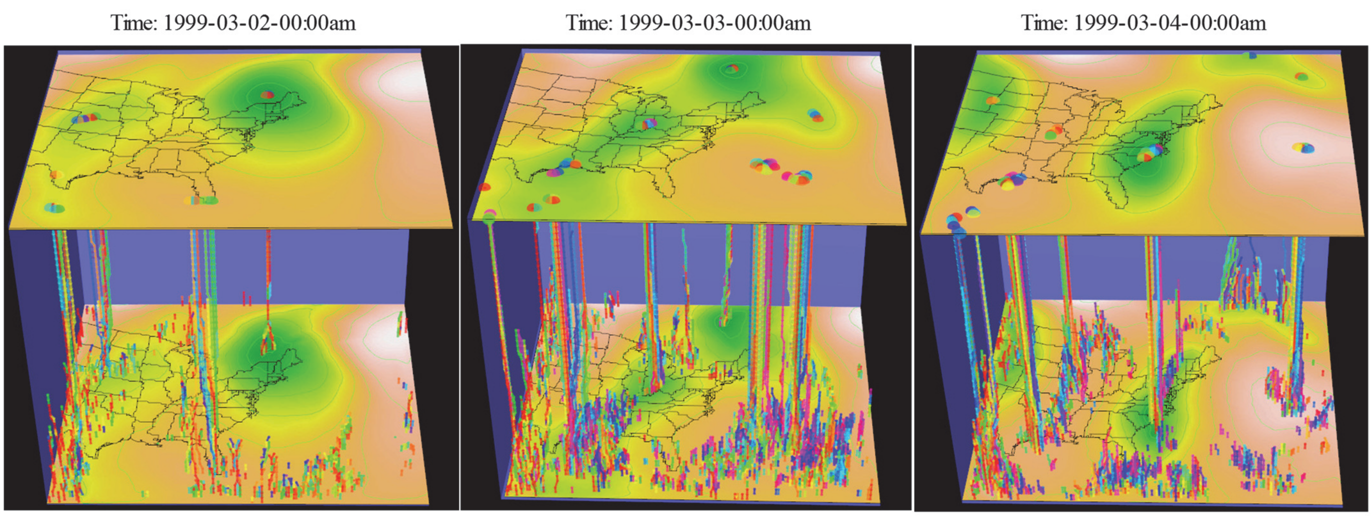


Figure 9: Extrema points evaluation in scale space. The mean sea level pressure of the cyclone data was used. Three time steps are shown in three columns. The images shown in three columns are the evolution of minima-maxima of 30 ensemble members differentiated by color. Scale increases from bottom to top. Big dots with color represent extrema in the highest scale.

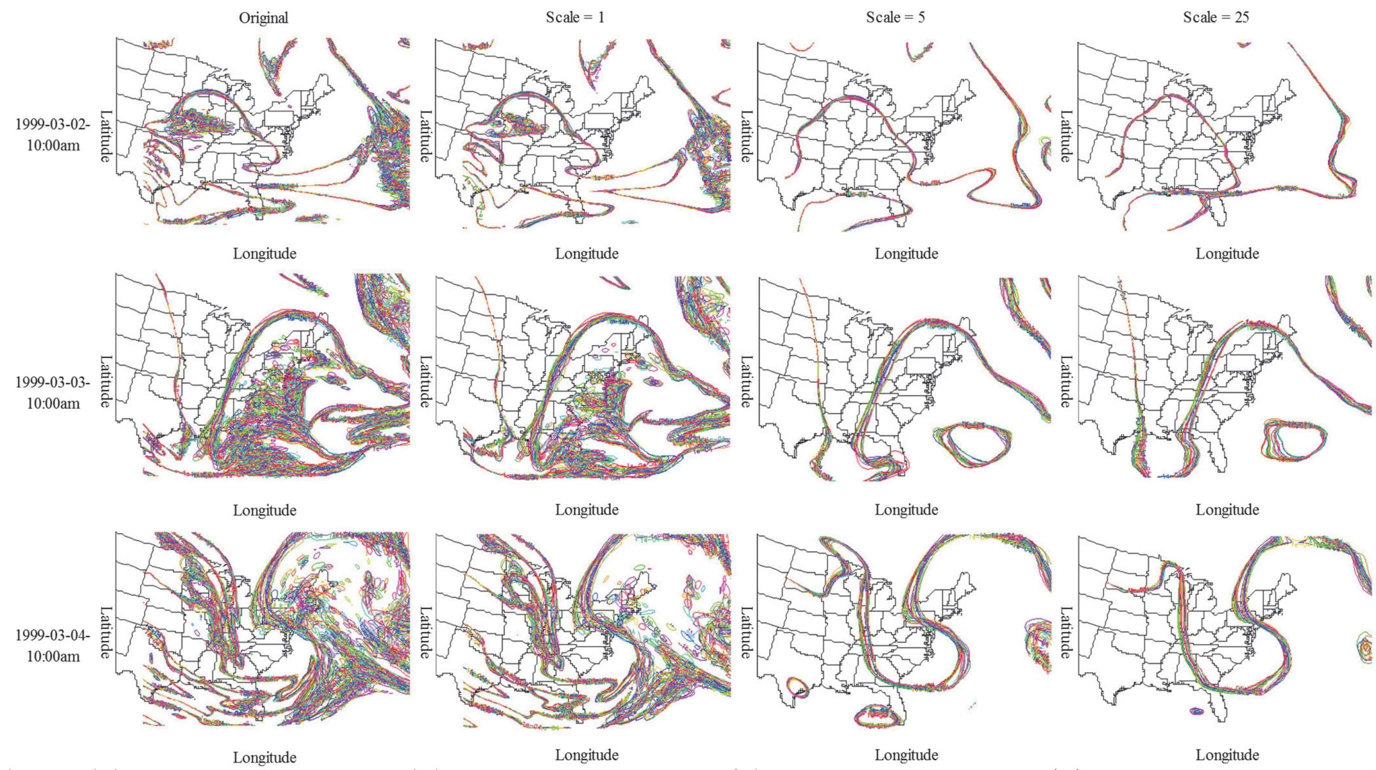


Figure 10: spaghetti visualization for all ensemble members each color represents one ensemble) of the cyclone data in scale space. Absolute vorticity value was used with the isovalue set to 0.0001s-1. Each row represents one time step. Columns are scales, increasing from left image to right image.

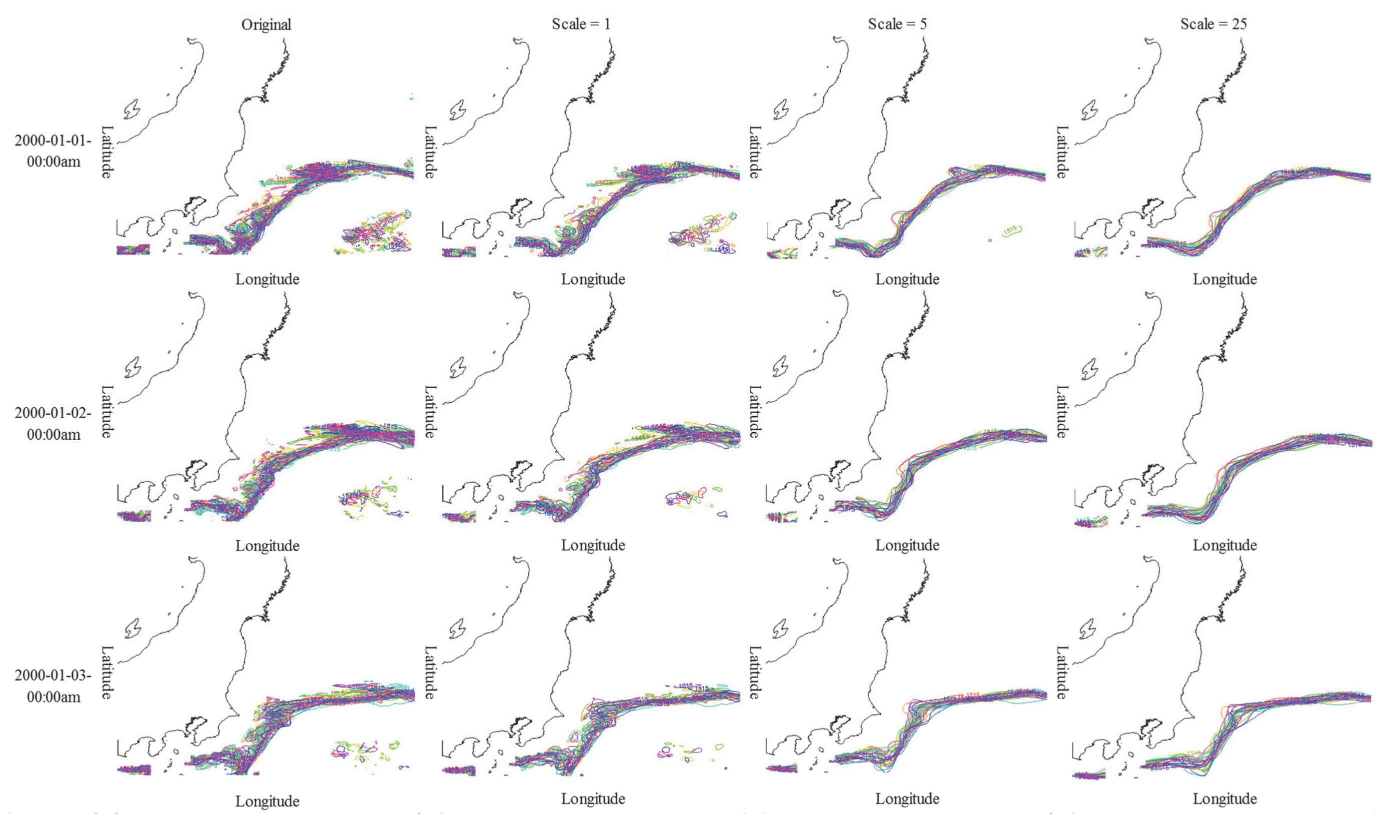


Figure 11: spaghetti visualization for all ensemble members (each color represents one ensemble) of the ocean data in scale space. Each row represents one time step. Columns are scales, increasing from left image to right image.

6 EVALUATION

To prove the effectiveness of the tool, we conducted an empirical evaluation with a meteorologist. He considers the tool to be very useful in a research environment, in particular for feature tracking and the identification of large-scale phenomena. Before the evaluation, the expert was instructed how to interact with the software and set parameters. He chose ensembles, variables, time steps, levels, maximum scale, isovalues for contours and isosurfaces to see the different outcome produced by the application. The absolute vorticity variable was of interest (Figures 8-9) due to its known large-scale characteristics in upper tropospheric flows. The initial scale space images (curvature extrema points results of Fig. 8) revealed the relative importance of shortwave features within the longwave, revealing three primary areas of activity where there are anywhere from 1 to 5 vorticity maxima (shortwaves) embedded within the longwave. These features are often important drivers of strong cyclones, and this technique depicts those features well. Additionally, he noted that, as time progressed, the 300 mb absolute vorticity contour tracked the 300 mb trough, both of which drove the evolution of the surface extratropical cyclone, and that this process was relatively well modeled by all ensemble members. This agreement between ensemble members becomes more apparent with contours at a higher scale (Figure 9). Automated feature tracking algorithms are somewhat limited in meteorology, particularly when tracking upper-level features, so this tool provides capability to do this tracking, which is important for forecasting the evolution of mid-tropospheric troughs that drive surface cyclone formation.

According to the expert evaluation, another interesting feature of the tool is the ability to diagnose the areal extent of all features within the ensemble framework, so as to establish model consistency of the simulation of the features. That is, the tool allows the meteorologist to see if the feature areal coverages are consistent among all ensemble members, which was previously not possible and could be important in forecasting applications. The extrema value points (Fig. 4, 9) supported these conclusions, showing exactly at what scale the features disappear and providing feedback specifically about how variable the shapes of the features are among the different ensemble members. In this example, the high degree of variability among the shapes suggests the model, while rendering upper-level processes effectively, is struggling with low-level features, which could be owing to the influence of the Appalachian Mountains on the cyclone at this time in its evolution. Previously, the only way this could be done was through estimation with a spaghetti plot, which is tedious, time consuming and only available at the original scale.

The expert evaluated the utility of the extrema value points tracking approach (e.g. Figure 9). Generally speaking, they noted that this approach allowed for feature tracking not only in space but also in scale and time. That is, the largest, most well defined features could be monitored through their entire life cycle in the simulation domain, while smaller scale features were able to be identified as forming and quickly dissipating. This could be anticipated by simply looking at the scale extent (vertical axis in Figure 9) of the features at any one-time step and by observing the same features over multiple time steps. As an example, in Figure 9, the large areas of high pressure (yellow) and low pressure (green) at mean sea level can easily be identified in the associated three-dimensional rendering of the local extrema points visualization. Additionally, several smaller-scale features are present in these fields that largely dissipate after a few time steps, despite the persistence of the larger-scale features. It is known that the largest scale features in meteorology seem to exist the longest, while smaller-scale features have short existence times, and this approach allows meteorologists a direct way to visualize this process precisely, which was a previously unavailable capability. Additionally, the interesting linear relationship among the ensemble members suggests certain biases in the positioning of the large-scale features by different ensemble members, which provide valuable insight about how the weather model handles feature positioning. This extrema value approach can provide research meteorologists with a new method of tracking features of different scales, accomplishing a primary research objective of this study.

The vorticity imagery provided in Fig. 6 reveals some interesting characteristics of the event's larger-scale synoptic structure. The clearest distinctions exist in the two broad areas of large absolute vorticity, which correspond strongly with deep troughs over the mid-latitudes. The presented pattern in Fig. 6a and 6d suggests one larger deep trough exists, but the long wave seen in those panels can be isolated into the shortwaves embedded within it by increasing the scale. This gives meteorologists a new way to distinguish shortwaves embedded within the larger scale wave, which have tremendous forecast implications in terms of precipitation occurrence and the formation of surface cyclones. Further, the vorticity fields reveal the well-known tilt with height of the synoptic-scale systems, particularly in panels 6c and 6f, which is a known feature that has not been visualized in this manner previously. Such additional information may be useful to establish cyclone and wave strengthening or weakening over time.

Finally, the scale selection component of the work (Figure 7) allows the user to gain an understanding of the dominant scale in place over a region. In the example figure, it provided insight into which areas are being influenced by the large-scale extratropical cyclone (denoted by the higher elevation of scale-selection surface) and which regions are largely being influenced by smaller-scale processes (areas where the surface has a lower elevation). This provides a visual rendering of the influence of the extratropical cyclone over the Appalachian mountain regions and points north, as well as the influence of smaller-scale sea-breeze and convective processes dominating atmospheric processes in the south (likely along the Gulf Coast). Such a distinction could not previously be seen and was inferred based on meteorological understanding of the processes that tended to dominate each region.

Despite the numerous positive contributions made by the software, the expert did note a few shortcomings to the approach. The three dimensional rendering makes it difficult to overlay multiple ensemble members or multiple contour levels, which requires the meteorologist to have foreknowledge of the desired contour levels and variables to consider. Most operational meteorologists have such knowledge from prior experience, but inexperienced meteorologists may face difficulties with this issue. These issues are related to the use of the tool in a forecast environment and are not a limitation of the scale-space methodology. Additionally, meteorologists are trained to consider 2D visualizations. The rendering of many of these processes in 3D will lead to increased interpretation challenges in the forecasting environment that will require additional training for end-users. The point-wise depiction of the contour curvature was interesting, but in a forecast context, the exact shape of a feature is not important (though the expert noted inconsistencies in size could suggest a relative lack of model precision). Finally, the tracking and scaling approaches worked very well in the presence of a large-scale feature, but often times such a feature is not present. The expert noted that often large-scale features are present, but some significant weather situations are not accompanied by a dominant large-scale feature and may not be properly visualized by these methods. Unfortunately, it was difficult to identify important features in the absence of a large-scale feature, which is beyond the scope of this paper, but nonetheless is being considered in future work.

7 DISCUSSION AND CONCLUSION

Several lessons can be learned from this work. First, scale space is an effective tool for reducing complexity and filtering small features in large ensemble data. Compared to topological methods, the history of the complexity reduction is recorded in scale space, allowing users to make inferences across scales. In addition, the evolution of key points through scale can be clearly traced in space. The simplicity and clarity of these point plots is often preferred by domain scientists over a complex 3D visualization.

Second, scale selection is the key to scale space data exploration. A user can apply prior knowledge to scale selection, effectively assuming that features smaller than a certain size are either noise or unimportant. Or a user may choose to explore the data with more than one scale when prior knowledge is lacking or uncertainty exists in the size of the target features. With interactive scale selection, a user can freely explore multiple scales at different locations, all in the same display. This flexibility is important for an effective data exploration. We plan to investigate automatic scale surface selection based on prior knowledge or learning in future work.

Third, the 3D data visualization in scale space remains challenging. 2D contours form a visible 3D isosurface in scale space while 3D isosurfaces become a 4D surface in scale space. How to organically show the evolution of isosurfaces in scale space is still an open problem.

In summary, several techniques are presented for extending 2D and 3D data into scale space and performing various analyses on those sets. We created extrema value points in scale space to examine how they evolve within scale space and how to make the comparison between ensemble members more evident. In addition, we implemented visualizations of multiple ensembles in scale space using contours and isosurfaces, which render the comparison of ensembles more effective.

We performed a qualitative evaluation of those techniques and generally received positive feedback. The technique presented has several unique applications, in particular, feature tracking and identification of large-scale phenomena. The tool can identify large-scale features embedded within noisy meteorological fields and reveal those features objectively and directly instead of quasi-subjectively.

REFERENCE

[1] K. Riley, D. Ebert, C. Hansen, and J. Levit, “Visually Accurate Multi-Field Weather Visualization,” *Proc. IEEE Vis. Conf.*, pp. 279–286, 2003.

[2] F. Breu, S. Guggenbichler, and J. Wollmann, “visualization of weather data,” *Vasa*, 2008.

[3] Y. Tang, H. Qu, Y. Wu, and H. Zhou, “Natural textures for weather data visualization,” in *Proceedings of the International Conference on Information Visualisation*, 2006, pp. 741–748.

[4] M. N. Phadke, L. Pinto, F. Alabi, J. Harter, R. M. Taylor, X. Wu, H. Petersen, S. a Bass, and C. G. Healey, “Exploring Ensemble Visualization.,” in *Proceedings of SPIE*, 2012, vol. 8294, pp. 1–12.

[5] A. P. Witkin, “Scale Space Filtering,” in *Artifical Intelligence (IJCAI)*, 1983, pp. 1019–1022.

[6] J. J. Koenderink, “The structure of images.,” *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.

[7] T. Lindeberg, *Scale-Space Theory in Computer Vision*. 1994.

[8] T. Lindeberg, “Scale-space theory: a basic tool for analyzing structures at different scales,” *J. Appl. Stat.*, vol. 21, pp. 225–270, 1994.

[9] A. Kuijper and L. M. J. Florack, “Using catastrophe theory to derive trees from images,” *J. Math. Imaging Vis.*, vol. 23, pp. 219–238, 2005.

[10] Z. Yang, W. Cui, and M. Zhang, “Deep structure of Gaussian scale space,” in *Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008*, 2008, vol. 6, pp. 381–384.

[11] K. Potter, A. Wilson, P. T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, and C. R. Johnson, “Ensemble-vis: A framework for the statistical visualization of ensemble data,” *ICDM Work. 2009 - IEEE Int. Conf. Data Min.*, pp. 233–240, 2009.

[12] T. Gneiting and A. E. Raftery, “Weather forecasting with ensemble methods,” *Science (80-.).*, vol. 310, no. 5746, pp. 248–249, 2005.

[13] H. Obermaier and K. I. Joy, “Future challenges for ensemble visualization,” *CG&A. IEEE*, vol. 34, no. 3, pp. 8–11, May 2014.

[14] O. S. Alabi, X. Wu, J. M. Harter, M. Phadke, L. Pinto, H. Petersen, S. Bass, M. Keifer, S. Zhong, C. G. Healey, S. Member, and T. M. Russell, “Comparative Visualization of Ensembles Using Ensemble Surface Slicing,” *Vis. Data Anal.*, vol. 8294, pp. 1–12, 2012.

[15] K. M. Smith, D. C. Banks, N. Druckman, K. Beason, and M. Y. Hussaini, “Clustered ensemble averaging: A technique for visualizing qualitative features of stochastic simulations,” *J. Comput. Theor. Nanosci.*, vol. 3, no. 5, pp. 752–760, 2006.

[16] L. Gosink, K. Bensema, T. Pulsipher, H. Obermaier, M. Henry, H. Childs, and K. Joy, “Characterizing and visualizing predictive uncertainty in numerical ensembles through bayesian model averaging,” *IEEE Trans. Vis. Comput. Graph.*, vol. 19, pp. 2703–2712, 2013.

[17] A. T. Wilson and K. C. Potter, “Toward visual analysis of ensemble data sets,” in *Proceedings of the 2009 Workshop on Ultrascale Visualization*, 2009, pp. 48–53.

[18] P. J. Rhodes, R. S. Laramee, R. D. Bergeron, and T. M. Sparr, “Uncertainty Visualization Methods in Isosurface Rendering,” in *Eurographics 2003*, 2003, pp. 83–88.

[19] A. E. C. Pece and A. D. Worrall, “A comparison between feature-based and EM-based contour tracking,” *Image Vis. Comput.*, vol. 24, no. 11, pp. 1218–1232, 2006.

[20] K. Pothkow and H. C. Hege, “Positional uncertainty of isocontours: Condition analysis and probabilistic measures,” *IEEE Trans. Vis. Comput. Graph.*, vol. 17, pp. 1393–1406, 2011.

[21] T. Pfaffelmoser, M. Reitinger, and R. Westermann, “Visualizing the positional and geometrical variability of isosurfaces in uncertain scalar fields,” *Comput. Graph. Forum*, vol. 30, pp. 951–960, 2011.

[22] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead, “Noodles: A tool for visualization of numerical weather model ensemble uncertainty,” *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 1421–1430, 2010.

[23] J. R. Holton and G. J. Hakim, *An introduction to dynamic meteorology*. Academic press, 2013.

[24] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, 1990.

[25] A. Rosenfeld and M. Thurston, “Edge and Curve Detection for Visual Scene Analysis,” *IEEE Trans. Comput.*, vol. C-20, 1971.

[26] P. Chaudhuri and J. S. Marron, “SiZer for Exploration of Structures in Curves,” *J. Am. Stat. Assoc.*, vol. 94, pp. 807–823, 1999.

[27] T. Klein and T. Ertl, “Scale-space tracking of critical points in 3D vector fields,” *Topol. Methods Vis.*, 2007.

[28] D. Bauer and R. Peikert, “Vortex tracking in scale-space,” *EUROGRAPHICS- IEEE TCVG Symp. Vis.*, pp. 233–240, 2002.

[29] C. Y. Ip and A. Varshney, “Saliency-assisted navigation of very large landscape images,” *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 1737–1746, 2011.

[30] L. Itti and C. Koch, “Computational modelling of visual attention.,” *Nat. Rev. Neurosci.*, vol. 2, no. 3, pp. 194–203, 2001.

[31] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[32] G. L. Kindlmann, R. S. J. Estépar, S. M. Smith, and C. F. Westin, “Sampling and visualizing creases with scale-space particles,” in *IEEE Transactions on Visualization and Computer Graphics*, 2009, vol. 15, no. 6, pp. 1415–1424.

[33] D. Shen, E. H. Herskovits, and C. Davatzikos, “An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures,” *IEEE Trans. Med. Imaging*, vol. 20, no. 4, pp. 257–270, 2001.

[34] H. Carr, J. Snoeyink, and M. Van De Panne, “Flexible isosurfaces: Simplifying and displaying scalar topology using the contour tree,” in *Computational Geometry: Theory and Applications*, 2010, vol. 43, no. 1, pp. 42–58.

[35] J. Michalakes, J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, “The weather research and forecast model: software architecture and performance,” in *11th ECMWF Workshop on the Use of High Performance Computing In Meteorology*, 2004, p. 13.

[36] T. Iijima, “Basic theory on normalization of a pattern (in case of typical one-dimensional pattern),” *Bull. Electr. Lab.*, vol. 26, pp. 368–388, 1962.

[37] M. Felsberg and G. Sommer, “The monogenic scale-space: A unifying approach to phase-based image processing in scale-space,” *J. Math. Imaging Vis.*, vol. 21, pp. 5–26, 2004.

[38] S. Shafii, S. E. Dillard, M. Hlawitschka, and B. Hamann, “The topological effects of smoothing,” *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 1, pp. 160–172, 2012.

[39] D. F. Watson, “Contouring,” *A Guid. to Anal. Disp. Spat. data. Comput. methods Geosci. New York, EUA Elsevier Sci.*, vol. 10, 1992.

[40] C. Maple, “Geometric design and space planning using the marching squares and marching cube algorithms,” in *Geometric Modeling and Graphics, 2003. Proceedings. 2003 International Conference on*, 2003, pp. 90–95.

[41] C. Ho, F.-C. Wu, B.-Y. Chen, Y.-Y. Chuang, M. Ouhyoung, and others, “Cubical marching squares: Adaptive feature preserving surface extraction from volume data,” in *Computer Graphics Forum*, 2005, vol. 24, no. 3, pp. 537–545.