Glyphs for Visualizing Uncertainty in Environmental Vector Fields

Craig M. Wittenbrink, Elijah Saxon, Jeff J. Furman, Alex Pang, and Suresh Lodha
Baskin Center for Computer Engineering & Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064

ABSTRACT

Environmental data have inherent uncertainty which is often ignored in visualization. For example, meteorological stations measure wind with good accuracy, but winds are often averaged over minutes or hours. As another example, doppler radars (wind profilers and ocean current radars) take thousands of samples and average the possibly spurious returns. Others, including time series data have a wealth of uncertainty information, that the traditional vector visualization methods such as using wind barbs and arrow glyphs simply ignore.

We have developed new vector glyphs to visualize uncertain winds and ocean currents. Our approach is to include uncertainty in direction and magnitude, as well as the mean direction and length, in vector glyph plots. Our glyphs show the variation in uncertainty, and provide fair comparisons of data from instruments, models, and time averages of varying certainty. We use both qualitative and quantitative methods to compare our glyphs to traditional ones. Subjective comparison tests with experts (meteorologists and oceanographers) are provided, as well as objective tests (data ink maximization), where the information density of our new glyphs and traditional glyphs are compared. We have shown that visualizing data together with their uncertainty information enhances the understanding of the continuous range of data quality in environmental vector fields.

Keywords: verity visualization, wind profilers, doppler radar, wind barbs, uncertainty glyphs

1 INTRODUCTION

Visualization is used to display large amounts of data, and to gain an understanding of the data. Generally, data have associated characterizations of quality or uncertainty resulting from the collection or processing. While the uncertainty is an essential part of the data, it has often been ignored while processing or displaying. There is a need to display the original data together with their *uncertainty* for accurate interpretation. There are several important questions in visualizing the uncertainty in environmental vector fields, including: How are different forms of uncertainty represented in vector fields? How are inaccuracies introduced in interpolations used in vector field visualization techniques? How can we visualize vector field uncertainty? And, how can we combine or multiplex uncertainty into standard vector field visualizations to improve understanding?

In a single dimensional data plot, such as a time series scalar plot, the uncertainty can be graphically represented with a glyph that shows the median, quartile, minimum, and maximum values. Two examples of such glyphs are Tukey's box plot[1] and Tufte's quartile plots.[2] These glyphs can graphically present the distribution of the sample point's value with economy. When going to higher dimensional data the box plot becomes unwieldy.

This paper is focused on different ways of mapping uncertainty parameters to visual cues in vector field visualization. For our work, we use a tuple to represent the dimensions of uncertainty. The challenging and novel aspects of this research are the integration of the data and its uncertainty vector to form a spatially and temporally accurate depiction; and the generalization of vector visualization methods to include uncertainty.

Our visualization work involves atmospheric and oceanographic data.[3–6] We show both simple techniques for visualizing uncertainty, the overloading approach, and new glyphs which do not require overloading to directly show the variation in bearing and magnitude. We call this non-overloading approach *verity visualization*, which while not always possible is a significant improvement, because it leaves overloading (color, transparency, and so on) for other variables.

The effectiveness of our glyphs are measured both by quantitative and qualitative measurements. Quantitative metrics include such metrics used in Tufte's principles of data–ink maximization,[2] information per unit of space, information per unit of ink (number of colors), number of multi–functioning graphical elements, and the data density for the entire graphic. These methods are used to objectively evaluate the effectiveness of placing larger and larger amounts of data into a visualization with economy of expression. Qualitative evaluations of the graphics by application specialists are possible, through our collaboration with MBARI, NPS, Long Marine Lab, and NOAA in the Monterey Bay Marine Sanctuary. This user evaluation provides important feedback, and indicates general trends in the effectiveness of our uncertainty visualizations.

Visualization without uncertainty is unreliable and misleading. We show that combined data and uncertainty visualization allows more precise interpretation, and scientific visualizations with combined uncertainty may be shown to be superior with quantitative metrics. Quantitative metrics may predict user interpretation, and visualizations without uncertainty are a special case of visualizing uncertainty, where uncertainty goes to zero. We believe visualizing uncertainty will significantly improve understanding of environmental phenomena, advancing the state of the art of scientific visualization. Our new uncertainty glyphs for environmental data may generalize to other scientific data visualizations.

2 BACKGROUND

Many definitions of uncertainty have been proposed.[7–10] Uncertainty is a multi-faceted characterization of data (value, range/domain–time, space, arbitrary dimensions) that encompasses many concepts including: accuracy, application, autocorrelation, confidence level, confidence intervals, covariance, difference, distribution, drop-off rate, error (acquisition, transformation, visualization, analysis), heuristic (expert), inaccuracy, lineage, logical, minimum/maximum, noise, precision, purity, purpose, quality, randomness, reliability, residual, resolution, spread, standard deviation, standard error, time, validity, and variance.

Error can be defined as the discrepancy between a given value and its true value.[10] Inaccuracy is the difference between the given value and its modeled or simulated value.[10] Data validity encompasses both the inaccuracy of the data and the procedures applied to the data. Data validity is measured by deductive estimates, inferential evidence, data consistency and comparison with independent sources, and it is ratified by testing.[10,7] Data quality is treated as an even more general term that includes data validity and data lineage. Data lineage refers to those characteristics of data such as collection circumstances and pedigree. Data quality can be defined as a three parameter variable, that consists of goodness or statistical measure, application or model resolution, and purpose such as analysis or communication.[8] Due to the importance of the presence of uncertainty in scientific data, NIST[9] has recently formulated guidelines for expressing uncertainty of data measurements. Data uncertainty, although consisting of several components, can be broadly classified into two categories according to the method used to estimate their numerical values: evaluation by statistical methods such as standard deviation or least squares, and evaluation by scientific judgement. The NCGIA initiative on "Visualizing the Quality of Spatial Information"[8] classified the sources of data uncertainty as source errors, process errors, and use errors.

Although advances have been made in defining and deriving uncertainty for data collected from instruments, the identification of the occurrence of uncertainty and its distribution in the visualization pipeline is crucial in identifying the effect of uncertainty on data interpretation. We illustrate the different sources of error that arise from the transformations that are applied from the data collection through the visualization. Figure 1 shows a data pipeline starting with the data *acquisition* where the physical phenomena are captured and recorded either through sensors or as output from numerical models. These measured phenomena may undergo *transformations* to produce derived data. Typical transformations include interpolation and sampling. The derived data are processed by *visualization* algorithms that generate images for the users to *analyze*. Data uncertainties are introduced and propagated at every stage in the pipeline. Measurement errors due to equipment limitations, calibration, effects of the environment, and others are introduced in the acquisition stage. Model errors result from model simplification, initial and boundary conditions, and stability in numerical algorithms. Due to the limitations of the devices or simulations used in the data acquisition stage, the resulting data can be sparse and possess less than ideal spatial and temporal measurement resolutions. These measurements typically undergo some form of transformation such as interpolation or approximation. Again, errors are introduced.

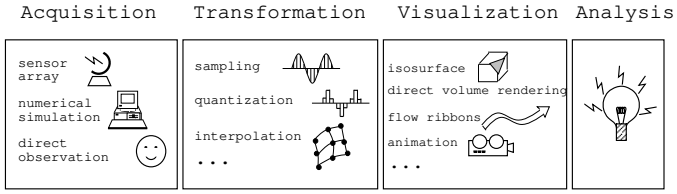

Figure 1: Data Pipeline

During visualization, the representation of uncertainty is followed by the challenging task of mapping uncertainty to visual cues. The task is complicated for several reasons. First, which aspect of the different parameters that constitute uncertainty is to be visualized? Second, which visualization primitives and rendering algorithms are to be utilized? These steps may introduce their own sources of errors due to sampling, quantization or interpolation, which may lead to aliasing artifacts for example. Compounding the problem, the errors at each stage may propagate and influence succeeding stages in the data pipeline. Identifying the sources of uncertainty and their distributions is critical in assisting the scientists to make accurate interpretations of the data.

2.1 Previous Work

Most of the work in visualization of data with uncertainty is in the field of Geographic Information Systems (GIS). See survey of methods in Hunter et al.[11] Surprisingly, little work in visualization of data uncertainty has been pursued in other areas, but we briefly survey the variety of uncertainty visualization techniques that we are familiar with below. For the following, we define glyphs as symbols that represent data through visual properties such as color, shape, size, and orientation. Glyphs are also called probes, geometrical primitives, stars, and boxes, but not icons. Icons are images or graphical elements that represent concepts, objects, or actions typically within a user interface. Specific examples of glyphs are vector arrows and wind barbs.

Glyphs are used to represent univariate data in Tukey,[1] Tufte[2],[12] and Cleveland.[13] Glyphs can also represent multivariate data shown by Cluff et al.'s survey[14] of stars, Chernoff faces, boxes, profiles, Kleiner-Hartigan trees, and Andrew's plots. Newer glyphs or probe techniques for vector and tensor fields are shown in de Leeuw et al.[15] Displacement vectors can be used to display difference or errors between two images such as in Peterson.[16] However, we have not found any glyphs used to display uncertainty information in vector fields.

A direct approach to visualizing uncertainty is to map it to different graphics attributes. Uncertainty can be seen as an additional variable(s) which is input to one of the properties of the graphic. This could be done with the multivariate glyphs just mentioned. Graphics attributes include: color, transparency, and line

width. Thus, examples that fall under this category include: varying contour widths depending on certainty,[17] mapping uncertainty parameters to different points in HSV space,[18] using cross hatches,[19] and our work in using transparency to indicate confidence in an interpolated field[3] to name a few. Most of the previous work on visualizing uncertainty is focused under this category.

Other graphics and rendering techniques may be used for visualizing uncertainty. Methods include overlaying (sandwich layers) or side-by-side comparisons of data and uncertainty information (Wills et al.[20]), pseudo-coloring of difference and error images such as our volume rendering comparisons,[21] haziness corresponding to uncertainty (Beard et al.[8]), and defocussing or Monte Carlo blurring (Fisher[22]). Animation by playing back sequenced images (Monmonier[23]), random dots (Fisher[24]), and segmenting and blurring (Gershon[25]) have been used to display data uncertainty. Even sonification has been proposed by Fisher[24] for perceiving uncertainty. Our glyphs typify a new class of technique we call *verity visualization* that we will describe shortly.

3 METHODOLOGY AND RESULTS

Our research on visualizing uncertainty in environmental vector fields is organized into four parts: collecting and characterizing different data, defining and deriving uncertainties, visualizing data with uncertainties, and evaluating the new visualization methods. The primary focus of the paper is on creating new visualizations for data with uncertainties (Section 3.3). While data acquisition and evaluation phases are essential ingredients, the secondary focus of the paper is the derivation of uncertainty from interpolation of noisy and scattered data (Section 3.2).

3.1 Data sources

Our REINAS (Real-time Environmental Information Network and Analysis System)[5] project has uncovered challenges in dealing with uncertainty from both instrument and numerical model data. The data that we investigated are from instruments, numerical models, and interpolation. Instrument data sources include radars, that measure wind,[26,27] or ocean surface currents, meteorological stations (wind), and sonar buoys (currents). We have access to output from numerical models such as the Navy Operational Regional Atmospheric Prediction System (NORAPS)[28] and the Princeton Mellor Model.[29] The quality of these data depend on the simplifying assumptions that were made in creating the model, numerical accuracy, spatial and temporal resolutions, and initial and boundary conditions. Typically the meteorological station and buoy data are sparsely sampled and contain fixed uncertainties, while the radar and circulation models are denser, gridded outputs with varying uncertainties. The last type of data is interpolated from the measured and simulated fields.

3.2 Data uncertainties

Figure 2 shows where uncertainties are introduced in the data pipeline shown in Figure 1. The synthesis of these uncertainties, along with the measured, derived, and vector visualization is the goal. Uncertainty for measured data are provided by instruments[30,27] or from the estimated accuracy of the instrument. As an example, the wind profiler measures wind velocity in three directions by sending out three beams. The radar processing from the beams outputs statistical information, including the signal to noise ratio, width of the radar signal return beam, and the maximum strength of the signal. These features are shown in the example radar spectral returns in Figure 3. By using thousands of samples from the same beam over a minute, one may take a consensus of the returns over an hour, the radar calculates a spectra for many heights as shown in Figure 3. The spectral data may

be averaged over ten minutes or an hour, and then we may calculate a derived uncertainty for the hour, (figure 4). We use the distribution, standard deviation, mean, minimum, and maximum as components of uncertainty for each beam's radial velocity. Three radial velocities are then used to calculate the planar wind velocity, and the magnitude and directional uncertainty.

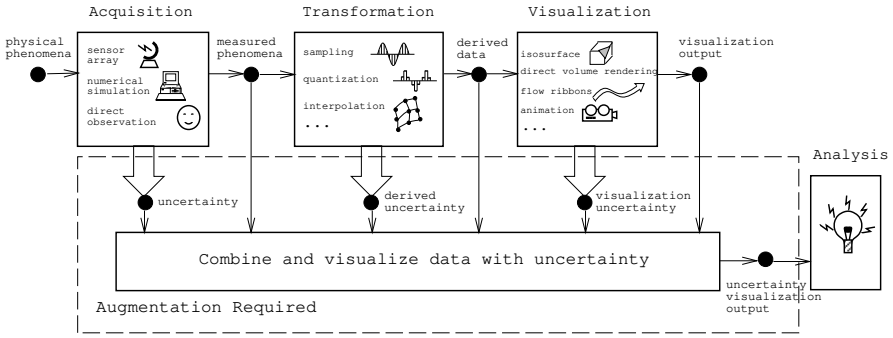

Figure 2: This augmented visualization pipeline shows measurement uncertainty, derived uncertainty, and visualization uncertainty. Our glyphs attempt to multiplex these uncertainties with the data sources from the visualization pipeline.

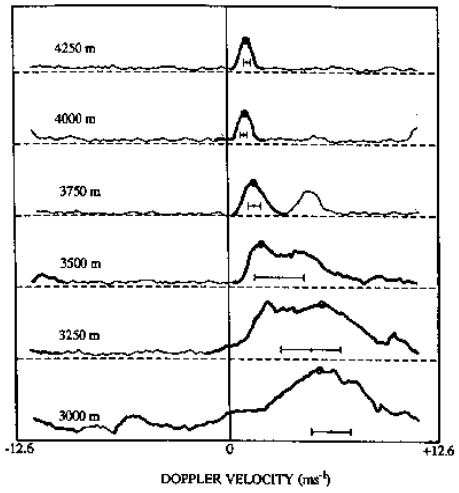

Figure 3: Spectral samples of wind velocity from a vertical wind profiler. The plot for each height level shows the frequency versus magnitude of signal return. Lower elevation returns show wider spreads indicating contamination from migrating birds and higher uncertainty (From[30] Figure 5).

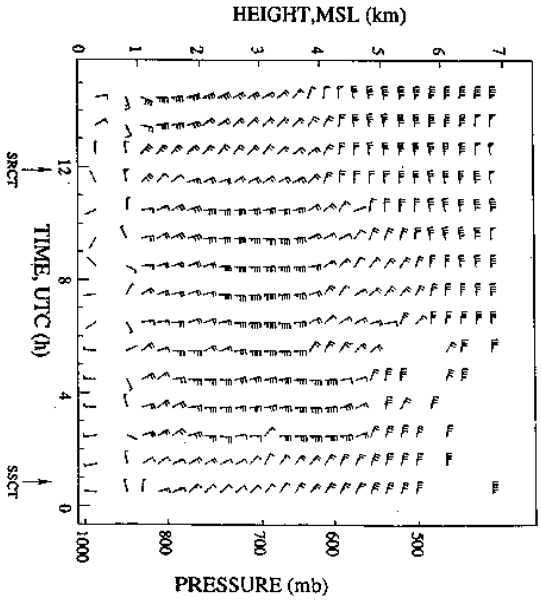

Figure 4: Time versus height display of hourly consensus-averaged wind using wind barb glyphs. Data is from vertical wind profiler measured at Lamont, Oklahoma on 26 September 1993. (From[30] Figure 9).

As another example, the meteorological data collected by the stations do not have any indication of uncertainty for each data point, other than an expected accuracy of the sensors. Time averages and statistics, however, are used similar to the wind profiler processing, to calculate the distribution of measurements over a longer time interval than the basic sampling rate. From this distribution the standard deviation, mean, minimum, and maximum are computed giving uncertainty for time averaged data.

For model data the uncertainty is obtained by multiple runs of the model, either with different starting

conditions, or with different grid spacings. The sparseness of much of the environmental data requires interpolation and extrapolation to assist scientists in visualizing the effects of environmental phenomena. Interpolated data provide a derived uncertainty from the measured data, and in many cases are acceptable to use for scientific visualization. Our early work focused on interpolation uncertainty using weighted interpolation from sparse meteorological stations.[3] The radar data provide an additional challenge, in that fields of data are returned from several radar sites, so that a maximum likelihood algorithm can calculate a best match field. Specifically, Codar Systems Inc. radars (SeaSonde (TM))[5] are now installed in three positions around Monterey bay which intersect, and provide three measurements of the ocean surface current. The best ocean surface currents are those that are measured at orthogonal directions from at least two sites, but the nearly nonorthogonal sites may also be useful if sufficient uncertainty information is provided. We have done some work to characterize the interpolation effect on data visualization, using a linear falloff assumption. (See example 3 in Section 3.3.) Our contribution is the definition of the augmentation required for uncertainty visualization, as shown in figure 2, and our new glyph for doing uncertain vector visualization to be described in the next section.

3.3 Visualizing data with uncertainty

In order to contrast our new visualizations, we first describe a common technique for visualizing uncertainty. Uncertainty can be an additional variable associated with the data and then one may use existing multidimensional visualization techniques. The additional visualization parameter tied to the uncertainty could be shading, normals, height, color, or some other aspect of rendering. With this overloading approach, using multivalued visualization, the user may still be able to confuse the data and the uncertainty information.

At the other extreme, we have also found that instead of treating uncertainty as an additional piece of data it can also be truthfully represented through *verity visualization*–in that it suggests the quality of the data that is exactly what it purports to be or is in complete accord with the facts. In this case, the uncertainty information is integrated with the data into the visualization graphic so that users cannot help but interpret the resulting image holistically. It should be noted that treating uncertainty information as additional data to be visualized or treating them through verity visualization is a continuum rather than a distinct dividing line in some instances.

The wind vector can be denoted in a visualization with a glyph, a symbol indicating the strength and direction of the wind. In meteorology, wind barb glyphs are used that encode the speed, figure 4. In many visualization tools, the glyphs are simply line segments whose lengths are scaled to the vector magnitude. In more sophisticated packages, the arrow heads are added to indicate the directions from which, or into which, the vector is flowing. There are several options available to indicate uncertainty associated with the wind vector. The most obvious is to overload the graphics attributes of the wind barbs such as with pseudo color and transparency. Alternatively, uncertainty may be shown with a color map. However, the resulting images still require a separation of interpretation between the two. In addition, to indicate uncertainty with overloading, one can plot the vector field with glyphs, and the uncertainty in the vector with a pseudo coloring of the image. Or one may color the glyphs themselves with the amount of certainty they have. Shading, transparency, and uncertainty contours, may all be used, if uncertainty is viewed as another datum to be plotted. But there is still a separation of the interpretation. One also loses the flexibility of using those shadings, transparencies, or contours to plot additional unrelated variables such as temperature or elevation.

Example 1 (Overloading Approach): Codar SeaSonde (TM), doppler current radars measure back scattering from the ocean surface, and using the doppler shift, calculate the speed the wave is travelling towards or away from the radar. Combining several radars' data, one can calculate ocean current surface vectors. The Codar fields shown in this paper are from November 24, 1994, 00:00 GMT, and were collected in Monterey Bay using a radar at the North end of the bay and one at the South end of the bay. Another illustration where new visualization strategies may help users identify uncertainty is with the use of bump mapping which is investigated in our companion paper.[4] Figure 14 shows a simple example of pseudo-coloring the magnitude uncertainty to the color

of the vector, and pseudo-coloring the ocean surface to the angular uncertainty. This figure was generated by our Spray rendering software.[6] This overloading gives the viewer complete knowledge of the uncertainty, but other data fields cannot be mapped to the coloring.

Example 2 (Verity Visualization Approach): To illustrate the point that uncertainty be treated through verity visualization, we describe a new uncertainty glyph which integrates vector information together with directional and magnitude uncertainty. After the examples we present the glyph design in section 4, and evaluation in section 5. Our glyphs are verity visualizations, because data overloading can be used to place even more variables into the visualization such as colored glyphs. The glyphs are arrows that indicate direction and magnitude. The width of the arrow heads indicate the range of possible directions at that location. Multiple arrow heads indicate the range of possible magnitudes. The process of combining the Codar sites' radial velocities provides an uncertainty in both magnitude and direction. Different methods for combining radial currents may be used with varying results. In figures 5 through 9 we show two methods (Method I and Method II) for combining radials from two sites, with the same radial data. The combining process is an averaging algorithm with a fall off in weighting of radials to create a combined vector. Due to space limitations we do not detail codar vector calculation Method I and II. Figure 5 and figure 8 display arrow glyphs. Figure 6 displays uncertainty glyphs with magnitude scaled to the length, and figure 7 and figure 9 display uncertainty glyphs using magnitude scaled to the area. Both methods of uncertainty may be useful, as the length scaling emphasizes uncertain areas, and the area scaling gives a more accurate depiction of current strengths.

The sparse current vector field in figures 5 to figure 7 (Method I) is helpful in describing the utility of the new glyph. In figure 5 no uncertainty is indicated, and we must take each arrow as equally valid. In figure 6, the highly uncertain vectors are highlighted showing that many of the measurements have from 20 to 160 degrees of angular uncertainty, and appreciable magnitude uncertainty shown by the upper deviation arrow heads. The glyphs in figure 6 are the same length as figure 5's arrows, and highly uncertain measurements are very large which could be perceived as magnitude.[2] Figure 7 shows uncertainty glyphs whose areas are proportional to the magnitude, which greatly reduces the size of the uncertain glyphs, and the perceived area is the mean magnitude. Comparing figure 5 and figure 7 one may see which vectors have uncertain measurements, get a good indication of the flow, and have an accurate perception of the relative current magnitudes.

Figure 8 and figure 9 (Method II) show the second method of combining Codar radial currents for the same time period. It's obvious that the radial combining method two gives much denser fields. The uncertainty calculated is only due to the combining process, though, and doesn't carry through all of the uncertainty that we discussed earlier. The vectors towards the edge of the fields can be seen to be more uncertain. It is also useful to compare figure 14 to figure 9 to contrast the overloading approach with the verity approach for the very same data.

Example 3 (Verity Visualization of Interpolated Winds): At meteorological stations, wind measurements are gathered at a one second sampling rate. Because at the scale of interest in Monterey Bay, these measurements are very sparse, we do an interpolation to provide a denser field of winds. The use of these winds follows an increase in uncertainty the further you get from a meteorological station. We use a linear falloff, and arrows versus angular uncertainty glyphs are shown in figures 15 and 16.

Figure 4 shows how wind vectors are typically represented as wind barbs. Such a representation of hourly averaged winds, gives no indication of uncertainty which we know is present from the spectral data, figure 3 In Figure 16, we show our new glyph, which changes depending on the uncertainty in direction. The wind speed is mapped to the area of the glyph. Using this glyph, the uncertainty is apparent within each vector presented. The viewer cannot detach the uncertainty from the vector itself. This makes the data analysis simpler. In addition, in comparison with Figure 15, we would argue that there is more information with a comparable amount of ink used, similar clutter, and a clearer understanding of the vector field. In section 4 we detail the design of the glyph, and in section 5 we evaluate the glyph's performance.

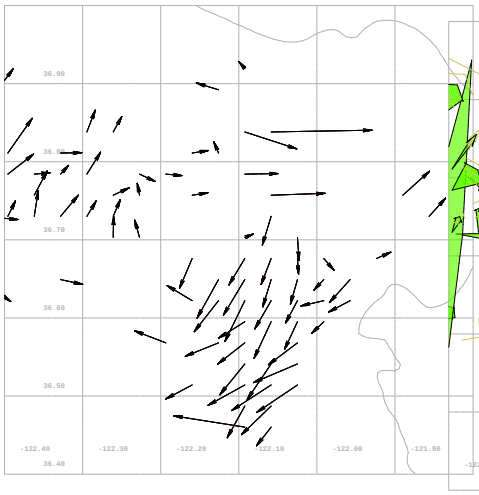

Figure 5: Arrow glyphs, length scaled to magnitude, Codar vector calculation Method I, Wed. 24 Nov 1994

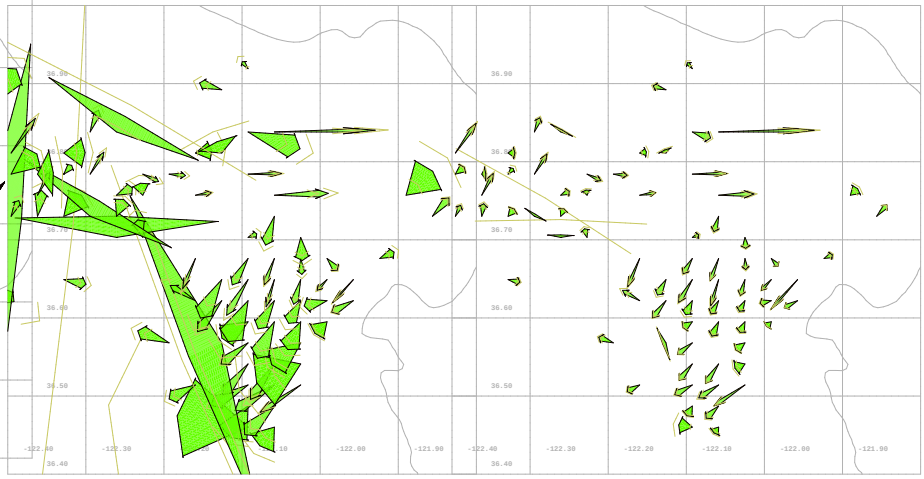

Figure 6: Uncertainty glyphs, length scaled to magnitude, Codar vector calculation Method I.

Figure 7: Uncertainty glyphs, area scaled to magnitude, Codar vector calculation Method I.

4 GLYPH DESIGN

Glyphs have an infinite number of variations in design. The designs that we have developed focus on the ability to clearly display uncertainty in magnitude and direction. Important to glyph design, is that they be evaluated in dense displays. A glyph which is apparently quite effective in isolation does not necessarily work well in a dense hedgehog style plot with 100's of glyphs. This perceptual element in glyph design can also be extended to graphics attributes. For example, line or surface thickness can be mapped to uncertainty levels. However, thicker or bolder lines also stand out more visually and tend to emphasize rather than de-emphasize fuzzy areas of the data as shown by the length scaled glyphs in figure 6 compared to figure 7. On the other hand, thicker lines or surfaces may be counter-balanced by lowering the brightness level. Likewise, textures can also depict areas of uncertainty, but we group these effects into the overloading approach.

Figure 11 shows a variety of possible glyphs to encode uncertainty. Whether glyphs encode uncertainty in magnitude or angle (figure 10) or both is indicated in the columns. While each glyph looks promising in isolation, placing them into a large field gives different results. We have also investigated the ability of the glyphs to encode uncertainty in different coordinate frames. The ability to have a single coordinate frame for uncertainty which subsumes all other possible coordinate frames for representing the vectors is important, so that the uncertainty has no secondary effects from the basis used to represent them. Shown in Figure 12 is the fact that uncertainty in directions x and y (dx, dy) can be converted to uncertainties in magnitude and direction, but the (x,y) breakdown is coordinate space dependent. It is for this reason that all of our vector uncertainties and glyphs represent the uncertainties in deviation in direction and deviation in magnitude.

We now show how to draw one of our most promising glyphs. We have reduced the drawing to first computation of a few parameters from area as magnitude, or length as magnitude. This determines the size of the glyph. We calculate the glyph oriented pointing at zero degrees, or to the right, and draw them in a counter clockwise fashion as shown in Figure 13, vertex 0, 1, 2, 3, 4, 5, 0. The glyph body coordinates to solve for are the main body length l_b, the arrow head length, l_a, the height from the baseline, h, and the winglets of the arrowhead.

The parameters that are given are the magnitude, m, the ratio of the head to the body, $r_{a/b} = l_a/l_b$ (or the

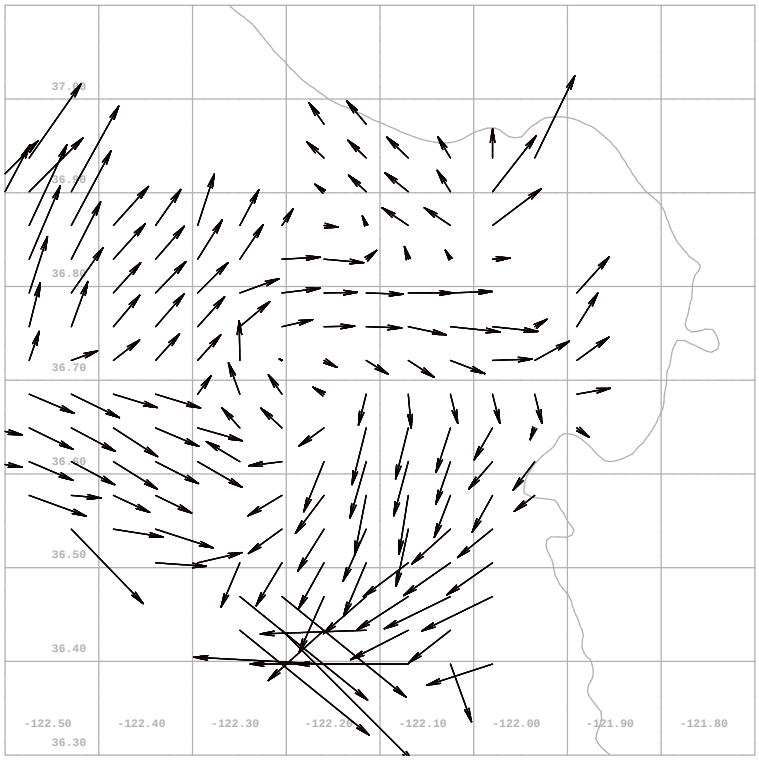

Figure 8: Arrow glyphs, length scaled to magnitude, Codar vector calculation Method II.

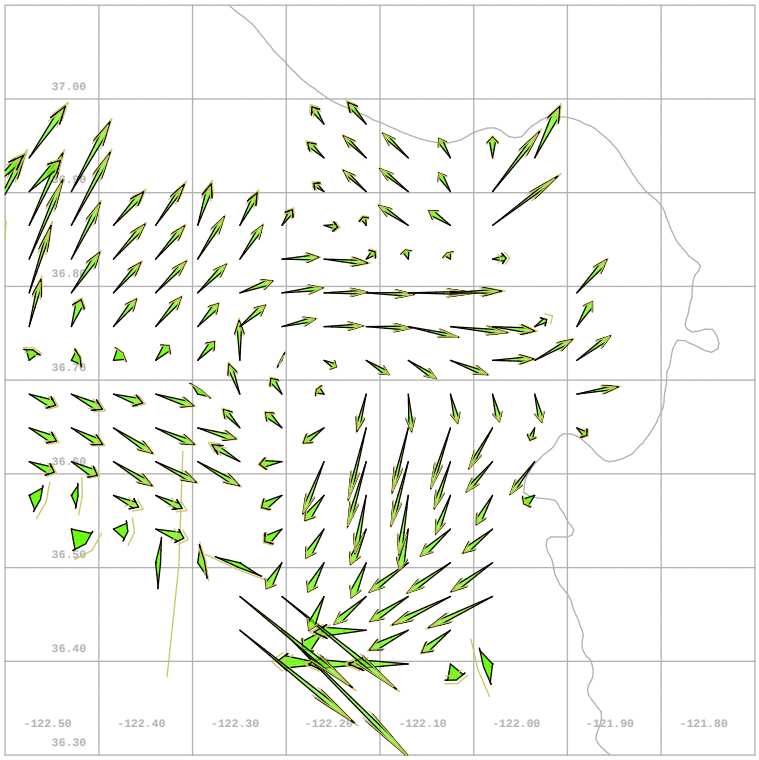

Figure 9: Uncertainty glyphs, area scaled to magnitude, Codar vector calculation Method II.

Figure 11: Variety of Glyphs Considered for Magnitude and angular uncertainty.

5 GLYPH EVALUATION

There are qualitative and quantitative methods to evaluate the effectiveness of any new visualization technique. We use both of these techniques. The quantitative approaches are less subjective and therefore simpler to use. We use the methodology proposed in Tufte's work[2,12] which consists of evaluating the visualization by the following metrics:

Figure 13: Glyph parameters.

1. Data–ink maximization. The ratio of data–ink (ink devoted to representing data information) to total ink (ink used to print the graphic) is maximized. Ink may correspond to the number of colors in color graphics.
2. Clutter or confusion minimization. The non–data–ink ratio, or the amount of unnecessary ink used can also be examined. Cluttered graphics have overlapping and/or confusing symbols that prevent seeing the data.
3. Moire pattern minimization. Moire vibration is caused by evenly spaced bars, lines, and cross hatches that make a graphic swim.
4. Multifunctional graphics elements. How many data values a symbol encodes is its multifunctionality. There is a danger of a graphic becoming confusing when overloaded with too much information.

One aspect of scaling the area of the glyph to the magnitude of the vector field, is that people perceive area better than length,[2] and therefore the perceived glyphs correlate to the true strength of the vector field. A side benefit of the glyphs, scaled by area is the fact that glyphs showing uncertainty, and those not showing uncertainty will have exactly the same area. This gives us a constant *ink* amount for plots of the same vector fields with each different glyph. The information content is greater however, because not only are we plotting *magnitude* and *bearing*, but we include in the uncertainty glyphs *bearing–deviation* for no increase in ink. Therefore the data to ink ratio improvement is $(2 * value/ink - -area)$ to $(3 * value/ink - -area)$ for a 3/2 improvement. When the magnitude deviation is also shown, there is a small increase in the amount of ink used for the upper deviation. The amount of increased ink usage is also analytically computable, but depends on the ratio $r_{a/b}$. In this case the *magnitude* and *bearing*, *deviation bearing*, and *deviation magnitude* are shown in the plot with a negligible addition in the ink used for approximately a factor of 2 improvement in the quantitative data/ink maximization ratio.

We have also done qualitative evaluations. Some of the qualitative metrics are clearly applied, others are not. The clearly applied ones are those which many people can easily agree on. The cluttered or confusing graphic and multifunctional graphics elements discussed above are partly qualitative evaluations, which are easier to agree upon. For other less clearly applied tests, we have taken advantage of collaborating organizations to gain access to meteorologists, oceanographers and computer graphics professionals who evaluated our techniques. We used comparison tests, where a traditional graphic from each field is compared to a graphic using one of our uncertainty techniques, and asked, one, whether the understanding of the data is improved; two, which graphic is preferable if only one could be retained; and, three, what is better about each graphic. A full scale statistically based survey would be an interesting follow on study to our research.

6 SUMMARY AND CONCLUSIONS

Scientific data from instruments, numerical models, or interpolation schemes almost invariably contain some degree of error or uncertainty. Display of such scientific data without uncertainty information can be misleading and may lead to erroneous conclusions. Visualization of data with uncertainty information allows more precise interpretations.

In this paper we showed scientific data collected from different sources, derived uncertainty information, and presented several new uncertainty visualization techniques. We also showed qualitative and quantitative evaluations of our new visualizations. Since no single technique of visualization works well for all data and applications, experiments are still needed to identify effective visualization strategies that work well in a given context. We do believe, however, that our new glyph visualization techniques will prove quite valuable when data analysts are looking into the validity of their data, and are confident that they may be proven superior, and come into more common use because of their ease of understanding, and improvements in data density and information presentation.

7 ACKNOWLEDGMENTS

We thank Dr. Daniel Fernandez for his extensive help, advice, and uncertain vectors. We thank Professor Wendell Nuss and Dr. Paul Hirschberg for their support in providing NRL's NORAPS output data from the Naval Postgraduate School. And we also thank all of those involved on the REINAS project. This project is supported by ONR grant N00014-92-J-1807.

8 REFERENCES

[1] John W. Tukey. *Exploratory Data Analysis.* Addison Wesley, 1977.

[2] Edward R. Tufte. *The Visual Display of Quantitative Information.* Graphics Press, 1983.

[3] Alex Pang, Jeff Furman, and Wendell Nuss. Data quality issues in visualization. In *SPIE Vol. 2178 Visual Data Exploration and Analysis*, pages 12–23. SPIE, February 1994.

[4] Alex Pang and Naim Alper. Bump mapped vector fields. In *SPIE & IS&T Conference Proceedings on Electronic Imaging: Visual Data Exploration and Analysis.* SPIE, to appear Jan 1995.

[5] P.E. Mantey et al. REINAS: Real-time environmental information network and analysis system: Phase III - systems design. Technical report, CIS Board, University of California, Santa Cruz, 1994. UCSC-CRL-94-08.

[6] A. Pang. Spray rendering. *IEEE Computer Graphics and Applications*, 14(5):57 – 63, 1994.

[7] H. Moellering. The proposed standard for digital cartographic data: report of the digital cartographic data standards task force. *The American Cartographer*, 15(1), 1988.

[8] M. Kate Beard, Barbara P. Buttenfield, and Sarah B. Clapham. NCGIA research initiative 7: Visualization of spatial data quality. Technical Paper 91-26, NCGIA, October 1991.

[9] Barry N. Taylor and Chris E. Kuyatt. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. Technical report, NIST Technical Note 1297, January 1993.

[10] M. Goodchild, B. Buttenfield, and J. Wood. *Visualization in Geographical Information Systems*, chapter Introduction to Visualizing Data Validity, pages 141–149. 1994.

[11] Gary Hunter and Michael Goodchild. Managing uncertainty in spatial databases: Putting theory into practice. *URISA Journal*, 5(2):55–62, 1993.

[12] Edward R. Tufte. *Envisioning Information*. Graphics Press, 1990.

[13] William S. Cleveland. *The Elements of Graphing Data*. Wadsworth, 1985.

[14] Elizabeth Cluff, Robert Burton, and William Barrett. A survey and characterization of multidimensional presentation techniques. *Journal of Imaging Technology*, 17(4), 1991.

[15] Willem C. de Leeuw and Jarke J. van Wijk. A probe for local flow field visualization. In *Proceedings of Visualization 93*, pages 39–45. IEEE Computer Society Press, October 1993.

[16] Jesper Kaae Petersen. Map processors. In *GIS/LIS '89*, pages 134 – 142. American Congress on Surveying & Mapping, November 1989.

[17] Geoffrey Dutton. Handling positional uncertainty in spatial databases. In *Proceedings 5th International Symposium on Spatial Data Handling*, pages 460 – 469. University of South Carolina, August 1992.

[18] Yee Leung et al. Visualization of fuzzy scenes and probability fields. In *Proceedings 5th International Symposium on Spatial Data Handling*, pages 480 – 490. University of South Carolina, August 1992.

[19] Mark Monmonier. Strategies for the interactive exploration of geographic correlation. In *Proceedings of the 4th International Symposium on Spatial Data Handling, Vol. 1*, pages 512–521. IGU, July 1990.

[20] G. Wills et al. Statistical exploration of spatial data. In *Proceedings of the 4th International Symposium on Spatial Data Handling, Vol. 1*, pages 491–500. IGU, July 1990.

[21] Craig M. Wittenbrink. *Designing Optimal Parallel Volume Rendering Algorithms*. PhD thesis, University of Washington, 1993.

[22] Peter Fisher. First experiments in viewshed uncertainty: The accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing*, 57(10):1321–1327, 1991.

[23] Mark Monmonier. Time and motion as strategic variables in the analysis and communication of correlation. In *Proc. 5th Int. Symp. on Spatial Data Handling*, pages 72–81. Univ. of South Carolina, August 1992.

[24] P. Fisher. *Visualization in Geographical Information Systems*, chapter Animation and Sound for the Visualization of Uncertain Spatial Information, pages 181–185. 1994.

[25] Nahum D. Gershon. Visualization of fuzzy data using generalized animation. In *Proceedings of Visualization 92*, pages 268–273. IEEE Computer Society Press, October 1992.

[26] L. K. Lewis, J. M. Wilczak, A. B. White, and Lingling Zhang. Newly developed x windows tools for analyzing the boundary-layer structure. Technical report, Environmental Technology Laboratory–NOAA, 1994.

[27] Daniel.C. Law. Effects of precipitation, convection, and waves on NOAA network profilers. In *Proceedings of the 25th International Conference on Radar Meteorology*, pages 43–46, Boston, June 1991. AMS.

[28] Richard M. Hodur. Evaluation of a regional model with an update cycle. *Monthly Weather Review*, 115(11):2707–2718, Nov. 1987.

[29] George L. Mellor. User's guide for a three-dimensional, primitive equation, numerical ocean model. Technical report, Princeton University, 1993.

[30] J.M. Wilczak et al. Contamination of wind profiler data by migrating birds: Characteristics of corrupted data and potential solutions. Technical report, Environmental Technology Laboratory–NOAA, 1994.

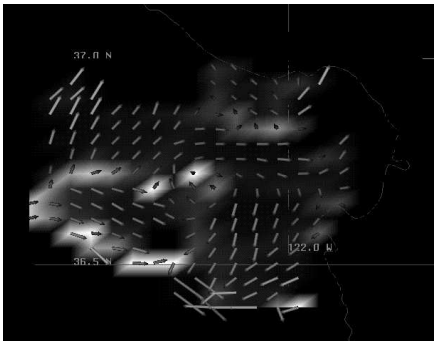

Figure 14: Ocean currents in Monterey Bay are shown with arrow glyphs whose color's are mapped to an HSV map on the magnitude uncertainty, and a pseudocolored image, from black to white illustrates the location of the angular uncertainty. Figure from Spray.[6]

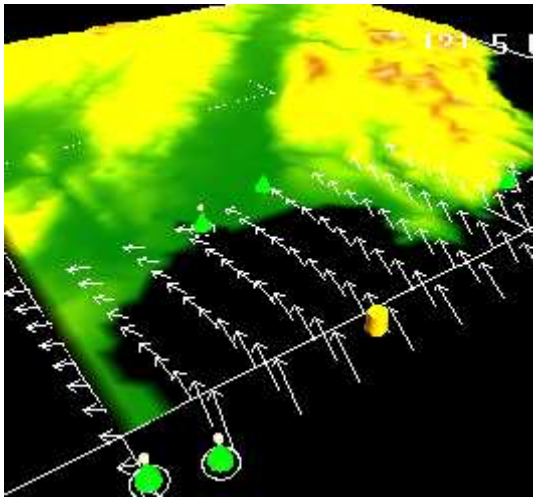

Figure 15: Visualization of wind velocity without uncertainty. Measured surface wind vectors from buoys and meteorological stations are interpolated and resampled over the Monterey Bay region. Regular arrow glyphs are used to represent the wind vectors.

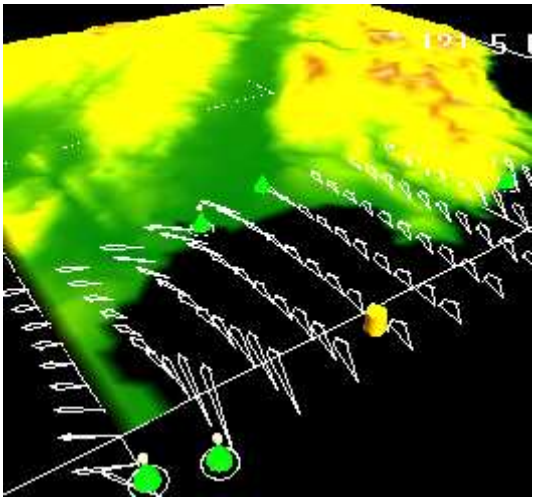

Figure 16: Same vector data with uncertainty that increases with distance from sensor sites giving a different impression. The new glyph incorporates directional uncertainty in the angle between the edges of the glyph. The magnitude is mapped to the area.