Designing a better weather display

Colin Ware [a] and Matthew Plumlee [b]
University of New Hamphshire, Durham, NH, 03824
[a]Center for Coastal and Ocean Mapping; cware@ccom.unh.edu
[b]Computer Science Dept.; mdp@cs.unh.edu

ABSTRACT

Weather maps commonly display several variables at once, usually a subset of the following: atmospheric pressure, surface wind speed and direction, surface temperature, cloud cover, and precipitation. Most often, a single variable is mapped separately and occasionally two are shown together. But sometimes there is an attempt to show three or four variables with a result that is difficult to interpret because of visual interference between the graphical elements. As a design exercise, we set the goal of finding out if it is possible to show three variables (two 2D scalar fields and a 2D vector field) simultaneously so that values can be accurately read using keys for all variables, a reasonable level of detail is shown, and important meteorological features stand out clearly. Our solution involves employing three perceptual "channels": a color channel, a texture channel, and a motion channel in order to perceptually separate the variables and make them independently readable. We describe a set of interactive weather displays which enable users to view two meteorological scalar fields of various kinds and a field showing wind patterns. To evaluate the method we implemented three alternatives representations each simultaneously showing temperature, atmospheric pressure, wind speed and direction. Both animated and static variants of our new design were compared to a conventional solution and a glyph-based solution. The evaluation tested the abilities of participants both to read values using a key, and to see meteorological patterns in the data. Our new scheme was superior, especially in the representation of wind patterns using the motion channel. It also performed well enough in the representation of pressure using the texture channel to suggest it as a viable design alternative.

Keywords: flow visualization, meteorology, human perception, visual channels.

1. INTRODUCTION

Weather maps are possibly the most common data visualizations on earth. A significant proportion of the human population views a weather map every day, either on the pages of a newspaper, on television, or a web-site. Non meteorologists may mostly be concerned with the forecast temperatures and whether it is likely to be overcast or clear or raining. People with meteorological knowledge are interested in the atmospheric pressure, because this helps them reason about the likelihood of storms, winds or persistent sunny weather. Some people have specialized needs for information such as sailors and pilots who are interested in wind speed and direction. Across all the various constituencies probably the most common variables of interest are pressure, temperature, wind speed and direction. Figure 1 shows a meteorological map that attempts to show pressure, temperature and wind speed, but is illegible over large areas. Of course a single example means little, but we have been unable to find any examples of meteorological maps that show these variables clearly on the same display.

In this paper we report on a project that had two goals. The first was abstract: to see if it was possible to simultaneously visualize two 2D scalar fields and a 2D vector field such that values could be read at any data point with reasonable accuracy. The second was to produce a measurably effective multivariate meteorological display.

The problem of visualizing continuous multivariate maps is common across many disciplines including geology, physics and microbiology as well as meteorology and oceanography. One class of solutions that has been proposed is based on glyphs[7,8,9,10,19]. A glyph is a graphical entity that through variations in its size, shape, orientation, texture and color can convey a number of variables. For example one variable can be mapped to color, another to the glyph size and a third to glyph orientation. When glyphs are used to represent a two-dimensional multivariate data field (such as is produced by a single layer of a weather model) the display map is densely packed with glyphs, either in a grid or using some other placement method. Some glyph displays claim to represent as many as 12 variables but, without a formal evaluation, it is not known to what extent the variables can be perceived.

Perceptual theory has a bearing on whether or not displayed variables interfere with one another[15]. Garner's theory of separable features suggests rules for the degree to which variables can be perceived independently and some work has been done to examine issues of the relative independence of pairs of variables[4]. But higher order perceptual interactions have not been studied. Another approach to the display of multiple scalar fields is the color mosaic display. A study by Hagh-Shenas et al.[6] examined the ability of subjects to read values from a multivariate map using a mosaic of small color patches displaying six variables, but the data were not continuous, instead consisting of relatively large patches representing states in the continental USA. In any case, for three variables the average error was around 8.5%. It is not known how well this method would perform with continuous variation. There is also a wealth of perceptual evidence that contrast effects can cause distortions in the perceived values whether displayed using a color sequence or a texture sequence. Sometimes the average error can amount to 20% of the full scale[14]. In the course of this paper, we discuss the extent to which these problems can be mitigated.

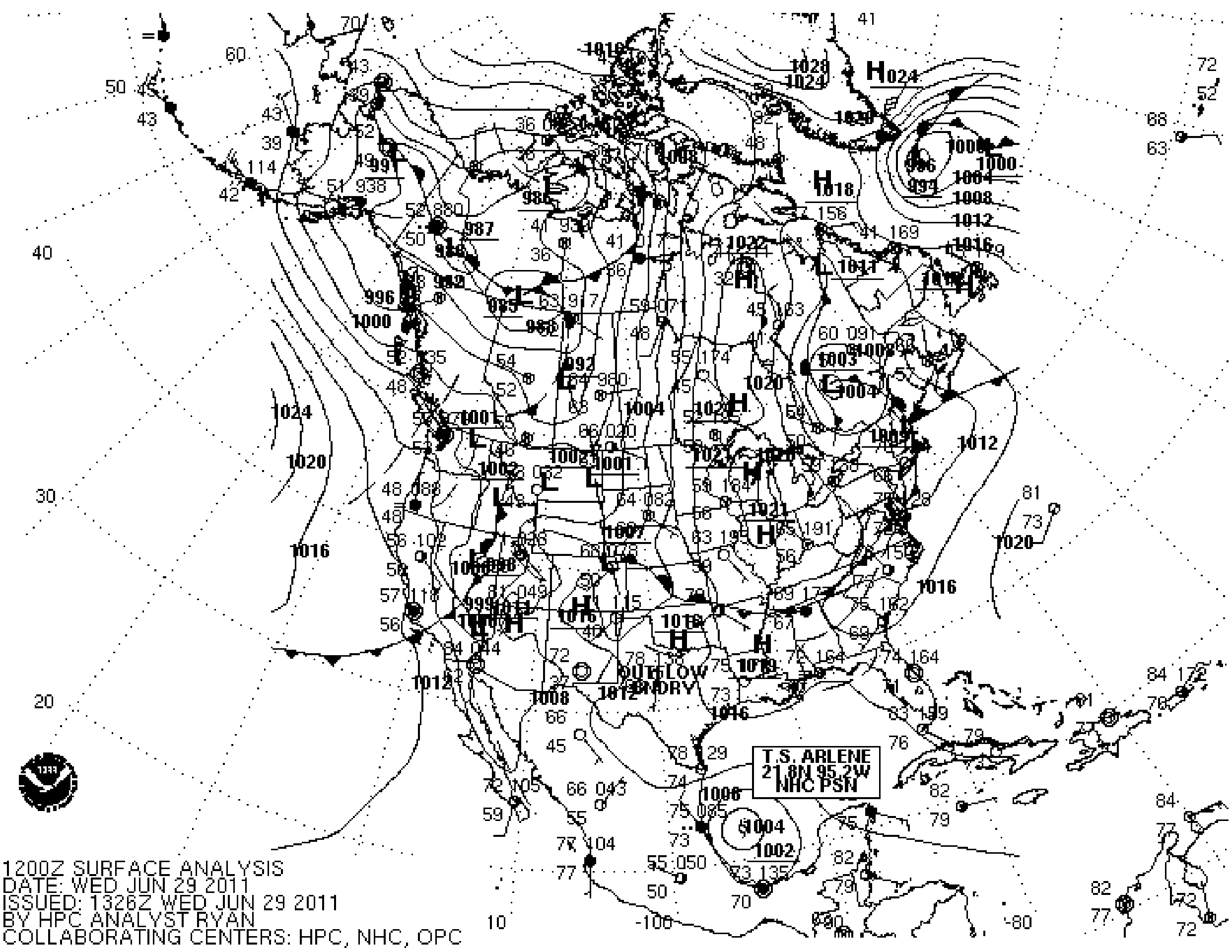


Figure 1. Image from the NOAA National Weather Service site. Contours show pressure contours, temperature and dew point are given with numbers, and wind speed and direction are shown using wind barbs.

We chose to concentrate on the display of meteorological data, partly because there are de facto standards for displaying such data. High resolution meteorological maps typically show atmospheric pressure using contours at 4 millibar intervals, and the range of pressures is commonly 50 millibars over the continental USA and adjoining oceans, meaning that 12 levels of contours must be shown (although in exceptional circumstances such as hurricanes the range can be considerably greater). Temperature is commonly shown using color codes at 5 degree (F) intervals and the range can easily be 60 degrees so 12 or more distinct values must be shown. Wind speed is usually shown by means of a standardized glyph called a wind barb (Figure 2). This allows 5 knot intervals to be shown. Wind barbs are capable of showing winds of hurricane speed in excess of one hundred and fifty knots, but ranges up to 60 knots are more common meaning that at least a 12 step scale is needed. Wind direction is given by the orientation of the barbs.

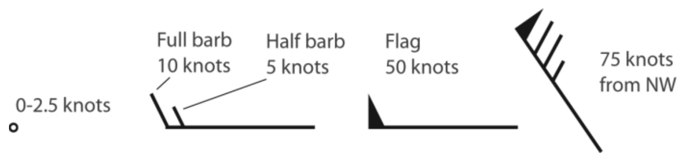


Figure 2. Wind barbs, a quantitative glyph used in meteorological maps to show weather information.

It was our goal to create an example of a multivariate weather map display where each variable was provably legible to at least the resolutions given above. In addition we intended that patterns important to meteorologists be visible in the display. For example, cyclonic and anti cyclonic wind patterns, as well as areas of high and low pressure should be clearly visible.

In the following sections we first briefly describe the data visualized in all designs. We then present the perceptual basis for our design and describe its implementation in a public weather display. Following this we present two design alternatives (classic and glyph-based) used in our empirical evaluation and describe the method and results of that evaluation. We finish with a discussion of the implications of those results.

2. NAM Model

The data set we chose for our public display and for the evaluation study was the surface layer of the National Weather Service's North American Mesoscale (NAM) forecast model covering the continental US. It uses a 614x428 grid in a Lambert conformal projection. Each forecast spans 84 hours in 3-hour increments (29 time-steps). A new, updated forecast is posted every 6 hours. The forecasts used for the images toward the beginning of this paper vary, but the forecast used for the evaluation started at 2:00 pm EDT August 4, 2010, and ended at 2:00 pm on August 8, 2010.

3. Our Three Visual Channel Design

A goal of our project was to see if it was possible to simultaneously visualize two 2D scalar fields and a 2D vector field so that values could be read at any data point with reasonable accuracy. The key theoretical idea behind our new design was to make use of the way the brain processes different kinds of patterns separately at early stages of processing by the visual system. Figure 3 illustrates a cross section of a small area of the primary visual cortex (V1) in the brain showing different sub-regions as classified by Livingston and Hubel[12]. All later stage visual processing depends on V1 where every part of a visual image is processed in parallel for a variety of different elementary features. There are three main classes of such features and these are reflected in the neural architecture: *form* (which processes local orientation and size information, providing the basis for texture discrimination), *color* (local color diffences are encoded), and *movement* (local simple motion patterns are encoded). Sometimes these feature classes are described as visual channels, because they provide relatively independent modes by which information is transmitted to higher levels of visual processing. In addition to the neurophysiological evidence for separate processing channels there is abundant support from psychophysical studies of human perception[11].

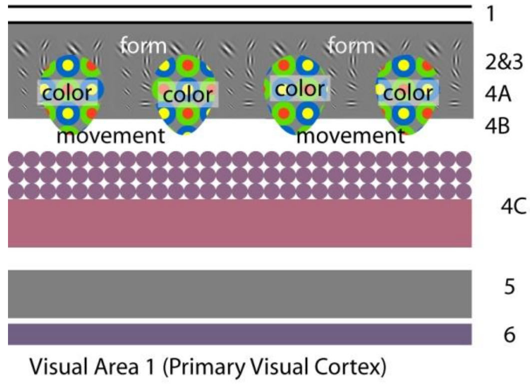


Figure 3. Adapted from Livingston & Hubel[11], a cross section through the primary visual cortex (Area V1) of the brain. This shows different regions specialized to process different kinds of information in layers 2-4B: form (texture), color, and movement. The other layers deal with inputs to and outputs from this cortical region.

For our initial design, we chose to map temperature to color, wind to animated traces (motion), and atmospheric pressure to texture (form), as illustrated in Figure 4. We chose to use color to represent temperature primarily because it is so pervasive in modern weather visualizations. We chose animated traces to represent wind because they provide the expression of directionality needed for a 2D vector field. We added numeric text to provide supplemental wind magnitude information, finding that this did not seem to interfere significantly with the other channels. Finally, we chose texture to represent atmospheric pressure, initially because it was the only channel remaining. We later found that it provided an excellent opportunity to mix the use of texture with the use of countours, which are commonly used for representing pressure in weather maps and make steep changes in pressure visually salient.

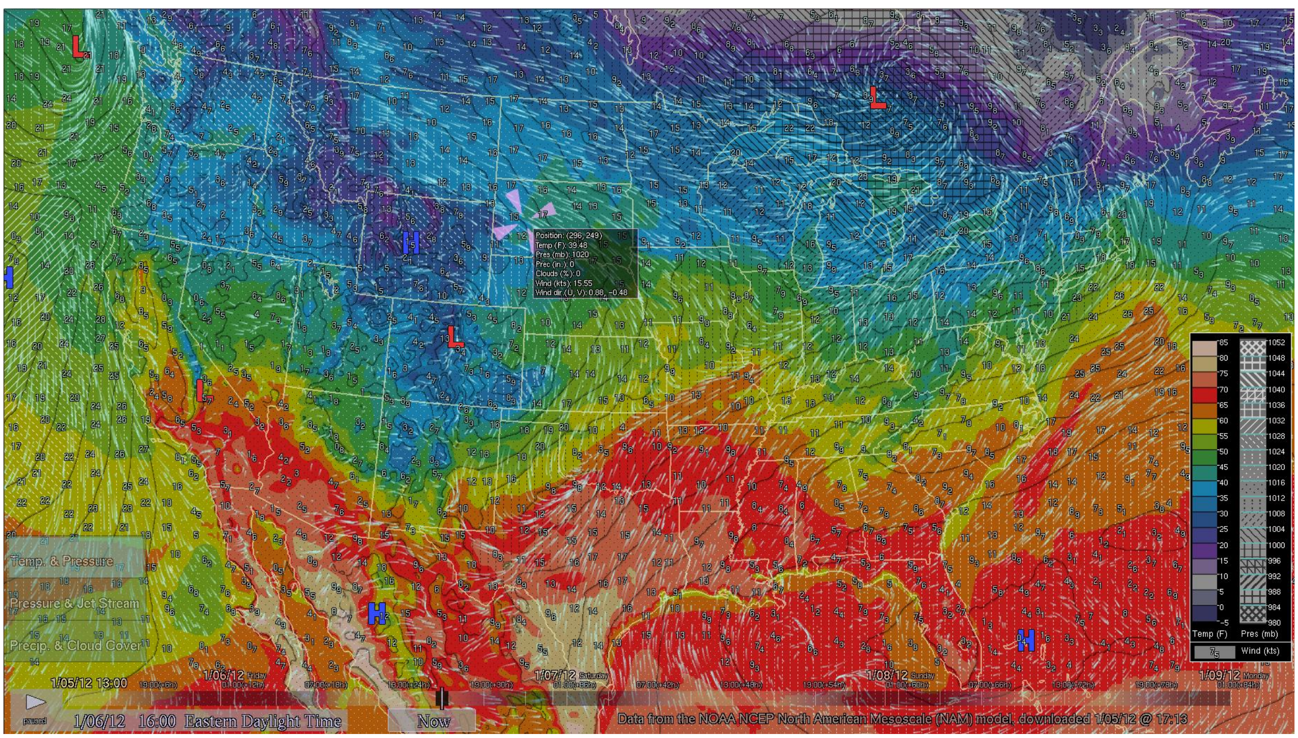


Figure 4. A still snapshot of our primary view design as implemented in a public display. Temperature is represented by color, while atmospheric pressure is given by contours and a sequence of textures. Animated wind traces show the wind direction and speed. Numbers show wind speed at sample points.

3.1 The color sequence design

When only temperature is shown on a meteorological map the most common representation is using a pseudo-color sequence, along with a key that can be read to reveal actual numbers. The color sequence used is often based on an approximation of the physical light spectrum with purples and blues showing low temperatures, reds showing high temperatures and yellows and greens between.

There has been much debate about the proper design of a color sequence for displaying a univariate map. Various researchers have argued that the spectrum approximation sequence is a poor choice on the grounds that to see the overall patterns in the data it is better to use a sequence that monotonically increases in luminance[1,14]. This is because the visual system mostly uses luminance variation to determine shape information, color providing a secondary attribute. Nevertheless, there are reasons for using a sequence that is somewhat like the physical spectrum to display temperature. Some colors have semantically intuitive mappings such as blue for cold and red for hot. But in order to get clearly distinguishable intermediate temperatures, other colors are needed and a transition through green and yellow provides many distinct steps. Also, people are familiar with spectrum sequences on weather maps. Finally, it is capable of accurately conveying more values than any straight line sequence through color space[15].

Our primary requirement was to enable people to read something of the order of 12 color steps accurately, as well as to broadly show areas of high and low temperatures at a glance. The sequence we designed is illustrated in Figure 4. It has a complex path in color space; its middle section is roughly a spectrum approximation with light pink hues added to represent the hottest temperatures, and blues and grays added to represent the coldest. We do not claim that this color sequence is optimal, only that it is a reasonable solution to the design problem.

3.2 Animated wind traces

We chose motion to represent wind direction and the pattern of wind speeds because there is an intuitive mapping between moving patterns and wind patterns. To generate motion, we use a modified particle tracing method, displaying particles with fading tails representing their recent history[3,18]. Our use of animation should not be seen as a way of reproducing actual particle advection trajectories, but rather as a method for making wind patterns clear, and also to judge wind speed and direction. Initially we used a linear mapping between wind speed and animation speed but found the results were either too fast in regions of strong winds or too slow in regions of light winds. To make the patterns clear in both high wind and low wind regions, we transformed wind speed (in m/s) according to the function $displayed\_speed = c \cdot model\_speed^{0.65}$. The exponent was chosen in a pilot experiment where participants were given control over a slider to adjust the exponent parameter while the animation was running and were asked to set a value that resulted in a good representation of strong, weak, and medium wind strengths. The value of c was set to be .01. This results in a mild 5 knot wind being displayed using a trace speed 1.9 cm per second, while a strong 30 knot wind is displayed at about 5.9 cm per second (given a 22 inch monitor).

To create the animated traces, particles are randomly seeded over the entire field such that approximately 9600 appear on the screen at a time. Each trace is formed by Lagrangian advection of an idealized particle according to the speed-transformed underlying vector field. Each particle trace is given a tail that increases in transparency. To eliminate distracting visual transients at the initiation and termination of traces, particles are randomly initiated in time. Also each individual trace begins as a point, grows to full length, and terminates by fading out, as illustrated in Figure 5. Traces “die out” when they had been animated for a pre-determined (constant) lifetime or when they reach the edge of the screen. A trace dies by stopping advection and allowing its tail to fade completely, at which point it is “reborn” in a new random location.

To allow for more precise judgments of wind speed a jittered grid of numbers shows wind speeds to two significant figures. These numbers are randomly jittered along a grid spaced at 10 data cells. The color of both the wind traces and the numbers is a pale cyan, chosen to be easily distinguishable from any temperature coloring.

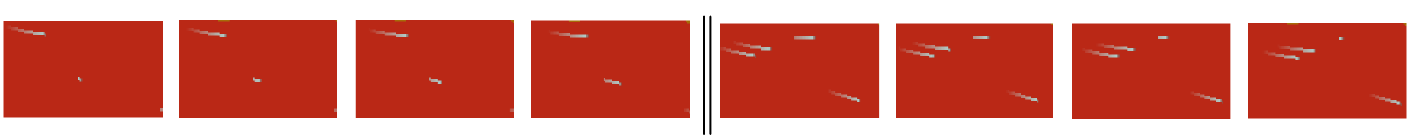


Figure 5. An illustration of the life cycle of wind traces, showing the first and last four frames of an 18-frame sequence. The first four frames show a trace coming into existence (center of each frame) and propagating forward. The last four frames show another trace (top-middle of each frame) shrinking out of existence.

3.3 Quantitative textures for pressure

We use texture to represent atmospheric pressure, designing a quantitatively coded sequence of textures[16]. The concept is to create a series of textures that vary monotonically in a perceptual sense from a low value to a high value and where each texture in the sequence is legibly distinct from the previous one. This allows for both pattern perception – the highs and lows are immediately apparent, and the ability to identify each texture in the sequence on a key and thereby “read off” the atmospheric pressure values using a key. The textures also have the quality of *laciness*[17] meaning that other information, such as the color coded temperature, can be perceived through the interstices. This is accomplished by creating (1) a sequence of textures made of black elements were each texture has less density than the previous one, then adding (2) a sequence of textures using white element where each elements has greater density than the previous one. The transition from black to white elements is made approximately at the average atmospheric pressure for the data shown. Texture transitions are made every four millibars. Both black and white are easily distinguishable from both the colors reserved for temperature and wind artifacts. The rightmost column in the key in Figure 4 shows the entire range of textures used.

In our initial design, we did not have contours. We added them because a meteorologist pointed out that pressure gradients are of critical importance in interpreting weather model output and the closeness of contours provides an excellent visual cue to the gradient. The contours also serve to better separate the different texture bands (see Figure 4).

4. IMPLEMENTATION IN A PUBLIC WEATHER DISPLAY

We developed an interactive public display of the forecast data with a touch screen interface. It automatically updates with the latest the 84-hour NAM forecast every six hours, uploading the latest model output from the National Weather Service NCEP FTP server. It supports user interactions for controlling time, for controlling the mapping of data to visualizations, and for querying the forecast for underlying values at forecast grid cells. This display has been running operationally in the center for Coastal and Ocean Mapping for more than a year and is being installed at the National Oceans and Atmosphere Administration in Washington, D.C.

4.1 Hardware

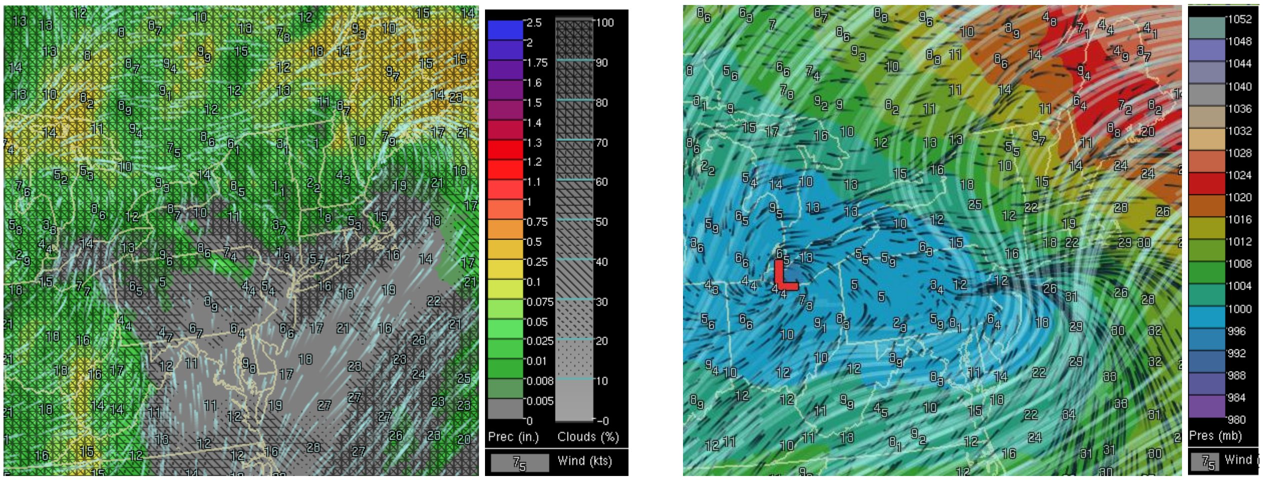
The public display uses a 40” touch screen, rendering at 1920x1080 pixels. The touch screen registers a single touch-point and supports only click and double-click operations (no dragging). The computer driving the display is an Intel Xeon-based personal computer running a 64-bit Windows 7 operating system, with an nVidia GeForce 400 series graphics card.

4.2 Alternative Views

We initially implemented a set of four alternative views in addition to the default. Then after it had been displayed for several months, we reduced the total number to the three. The default was the one we have already described showing surface temperature, atmospheric pressure and near surface wind patterns. The remaining two alternatives are illustrated in Figure 6. Each has a corresponding customized key to allow users to understand the values being shown. For example, millibars in the case of pressure, and inches of water per three-hour period in the case of precipitation.

4.2.1 Wind, Precipitation and Cloud Cover

People are often interested in whether it will be cloudy and when it will rain. Figure 6a shows a view that shows patterns of cloud cover using texture and precipitation using color. The cloud-cover mapping uses a different texture sequence designed to have less distinguishable boundaries from the other texture sequences.



(a) (b)

Figure 6. The two most successful alternative views, alongside their custom keys. (a) Precipitation is color coded and cloud cover is textured. (b) The jet stream level winds are shown in white, surface winds are shown in black and ground level pressure is color coded.

4.2.2 Pressure, Surface Winds and Jet Stream Winds

People with more meteorological sophistication are often interested in how jet stream level winds steer the weather patterns that sweep across the continental United States. The view shown in Figure 6b displays winds at the 250 mbar pressure level and uses color coding to show atmospheric pressure. To simultaneously show surface level winds with jet stream level winds we use *counterphase* luminance levels for the streamlets. Perceptual research shows that patterns that are black on a mid-level background can easily be discriminated from patterns that are white[13]. It also helps that the winds in the upper atmosphere have much smoother trajectories than surface level winds. The jet stream component of the display is purely thematic; it does not allow wind speeds to be read off. In order to allow the jet stream pattern to stand out clearly, we chose linear scaling instead of the non linear scaling applied to surface level winds: $displayed\_speed = 0.0022 \cdot model\_speed$. For example, an 80 knot wind is displayed using a trace with a speed of just over 9 cm per second.

In order to highlight only the wind speeds considered to be part of the jet stream, a filter is applied such that a trace only appears at full opacity when its head represents a wind speed over 40 m/s (~78 knots). Lesser wind speeds cause the trace to have an opacity equal to the proportion of its speed over 40 m/s. Even though all wind speeds above 78 knots are drawn at the same opacity, regions of significantly strong winds appear denser due to the greater probability of there being a trace that will transit through the region.

4.3 Time progression

The applications has two modes, one in which time progresses (at the rate of a little less than one model step or three hours per second) and a mode when time does not progress. The former allows for perception of time varying patterns, such as fronts sweeping across the country, while the latter allows for more detailed examination of a particular snapshot in time. When time is progressing and the model reaches the last forecast time, the application loops back to the initial model time and the visualization repeats.

4.4 Interactive Controls

A number of interactive controls are provided to give the user the ability to change one of the alternative views, change the time of the forecast, or access more detailed information at a particular time and place. The semi-transparent menu in the lower-left corner of the display (as shown in Figure 4) makes the alternative data visualizations easily accessible. The menu becomes fully opaque when the user hovers the mouse over it.

4.4.1 Time Controls

The time bar, illustrated in Figure 7, allows the user to start and stop the animation, and provides an indication of the time of the forecast is being rendered, and a way for the user to stop the animation at a particular time. The time bar has a tick-mark at every three-hour step of the forecast, with times and dates at selected positions along the bar. The color of the time bar varies between dark and light to indicate night and day hours, respectively. Clicking on the time bar "snaps" a time indicator to the time of the nearest tick-mark, halts the progression of time, and causes the corresponding forecast step to render. The animation of the streamlets does not cease, however, when time progression is halted. Instead, it continues in order allow the perceptual separation of wind patterns from the other layers of information. Dragging the time indicator along the bar causes forecasts to be linearly interpolated for the indicated times. Whenever a new time is selected, a textual display below the time bar is updated.

Figure 7. The left end of the time bar, showing date boundaries and selected times from the forecast model, as well as a dark indicator that can be dragged to a new forecast time. Also visible are a buttons to play/pause forecast time and to set the forecast time to the actual clock time, as well as a readout of the forecast time currently being displayed.

A button to the left of the time bar toggles the advancement of time between "paused" and "playing" states. The application enters the "paused" state either when the toggle button is pressed, or the user selects a time on the time bar. In the "paused" state, wind traces continue to advect to indicate wind pattern.

Below the time bar, there are two information displays and a button. The first information display is the forecast time currently being rendered, shown in the local time zone. The second information display gives the source of the forecast data, and the approximate time that the data was downloaded from the source. Finally, there is a "Now" button that changes the time step of the model to match the current time reported by the computer's clock. Pressing the "Now" button causes the application to enter the "paused" state.

4.4.2 Data point summary

Wherever the user touches the map on the screen, a point summary appears with scalar and vector data at the nearest forecast node, linearly interpolated exclusively with respect to time. The point summary is offset from a cursor that slowly spins about the central point being summarized as shown in Figure 8. The summary information continuously updates with respect to the time being rendered. To allow the display to return to an un-annotated state between users the summary and cursor both start to fade after a minute, disappearing completely after several more seconds unless the user touches a new point.

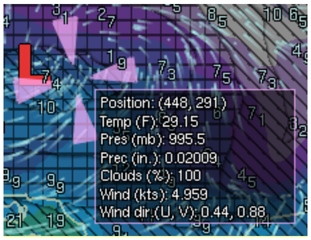


Figure 8. The data point summary consists of a rotating pink cross that highlights the selected model forecast point, and a small window that lists the loaded model data for that point.

4.5 Other Display Features

4.5.1 High and low pressure labels

High and Low pressure labels are automatically generated for each time step using the following procedure applied once for highs, and once for lows.

1) Local maxima are first identified by comparing each cell with its surrounding two rows and columns (a 5x5 grid). The maxima are sorted and the positions of the largest 25 pressure values are preserved.

2) Filter out any values in the ranking whose distance is "too close", meaning that if two values are within 45 forecast

grid cells of each other, the less extreme value is dropped from consideration.

3) In some cases, two values are linked by a pressure ridge. The following method is used to ensure that a label appears only at the most extreme of the two in such a case. Lines are traced between each pair of the remaining values. The path goes from the less-extreme value toward the more-extreme one, and both values are kept only if there is a segment along the way that demonstrates a “dip” toward a less-extreme value by at least 2.2 mb.

4) Remaining values are labeled with a blue “H” for highs.

5) Repeat the process for minimum values and label with a red “L” for lows.

4.5.2 Mapping-aware key

A panel on the lower-right side of the display provides a key for interpreting information shown on the map. This is customized to the particular view being shown.

5. IMPLEMENTATION DETAILS

The continuous color background with texture overlay, as well as the contours are generated using a fragment shader on the GPU. Temperature and pressure data are first normalized to [0, 1] according to the bounds of their respective scales. They are then loaded onto the GPU as 2D scalar textures (using GL_DEPTH_COMPONENT rather than a method that maintains color information). Each time a new animation frame is rendered to the screen, a single quadrilateral is drawn to cover the entire screen using texture coordinates to map the data in these textures to their proper screen positions. While the method below applies for any mapping of 2D scalar fields to color, texture, and contours, the following description assumes for clarity that temperature is mapped to color and pressure is mapped to texture.

5.1 Textures

To color a given pixel, the fragment shader first uses the texture coordinate of a given pixel to perform bilinear interpolation on the temperature and pressure data. A base color is then selected from the temperature color scale (1D texture) using the interpolated temperature value as an index. If a texture overlay is in use, a texture color is selected from the pressure texture scale (3D texture) using for indexes the pixel’s screen position and the interpolated pressure value. The texture and base colors are blended using an alpha value of 0.7 (using 70% of the texture color and 30% of the base color where the texture is opaque, but using 100% of the base color where the texture is transparent).

5.2 Contours

Contours are generated in the final step of the fragment shader after color and texture are blended. The resulting color from the previous step is blended with black if it was sufficiently near a boundary on the pressure scale. To do this, we adapted a technique game developers use[5], approximating screen-space derivatives to maintain contours of relatively constant size. Screen-space derivatives are approximated by looking up pressure values in the 2D scalar texture for a position one pixel to the right and another one pixel down. If a given pixel was not sufficiently near a contour, the underlying color goes unchanged. However, the pixel is darkened toward black if it is within a certain number of pixels (*n*) away from where the derivatives predict a boundary would be. The graphics card’s smoothstep function is used to modify the alpha value of the darkening to provide control over how “hard” or “soft” the contour will appear. For our public display implementation, *n* is 2 and the smoothstep function goes from half-transparent black to fully transparent over the course of roughly 2 pixels.

5.3 Additional layers

After the base color layer is generated by the fragment shader, additional layers are rendered on top of it in the following order: vector coastline and state boundaries, wind traces, wind speed text, and wind barbs.The vector coastline and state boundaries are rendered with light grey lines with a width of one pixel. When present, wind traces are rendered as line strips of up to seven segments, with a width of two pixels. A texture is used to change the color along their length, going from a transparent pale cyan at the tail vertex to a slightly more saturated cyan with alpha 0.8 at the head vertex. When present, wind speed text is rendered using the GLUT bitmap font Helvetica 10 in pale cyan. Numbers are rendered with a shadow that is created by rendering the numbers again in black, shifted by one pixel up, down, left, and right of the original. Finally, when present, wind barbs are rendered as lines the same color as the head of the wind traces, with a width of four pixels. Barbs are rendered with an outline that is created by re-rendering the barb with a in a semitransparent black width of six pixels. In the case of both the numbers and barbs, the depth test is used to keep a shadow from modifying any pixels already covered by a number or barb.

5.4 Software environment

A separate scheduled process is run every six hours to download the latest version of the NAM forecast. On completion, this process replaces the data file that our display uses as its primary data source. The display application checks for an updated file periodically, and reads the new data file when detected.

6. EVALUATION: COMPETING DESIGNS

In order to evaluate the effectiveness of our weather data visualization, we conducted a study in which subjects were asked to perform a set of tasks using four different designs.

The default view from our public display (designated the *Animated* design) was compared to three alternatives. There were small difference from the version used in the public display: 1) vector coastlines and state boundaries were drawn first, at a line width of two pixels, giving them priority over coloring from the fragment shader; 2) texture transitions were made every two millibars; and 3) the grid of wind-speed numbers is populated by first filling in wind speeds that are local maxima within a 24 data-cell area, and then filling the unpopulated grid positions with a randomly located wind speed value from within the central 4x4 region of data cells.

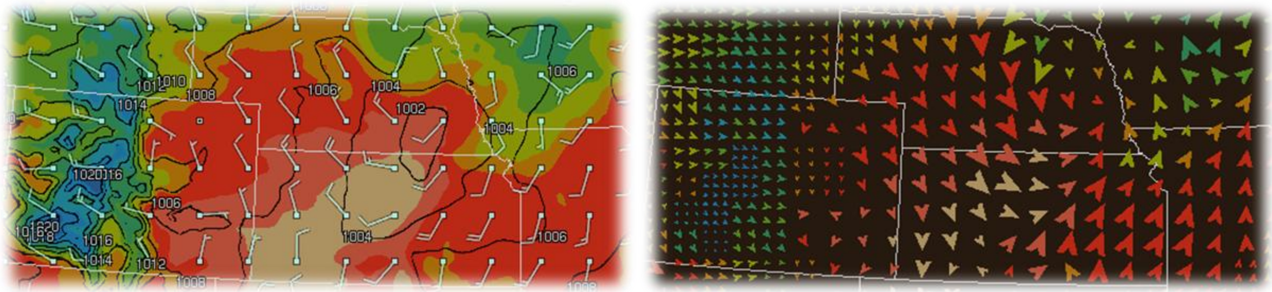
We now describe three additional designs that were used in the evaluation: *Classic*, *Glyphs*, and *Static*.

6.1 The Classic Design

The *Classic* design is our attempt to re-design techniques commonly used in meteorological maps to enhance clarity. Although displays such as that shown in Figure 1 are difficult if not impossible to interpret, we thought it possible that with careful design we could come up with a solution using common existing methods to meet our requirements. Standard practice in meteorological displays is to use contours for atmospheric pressure. Figure 1 shows an example of a meteorological map showing contours for pressure, with wind barbs for wind speed and direction. As mentioned before, temperature is often represented using a pseudo-color sequence, together with a key that can be read to reveal actual numbers.

Figure 9a illustrates a small sample from our *Classic* design showing the weather over northeastern Colorado. Pressure is represented using two millibar (mb) contours labeled at even-numbered intervals. The contours are generated using the fragment shader program described in the implementation details for the public display, but where n was .8 and the smoothstep function went from fully opaque black to transparent in less than half a pixel width. This had the effect of making the contours look as "hard" as contours shown in Figure 1, in contrast to the "soft" contours employed in the *Animated* design. Contour labeling is accomplished by first assembling the same collection of points representing the highest and lowest pressures that we used in the public display to generate high and low pressure labels, plus a point for each corner of the screen. Second, imaginary lines are run between every point in the collection, and the intersections of these lines with contour boundaries are recorded. Finally a coarse grid is used to filter out extraneous intersections: only one intersection at each contour level is allowed to remain, and the one closest to the center of the grid cell is the one that would represent that contour for the cell. The surviving intersections resulted in labels. The result does not label every contour, nor does it avoid the occasional occurrence of a label overlapping multiple contours in high-gradient areas.

The *Classic* design uses traditional wind barbs to represent wind direction and speed. The wind barbs are evenly spaced 10 data cells apart. Each half-length barb represents five miles per hour (mph) of wind speed, while each full-length barb represents 10 mph. Wind barbs are generated to represent ranges centered on integral multiples of 5 mph, with a range of plus or minus 2.5 mph. Wind speeds of 0-2.5 mph are represented by square dots. The spacing of both the wind barbs and the pressure levels was chosen to best approximate what the spacing and resolution might be if the weather maps such as those shown in Figure 1 were designed take up a full screen. The spacing of wind speed numbers in the *Animated* design was chosen to match the spacing of the wind barbs here.



(a) (b)

Figure 9. (a) The *Classic* design. Labeled contours are used for atmospheric pressure. Wind barbs show wind speed and direction. (b) The *Glyphs* design. Pressure is shown by the number of arrows in a cell, wind direction by the way the arrows are pointing, and wind speed by the relative size of arrows. Both images show the same data.

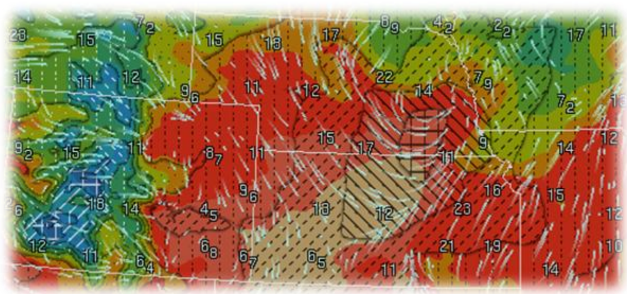


Figure 10. The *Animated* and *Static* designs. Pressure is given by contours and a sequence of textures. In the *Animated* design, animated traces show the wind direction and speed. In the *Static* design, static traces show the wind direction without animation, as shown. Numbers show wind speed at sample points. This is the same data as shown in Figure 9.

6.2 Glyph-based Design

We chose to base our glyph design on the work of Healy et al.[8], mainly because their work targeted meteorological data, and their design could be used with relatively minor modifications. The *Glyphs* design is illustrated in Figure 9b. It employs regularly spaced arrowheads to represent the underlying data, in a manner suggested by Healey et al.[8] The glyphs are evenly spaced six data cells apart, each one centered over the location of the single data point it represents—no averaging of nearby data cells is done. Temperature is mapped to glyph color using the sequence already described, and wind direction is mapped to the orientation of the delta-shaped glyphs (Healy et al. used orientation for scalar variables). In order to represent pressure, the number of arrowheads that make up a glyph is varied (they refer to this as “density”). For areas of pressure below 1008 mb, a single arrowhead is drawn. Between 1008 and 1016 mb, four arrowheads are drawn at half-size in each linear dimension, filling a 2x2 grid within the glyph. Finally, above 1016 mb, nine arrowheads are drawn at one-third-size in each linear dimension, filling a 3x3 grid. The center of each arrowhead remains at a fixed location, so any variation in orientation or size of the arrowheads is with respect to that fixed center.

Finally, wind speed is represented in the *Glyphs* design as the relative size of the arrowheads—effectively the amount of screen real-estate that was covered by a glyph. The size of arrowheads increases linearly every 8 mph in the range of 0-40 mph, for a total of five different sizes. The glyph spacing was chosen to be as small as possible while still making it possible to determine wind direction at high pressures and moderate wind speeds (sacrificing low wind speeds in high pressure areas in some cases, as seen in the bottom left region of Figure 9b).

6.3 Static version of the *animated* design

To evaluate the effectiveness of animation, we added the *Static* design, illustrated in Figure 10: a variation on the *Animated* design in which the wind traces do not move. This was the only difference between the *Animated* and *Static* designs.

7. EVALUATION: METHOD

To evaluate the four different display alternatives, we carried out an experiment with two parts. The first part involved subjects repeatedly estimating the temperature, pressure, wind speed and wind direction for different points on a weather map. In the second part, subjects judged how clearly they could see different features of meteorological interest.

7.1 Experiment Task

On each trial, a cursor appeared and the subject's task was to estimate the temperature, pressure, wind speed and direction at that point. Figure 11 shows an example of what a subject would have seen for the *Classic* design. The subject was responsible for making selections in the key that reflected their best judgment of the underlying data at the center of a slowly spinning cursor. For example, if the cursor was centered over a boundary between light and dark green (as in Figure 11), the subject should have selected a temperature of about 70 degrees from the key. Likewise, the subject should have traced the contour line to find the 1018 and selected 1018 mb for the pressure, and used the nearby wind barbs to estimate wind direction and speed and made appropriate selections in the key. Once all selections were done, the subject was responsible for clicking on the "Done" button in the lower-right corner.

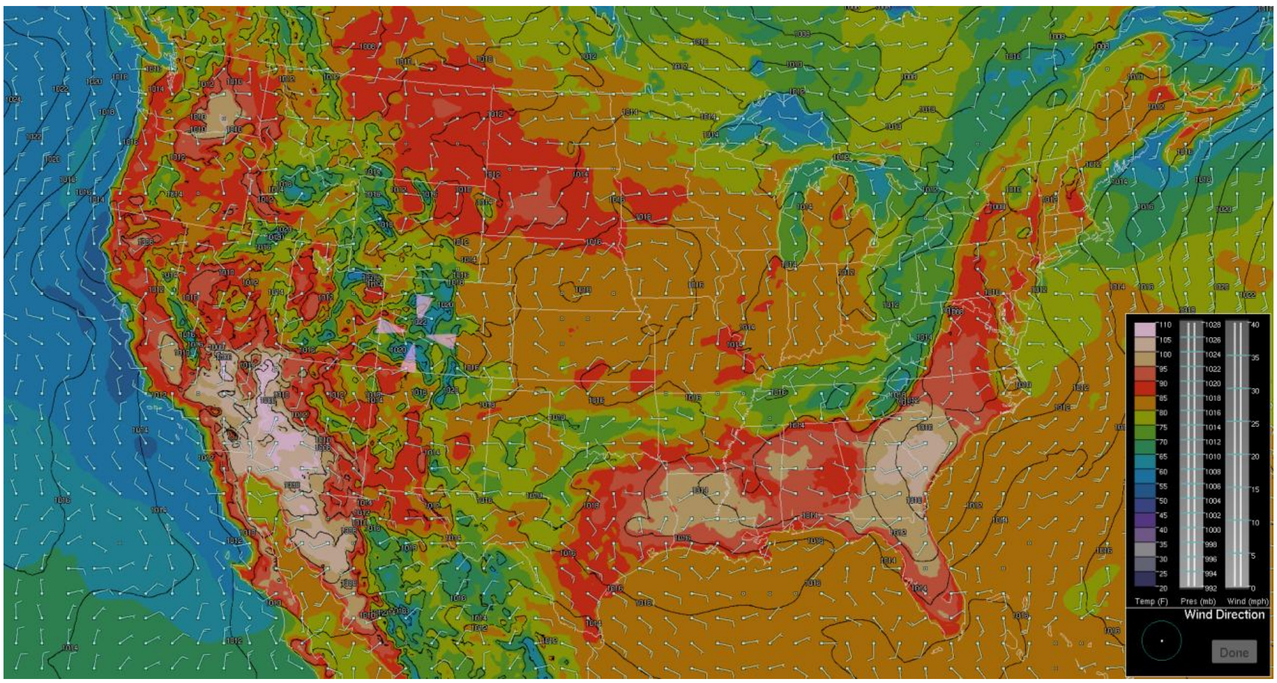
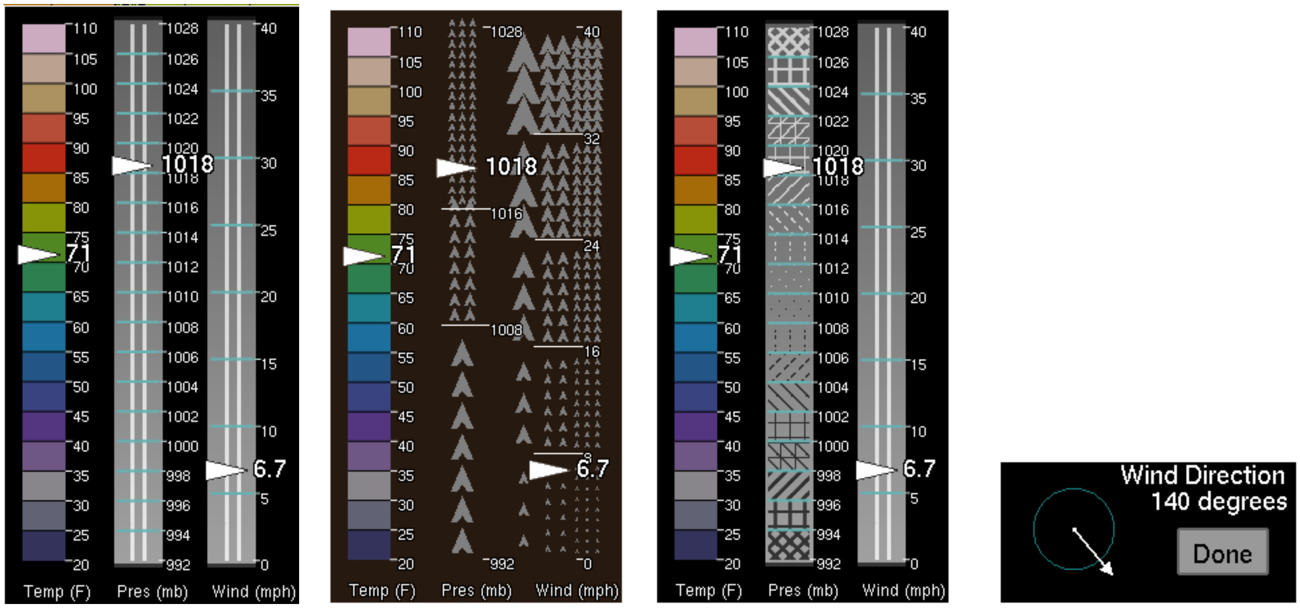


Figure 11. An example of what was seen by subjects after the one-second orientation period for the *Classic* design.



(a) (b) (c) (d)

Figure 12. Keys used for the task. The upper portions of the keys presented for the (a) *Classic*, (b) *Glyphs*, (c) *Animated* and *Static* designs. Each illustrates selections already made by a subject at 71°F, 1018 mb, and 6.7 mph for temperature, pressure, and wind, respectively. (d) The lower portion of the key presented for all designs. It illustrates a selection made at 140 degrees clockwise from the North position (with North pointing straight up).

The subject selected temperature, pressure, and wind speed values by using the mouse to move numeric indicators along three vertical scales in the key as in any of the keys shown in Figure 12a-c. An indicator appeared whenever the subject pressed a mouse button with the mouse cursor along a scale, and the subject could drag the indicator to the desired value. The indicator consisted of a triangle on the left and a numeric indication of the value on the right. The subject was free to adjust a selection by clicking on the scale again, at which point the indicator would “snap” to the position of the mouse cursor.

The subject selected wind direction values by using the mouse to move an arrow indicator around a circular scale, as shown in Figure 12d. The arrow indicator appeared whenever the subject pressed a mouse button near the circular scale, and it followed the mouse cursor around the scale such that an arrow always pointed from the center of the scale to the mouse. As with the vertical scales, the subject was free to adjust this selected value up until the “Done” button was pressed.

7.2 Displays

Subjects were presented with the four designs during the course of the experiment. A fixed 490x271 sub-grid of the full NAM dataset was used for display (showing the continental US). The display of this data was scaled to fill the entirety of 23” widescreen LCD monitor at a resolution of 1920x1080.

As already discussed, bilinear interpolation of the underlying temperature data was used to color pixels between actual data points, which tended to be roughly 4 pixels apart. For all designs, temperature was represented using an 18-color scale, as seen in Figure 12.

While the source data for the maps ran the entire gamut of the scale, subjects were most often asked to make judgments about data that was on or near land. Therefore, judgments were only made about temperatures in the 50-100° range, with rare cases in the 40°-50° and 100°-110° ranges, making the effective size of the scale roughly ten colors. Furthermore, the random locations selected for trials never included wind speeds at or above 24 mph with this design, meaning subjects never had to make judgments using the two highest-end glyphs. Finally, the pressures that subjects were asked to make judgments on were mostly in the 1004-1022 mb range, with rare cases in the 1000-1004 and 1022-1026 ranges.

7.3 Experiment Design

Each experimental subject was first trained using four trials with each of the designs. Following this, the experiment consisted of four experimental blocks each containing 16 trials. Each block was partitioned into four sub-blocks of four trials where each sub-block presented a different design. Sub-blocks were arranged in a Latin squares design such that each design was first within exactly one block of trials. The Latin squares were generated randomly such that different subjects saw different orders of presentation.

For each block of trials, data for the maps presented to the subjects were selected from a different time-step of the same NAM forecast model output. At the beginning of each block, a random time-step t (in the range 0 to 28) was chosen and the first map was generated from data at time t. Each subsequent trial presented the subject with a map generated from data taken at a time seven steps later, wrapping back around to the beginning as necessary (mod 29). Each time a block ended, the program paused to load a new time-step and the subject was given a short break until they were ready to proceed.

Each trial began by showing an outline of the United States over a black background with a key and cursor. On the first trial of each block, the following message appeared in the bottom left: "Press SPACE when you are ready to begin the next block of trials." Once the subject pressed the space bar, the first trial of a block was like any other. The black-background map was shown for one second to make it easy for the subject to spot the magenta cursor that slowly spun in a new random location. Then this was replaced by the graphic display corresponding to the condition. The key was always present, but was updated to reflect the design for the current trial.

At the end of the one-second orientation period, the subject was presented with a fully populated map and was responsible for completing the experimental task. Task completion time was measured starting at this point, and ended when the subject clicked the "Done" button. There was no limit to the time a subject could take completing the task, but subjects were asked to try to complete the task within about 30 seconds. As a visual cue after 30 seconds, the spinning cursor would slowly pulse, gradually cycling color from red to white. The amount of time the subject took for the task was recorded along with the selections the user made in the key and the actual data values at the point indicated by the cursor.

At the end of a trial in a training block, the subject was provided with feedback on how well their answers matched with the actual values at the center of the spinning cursor. In order to minimize any training effects as well as to keep the total amount of time a subject spent on the experiment reasonable, this feedback was not given at the end of trials in the experimental blocks.

7.4 Questionnaire

After all the experimental trials were completed, each subject was asked to fill out a questionnaire to rate how easily certain features could be seen. Features for the questionnaire were made in consultation with a weather expert from NOAA, but were worded so that subjects could understand. Subjects were asked to circle a dot along a 9-point scale that progressed at half-point intervals from 0 ("Can't see it") to 4 ("It's very clear") for all four designs on each question. These questions asked how well the subject could see the following items:

Q1. A large low pressure system

Q2. A long line of converging winds

Q3. The land-water temperature differences along a few coasts

Q4. A moderate-sized area of (locally) highest temperature

Q5. A small area of lowest temperature

Q6. A moderate-sized area of (locally) highest wind speeds

Q7. A large area of lowest wind speed

Q8. A large-size area of (locally) high pressure

Q9. A moderate-to-large-size area of low pressure

For each question, the subject was shown where the feature in question was (pointed to by the experiment proctor), and was given the ability to cycle through all four designs as many times as desired. Each design was labeled in the bottom-left corner of the screen, and the same labels were used on the questionnaire. For a given question, every subject saw maps of the exact same feature on the exact same time step of the same forecast as the other subjects did, but randomized elements with the *Static* and *Animated* designs made the maps appear slightly differently (locations of traces and jittered wind numbers).

Subjects were proctored through the experiment by a paid undergraduate research assistant.

8. EVALUATION: RESULTS

8.1 Subjects

There were 13 subjects, eleven of whom were college students of varying backgrounds who were paid for participating. The other two were staff members of the Center for Coastal and Ocean Mapping who had little or no prior training in reading weather maps.

8.2 Quantitative Results

The error data is summarized in Figure 13. This shows the mean absolute errors for each of the four estimates. For temperature, the errors were 1.9 degrees or less for each of the conditions except for *Glyphs* where the errors were 2.75 degree. An ANOVA with a Tukey HSD test showed the data to fall into two groups. The *Classic* and the two new designs (*Animated* and *Static*) were all statistically indistinguishable, forming one group. The *Glyphs* design was the sole member of a second group and was significantly worse than all members of the first group (all significant levels < 0.001). The greater error that occurred with the *Glyphs* design is likely due to the coarser spatial discretization that resulted from the use of colored glyphs.

Errors in pressure responses showed the same pattern of results. All the methods yielded statistically equivalent results (< 1 mb mean error), except for *Glyphs* which resulted in a much greater error of around 4.7 mb. This much greater error can be attributed to Healey et al.'s design having only three levels of texture density.

The results for wind speed showed the lowest errors for the new methods (*Animated* and *Static*), while *Glyphs* were the least accurate. The mean errors for the new *Animated* and *Static* designs were 1.6 and 1.5 knots respectively; they were 2.35 knots for wind barbs (*Classic*) and 5.2 knots for the *Glyphs* design. These fell statistically into three groups. The two new methods (*Animated* and *Static*) produced errors that were significantly lower than the *Classic* design using wind barbs, and wind barbs were more accurate than *Glyphs*.

The wind direction errors were the lowest for the new *Animated* wind traces (19.3 deg) with *Classic* wind barbs next (24.9 deg). Both the *Glyphs* and the new *Static* traces produced errors greater than 30.0 deg. The results did not fall into clean statistical groups according to the Tukey HSD test. Differences were significant between the *Animated* group and the (*Glyphs*, *Static*) group, but they were not significantly different between the new *Animated* traces and the *Classic* wind barbs. Neither was the *Classic* design significantly worse than either member of the (*Glyphs, Static*) group. We did note, however, that 11 out of 13 subjects were more accurate with the *Animated* traces than with the *Classic* wind barbs, and this is significant by a 2-tailed binomial test ($p < 0.05$).

The average time to perform each trial was 22.2 sec. There were no significant differences depending on the display design.

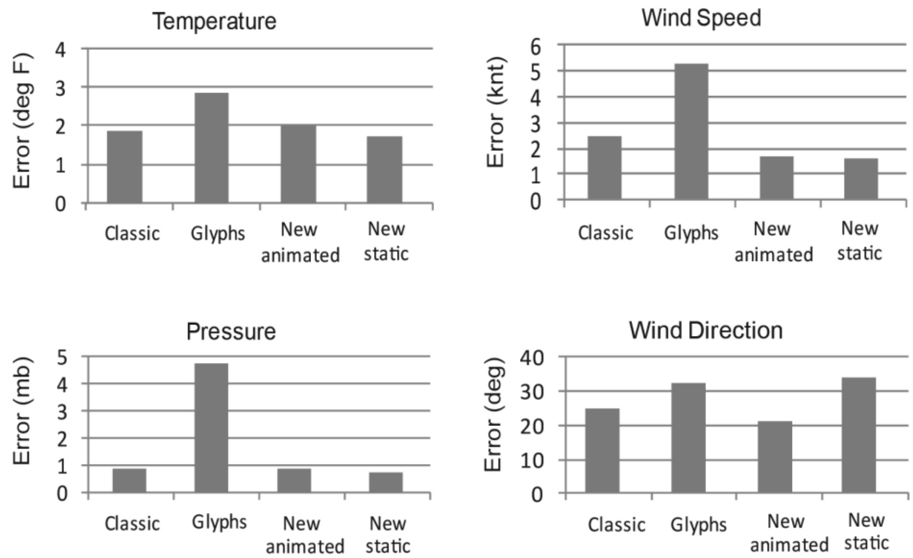


Figure 13. The mean errors for each of the measured variables for each of the four designs.

8.3 Qualitative Results

We now turn our attention to the subjective rating data. ANOVAs and Tukey HSD tests were run independently for each of the nine questions. Three of these yielded no significant differences. Ratings for the remaining six questions are given in Figure 14. In the following discussion we use the symbol “>” to mean “rated more effective than”.

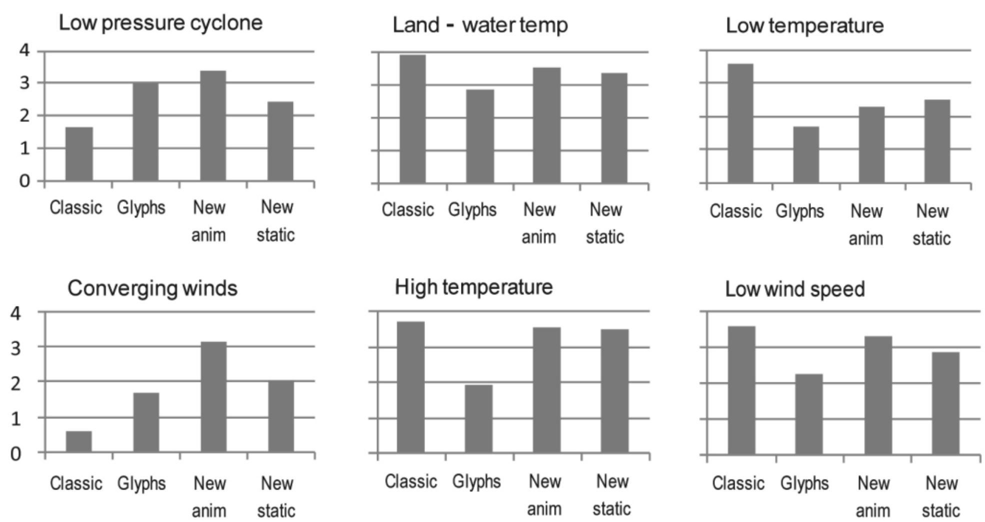


Figure 14. Subjects’ mean rating of how clearly various features could be seen with each of the four designs.

Two questions concerned wind movement patterns (Q1—*Low pressure Cyclone* and Q2—*Converging winds*). For both of them, the new *Animated* design with wind traces was rated as showing these significantly more clearly than the *Classic* design using wind barbs. The significant differences were as follows: Q1 (*Animated* > *Classic*, *Glyphs* > *Classic*); Q2 (*Animated* > *Classic*, *Animated* > *Glyphs*, *Glyphs* > *Classic*, *Static* > *Classic*).

Three questions dealt with temperature patterns (Q3—*Land-water temp*, Q4—*High temperature*, and Q5—*Low temperature*). In each of these the *Classic* design was generally rated highly and the *Glyphs* was rated poorly. The significant differences were as follows: Q3 (*Classic* > *Glyphs*, *Animated* > *Glyphs*); Q4 (*Classic* > *Glyphs*, *Static* > *Glyphs*, *Animated* > *Glyphs*); Q5 (*Classic* > *Glyphs*, *Classic* > *Animated*, *Classic* > *Static*, *Static* > *Glyphs*).

The remaining question (Q7—*Low wind speed*) concerned the perception of an area of low wind speed. In this case there was only one significant difference: Q7 (*Classic* > *Glyphs*). We attribute this to the way speeds of less than 2.5 knots are displayed using wind barbs through the use of a symbol that appears very distinct when barbs appear everywhere else.

9. DISCUSSION

The results suggest that our new design offers significant advantages for the display of wind information over the classic use of wind barbs. Subjects judged it much better at revealing the wind front defined by a line of converging winds and also better at showing the cyclonic pattern. In addition, objective measurements revealed greater accuracy in both judged wind speed and direction. Regarding the perception of temperature and pressure, the objective results showed little difference between the new designs and our re-designed version of the classic. In the case of temperature, this is not surprising since, in this respect, the two designs used identical color sequences. The subjective measurements, however, showed that the classic was judged to show temperature more clearly than the other methods, presumably because in the new designs texture obscured more of the temperature field, while in the glyph design the temperature field was fragmented.

In our design of a color sequence we found it impossible to create the kind of spiral sequence that monotonically increases in luminance that has been recommended because it shows visual form better[1,14]. Some of the reasons for this were as follows. We needed to resolve up to 25 five degree steps since our visualization covered temperatures zones from the sub-arctic to the southern United States. In order to distinguish adjacent colors, a significant luminance change is needed at each step in the sequence. Because the color sequence was a background for all other information, it was important to maintain luminance contrast with the numbers and the streamlets showing wind patterns and this limited the luminance range we could use. All of these combined to rule out a simple monotonic sequence. Nevertheless, in cases where a smaller number of colors are needed we still recommend the spiral sequence.

In the case of pressure, there were no differences between our new design and the classic, although the actual values came in one case from contours labeled with numbers, and in the other from our sequence of textures combined with a key. One way of looking at this result is that by using the sequence of textures we made numbers available for other purposes, and we chose to use them to represent wind speeds, which in turn resulted in more accurate readings of this variable. We attribute the more accurate readings of wind speeds to the simple fact that we used numbers that resolved one knot, significantly better than the 5 knot resolution provided by wind barbs. In any case, our use of texture to represent pressure proved to be a viable alternative to labeled contours .

Overall our new design proved the best for the perception of wind patterns, but the classic did not fare as poorly as we had expected. The use of contours to show pressure and color to show temperature are clearly effective and the study participants, most of whom were unfamiliar with wind barbs, nevertheless had little trouble reading them. This suggests that weather displays like the ones illustrated in Figure 1 could be substantially improved.

The glyph-based design was significantly worse on all measurements. Would another glyph-based design have fared better? We chose the Healey et al.'s [8] design because it had been specifically customized to meteorological data and we could adopt it with minimal changes. But other glyph techniques can be expected to suffer from some of the same problems[7,9,10,19]. None of the prior glyphs we are aware of, except for wind barbs, enable values to be read like our texture sequence. Instead they all rely on continuously varying values, such as the size of glyphs. Simultaneous contrast effects apply to particle size judgments, as much as color and lightness, so any display that uses element size in a dense pattern can be expected to produce large errors[15]. The same is true for orientation and elongation. This does not mean that glyphs should not be used. In fact our own design can be thought of as glyph-based, although we are not using glyphs to fuse variables. Instead we use animated glyphs for wind speed and the sequence of texture glyphs for pressure. In general, we suggest that glyphs that encode varables using simple mappings to attributes, such as size, are most suitable when fine discriminations are not needed and only qualitative judgments must be made. When quantitative as well as qualitative judgments are needed, then either numbers or some other discrete coding device should be added. This is what we did in the case of wind speed where speed of animation provides a qualitative indicator and the numbers provide quantitative information.

This project began as a design exercise and it should be evident that what emerged is highly customized to the application. In so far as our design, or the classic design, is successful, this can be attributed to the specific way that data variables (temperature, pressure, wind vectors) are mapped to display attributes. As a general lesson this suggests that good designs are likely to be highly customized. It matters very much that wind is mapped to animated motion and temperature is mapped to color, the reverse mapping would be extremely confusing. In addition much thought and many design iterations went into the exact design for the animated traces, for example determining that they should move at a rate proportional to the actual wind speed raised to the power of 0.65. The addition of contours to our texture sequence design came from a meteorologist's comment that showing pressure gradients is important. The designer must understand the application and the user's requirements in detail to ensure that the critical features are clear. Our design solution is dependent on animation. This is not suitable for delivery to all media, but the increasing prevalence of digital media makes animations, even of non time-varying data, more attractive as a design alternative. Nevertheless, we are working on alternative designs that are more suitable for the static representation of wind patterns (Pilar and Ware, submitted for publication).

We feel that the use of distinct perceptual channels was successful, although of course we cannot claim to have proven this. Many viewers of our public display have remarked on the effectiveness of the animated winds and how these allow for wind pattern features such as fronts to become visible is a way that they have not seen on other weather maps. We do not claim to have proven that our display is the best that could be made, but we do claim to have successfully created a map that represents two 2D scalar fields and a 2D vector field in a way that is highly legible and shows important patterns. We know of no other example that has been empirically tested to meet such rigorous requirements.

ACKNOWLEDGMENTS

The authors wish to thank John Kelly for advice on matters pertaining to meteorology and Alan Hubbe for implementing a preliminary version of this design. This work was supported by NOAA Grant # NA05NOS4001153.

REFERENCES

[1] D. Borland, and M.R.Taylor 2nd. Rainbow color map (still) considered harmful. IEEE Computer Graphics and Applications 2007 Mar-Apr;27(2):14-7 (2007).

[2] C. Cavina-Pratesi, R.W. Kentridge, C.A. Heywood and A.D. Milner. Separate Channels for Processing Form, Texture, and Color: Evidence from fMRI Adaptation and Visual Object Agnosia, Cerebral Cortex, 20 (10): 2319-2332. (2010)

[3] D. Fowler and C. Ware. Strokes for Representing Univariate Vector Field Maps. Graphics Interface Proceedings, 249-253. (1989).

[4] W.R. Garner. (1974) The processing of information and structure. Lawrence Erlbaum: Hillsdale, NJ.

[5] Gamedev.net discussion thread, "Terrain: Contour Lines using pixel shader." http://www.gamedev.net/topic/529926-terrain-contour-lines-using-pixel-shader

[6] H. Hagh-Shenas, S. Kim, V. Interrante, and C.G. Healey. Weaving Versus Blending: A Quantitative Assessment of the Information Carrying Capacities of Two Alternative Methods for Conveying Multivariate Data With Color." IEEE Transactions on Visualization and Computer Graphics 13, 6, 1270–1277 (2007).

[7] C.G. Healey, Formalizing Artistic Techniques and Scientific Visualization for Painted Renditions of Complex Information Spaces. In Proceedings International Joint Conference on Artifical Intelligence. 371-376. (2001).

[8] C.G. Healey, S. Kocherlakota, V. Rao, R. Mehta, and R. St. Amant, "Visual Perception and Mixed-Initiative Interaction for Assisted Visualization Design." IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 2 , pp. 396-411. (2008)

[9] G. Kindlmann and C.F. Westin. Diffusion Tensor Visualization with Glyph Packing. IEEE Transactions on Visualization and Computer Graphics, 12(5):1329–1336, (2006).

[10] D. H. Laidlaw, E.T. Ahrens, D. Kramers, M.J. Avalos, C. Readhead, and R.E. Jacobs, Visualizing diffusion tensor images of the mouse spinal cord. Proceedings of the IEEE Visualization Conference. 127-134. (1998).

[11] M.S. Livingston, and D.H. Hubel, Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. Journal of Neuroscience. 7, 3416-3468. (1987).

[12] M.S. Livingston, and D.H. Hubel, Segregation of form, movement and depth: Anatomy, physiology and perception. Science 240: 740-749. (1988)

[13] Theeuwes, J. and Kooi, F.L. (1994) Parallel search for a conjunction of contrast polarity and shape. Vision Research 34 (22) 3013-3016.

[14] C. Ware. Color sequences for Univariate Maps: Theory, Experiments and Principles. IEEE Computer Graphics and Applications 8(5), 41-49.(1988).

[15] C.Ware. Information Visualization: Perception for Design. Morgan Kaufman: San Francisco, (2004).

[16] C. Ware. Quantitative texton sequences for legible bivariate maps. IEEE Transactions on Visualization and Computer Graphics. 15(6) 1523-1529. (2009).

[17] T. Watababe and P. Cavanaugh. Texture laciness: the texture equivalent of transparency. Perception 25, 293-303, (1966).

[18] R.Wegenkittl and E. Gröller. Fast Oriented Line Integral Convolution for Vector Field Visualization via the Internet. In Proceedings IEEE Visualization '97, pages 309–316, Los Alamitos, Oct. 19–24 (1997).

[19] Weigle, C., Emigh, W., Liu, G., Taylor, R., Enns, J., and Healey, C. Oriented texture slivers: A technique for local value estimation of multiple scalar fields. Proc. Graphics Interface 2000, 163-170. (2000).