

# Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models

Hadley Wickham<sup>1</sup>, Heike Hofmann<sup>2</sup>, Charlotte Wickham<sup>3</sup>, Dianne Cook<sup>2</sup>

<sup>1</sup>Department of Statistics, Rice University

<sup>2</sup>Department of Statistics, Iowa State University

<sup>3</sup>Department of Statistics, Oregon State University

## Abstract

Understanding the multiple aspects of climate data is a challenge. There are both spatial and temporal components to examine, and often multivariate dependencies. This paper proposes a new display for climate data: glyph-maps. Glyph-maps are a specialization of multivariate glyph plots; each spatial location is displayed with one glyph that represents the measurements recorded over time at that location. Glyph-maps allow the discovery of both local and global structure, with a particular focus on temporal relationships.

The instructions for computing and displaying glyph-maps are described, along with the design of reference frames, and the effects of different scaling of the data on perception of structure. The methods are developed for rectangularly gridded data, but with some clever processing, which is explained, it is possible to get good glyph-maps of irregularly gridded data.

## 1 Introduction

Climate data is composed of measurements, such as temperature, precipitation, and winds, with a spatial and temporal context. The classic display for data of this type is “heatmap”, a tiled plot using color to display the value of the variable of interest at each location. When measurements are made at multiple time points it is common to display the data with small multiples, such as in Figure 1. In this plot, a separate map is drawn for each month (columns) and year (rows), over a six year period of remotely sensed temperature data above Central America (Murrell, 2010). Color is used to display de-seasonalized temperature, with red mapped to high values and blue to low. The most noticeable feature is the strong red patch in the equatorial

Pacific beginning mid-1997 and tapering out during 1998. This is the El Niño event, a major temperature anomaly. More work is required to see more localized patterns, like the cooler land temperatures seen in early years.

Reading this type of plot is cognitively challenging: the reader must play “spot the difference” from one small image to another. Large structures such as El Niño are clear but it is very difficult to mentally difference the images and read off the long-term trend or to notice local deviations. Using a movie to render these small multiples can help, but small trends still fail to draw readers’ attention and escape unnoticed (Simons et al., 2000).

This paper develops a new type of display to resolve these problems: the *glyph-map*. The glyph-map uses a small glyph, or icon, to represent multiple values at each location. Glyph maps are an adaptation of glyphs, tools developed for display multivariate data. Multivariate glyphs include the star plots of von Mayr (1877), the semi-graphic displays of Anderson (1960), and the infamous Chernoff faces (Chernoff, 1973). A glyph is produced by mapping each variable to some graphical feature, such as the length of a line. The glyph-map can also be thought of as a small multiple display Tufte (2001) of time series with a geographic context.

Figure 2 shows the glyph-map equivalent of Figure 1: each location is represented by a small time series glyph. The primary visible structure is still El Niño, the regions in the Pacific with bump in the middle of the series, but other features are also visible: temperatures are generally more varied over land and a few locations have dramatically increasing temperature in the later years typically in high elevation areas, along the Andes and the Sierra Madre Occidental in Mexico.

One problem with multivariate glyph is that lack of natural ordering of each glyph, or of the mapping between variables and glyph properties. This leads to a combinatorial explosion of possibilities, each of which may have dramatically different perceptual properties. This problem has prevented the widespread adoption of glyph displays, despite some proposed remedies (Kleiner and Hartigan, 1981; Hurley and Oldford, 2010). Using glyphs with space-time data eliminates these problems: glyphs are placed according to the location of the measurement, and variables are ordered by the time they were collected. Additionally, climate data usually has high correlations between nearby locations and times. This imposes a degree of smoothness that gives the plot an appearance of a textured landscape, making it easier to digest patterns. Others Pickett and Grinstein (1988) have used glyphs on maps to display multivariate spatial data. This area of work has morphed into the field of metaphorical data displays, which create abstract landscapes of spatial data, a digression from glyph-maps. Gribov et al. (2006) describes the use of glyph-maps for multivariate data also, with emphasis on the graphics software Gauguin.

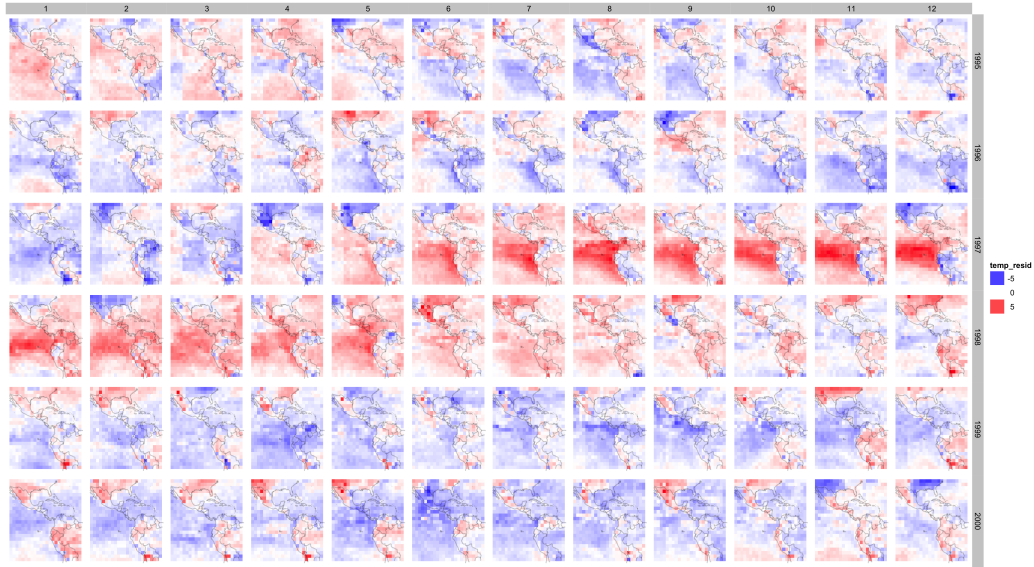


Figure 1: Facetted map of de-seasonalized temperature. The dominant feature is the El Niño warming in the southern equatorial region in the last half of 1997 and first half of 1998. Smaller features are only noticeable on closer inspection, or if pointed out: such as the relative warming on the mountain regions in south and north America in the later years.

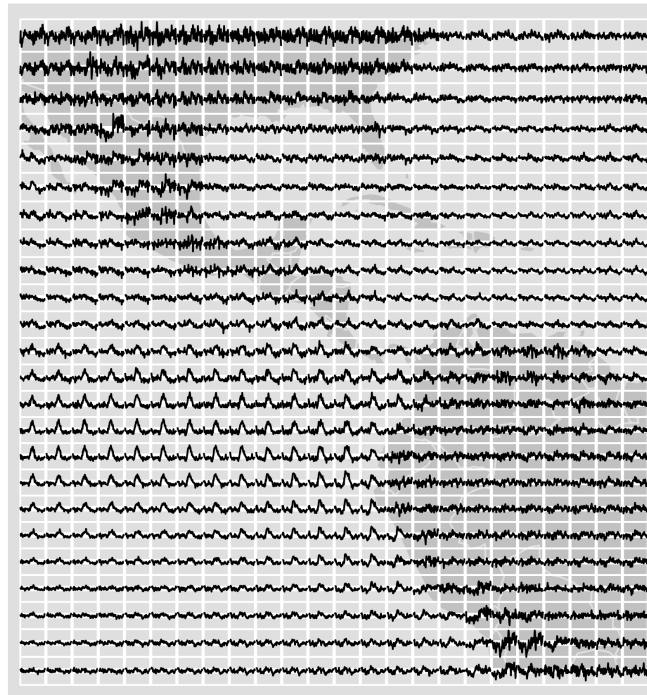


Figure 2: Glyph-map of de-seasonalized temperature, to compare with Figure 1. Time series of the six years of monthly temperature are plotted at each spatial grid location. The El Niño appears as a bump in the middle of the time series, in the equatorial Pacific region. Large variations in temperature can be seen in areas over land, while being fairly constant over water.

Two types of glyph – lines and stars – are especially useful for temporal displays. Figure 3 displays 12 iconic time series shapes with line- and star-glyphs. The data underlying each glyph is measured at 36 time points. The line-glyphs are time series plots. The star-glyphs are formed by considering the 36 axes radiating from a common midpoint, and the data values for the row are plotted on each axis relative to the locations of the minimum and maximum of the variable. This is a polar transformation of the line-glyph.

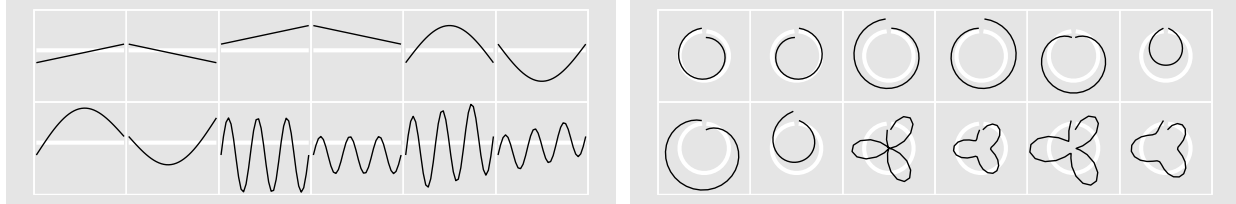


Figure 3: Icon plots for 12 iconic time series shapes (linear increasing, decreasing, shifted, single peak, single dip, combined linear and nonlinear, seasonal trends with different scales, and a combined linear and seasonal trend) in Euclidean coordinates, time series icons (left) and polar coordinates, star plots (right).

The paper is structured as follows. Section 2 describes the algorithm used to create glyphs-maps. Section 3 discusses their perceptual properties, including the importance of a visual reference grid, and of carefully consideration of scale. Large data and the interplay of models and data are discussed in Section 4. Many spatiotemporal data sets have irregular spatial locations, and Section 5 discusses how glyph-maps can be adjusted for this type of data. Three datasets are used for examples:

**data-expo** The ASA 2009 data expo data (Murrell, 2010) consists of monthly observations of several atmospheric variables from the International Satellite Cloud Climatology Project. The dataset includes observations over 72 months (1995–2000) on a 24 x 24 grid (576 locations) stretching from 113.75°W to 56.25°W longitude and 21.25°S to 36.25°N latitude.

**GISTEMP** surface temperature data provided on 2° x 2° grid over the entire globe, measured monthly (Earth System Research Laboratory, Physical Sciences Division, National Oceanic and Atmospheric Administration, 2011). Ground station data was de-seasonalized, differenced from from the 1951-1980 temperature averages, and spatially averaged to obtain gridded measurements. For the purposes of this paper, we extracted the locations corresponding to the continental USA.

**USHCN** (Version 2) ground station network of historical temperatures (National Oceanic and Atmospheric Administration, National Climatic Data Center, 2011). Temperatures from 1219 stations on the contiguous United States, from 1871 to present.



Supplementary material contains the datasets and code to reproduce the plots in this paper. R 2.13.1 (R Development Core Team, 2011) with the packages `ggplot2` (Wickham, 2009) and `plyr` (Wickham, 2011) were used.

## 2 Construction

Creating a glyph-map is elementary. It involves recognizing that they are a linear mix of two structural components of the data: spatial location and data values. The spatial location is the major positioning component, while the data values are minor adjustments to those positions. For spatiotemporal data, the major axes are latitude ( $y_{\text{major}}$ ) and longitude ( $x_{\text{major}}$ ), and the minor axes are time ( $x_{\text{minor}}$ ) and some measurement ( $y_{\text{minor}}$ ), for example, temperature, or predicted temperature. Assuming the minor axes are rescaled to  $[-1, 1]$ , the final coordinates  $(x, y)$  on the chart are the linear combination, given by:

$$\begin{aligned} x &= x_{\text{major}} + \frac{w}{2} \cdot x_{\text{minor}} \\ y &= y_{\text{major}} + \frac{h}{2} \cdot y_{\text{minor}}, \end{aligned} \tag{1}$$

where  $w$  and  $h$  are maximal width and height of the glyph, respectively. For gridded data, it is convenient to make use of the resolution of the major axes, that is, the minimum difference between neighboring locations.

Since the coordinates are linear combinations of two variables, it is possible to build these line-glyphs interactively, illustrated by software packages that have incorporated tours (as described in Cook et al. (2008)), such as DataViewer (Buja et al., 1986), XGobi (Swayne et al., 1991) or GGobi (Swayne et al., 2003). This technique is used to good effect in Buja et al. (1996), and was used to examine the climate data in Hobbs et al. (2010).

Coordinates for star glyphs are computed with a polar transformation, with minor axes scaled to  $[0, 1]$ :

$$\begin{aligned} x &= x_{\text{major}} + y_{\text{minor}} \cdot \frac{w}{2} \sin(2\pi x_{\text{minor}}) \\ y &= y_{\text{major}} + y_{\text{minor}} \cdot \frac{h}{2} \cos(2\pi x_{\text{minor}}). \end{aligned} \tag{2}$$

This is a non-standard conversion to polar coordinates, but it creates a timeline that starts at 12 o'clock and proceeds clockwise.

Figure 3 gives some small examples of time series with prototypical trends, shown both in cartesian coordinates and polar coordinates. Differences between linear and nonlinear trend are more apparent in the line-glyphs, and are effectively lost in the star-glyphs, which are most effective in exposing cyclical patterns

due to seasonality. Star-glyphs show seasonality as floral cartoons, with peaks forming petals. It is surprising that time series with opposing trends in line-glyphs don't show this symmetry when displayed with star-glyphs. The area of star-glyphs coordinates mainly reflect the average value of a time series, while its shape shows deviations from the average.

### 3 Perception

All plots facilitate some comparisons and impede others. For a plot to be useful for a particular data analytic task, primary comparisons need to be made the simplest by the display. For climate data, changes in slope, or trend, average value, and variance over time, are the primary tasks. Glyph-maps support these tasks because the pieces of the graphical elements that are needed to make the temporal comparisons are organized, and grouped together.

Glyph-maps allow time trends to be read directly from the plot. Values are mapped to position, one of the easiest properties to perceive (Cleveland and McGill, 1984). Perceiving time trends in faceted heatmaps is much more difficult: not only do you need to read value from color (a relatively difficult task), you also need to spot the difference between the different maps. This is challenging from a cognitive perspective (Healey and Enns, 2011; Busey, 2011) and comparisons can suffer from *change blindness*, often leading to silent failure. To see this, examine Figures 1 and 2. The heatmap requires comparing color values across maps to assess temporal change but the glyph-map allows direct reading of the temporal trend at each location, and comparison across spatial neighbors.

One draw-back of the glyph-maps, particularly if linear trend is displayed as an icon, is they may suffer from the Zöllner Illusion (Zöllner, 1860), visible in Figure 8, which makes straight lines look crooked.

An third alternative is to calculate the relevant statistic for each location and display these values as single heatmap. For example, to study long-term trends, we could calculate the slope of a linear model at each location and color tiles across the map by the value of the slope. The disadvantage of this approach is that it is not exploratory, and requires that we have a very precise definition of the measurement we are interested in, and it still suffers from the relatively difficult color perception.

Two other factors are critical for accurate perception of change: reference frames and scaling. These are described in the following two sections.

### 3.1 Reference frames

The structured spatial arrangement of icons in a glyph-map helps to compare the patterns of shape, like slope, intercept, or size, across icons. However, additional clues can make comparisons easier, converting the perceptual task from comparing length to the easier position along a common scale (Cleveland and McGill, 1984). This is also called a *visual reference grid* (Cleveland, 1993).

Each glyph is small, so there is not enough space for a full set of axes. Instead, minimal reference lines and boxes are incorporated. These need to be minimally perceptible, post-attentive (Healey, 2011) and de-emphasized, so as not to detract from the data. The reference grid is a box, framing each icon, which represents the spatial grid. The reference line is a horizontal line at mid-range. Both help to read differences in slope and intercept. Slope is read by comparing the position of left and right endpoints on the frame, or by reading the angle between the data line and the reference line. Intercept is read from “average” position of the line in the box: is it near the top, or near the bottom? Figure 4 demonstrates both guides. Reference elements are drawn in white, so they are minimally perceptible relative to the black data.

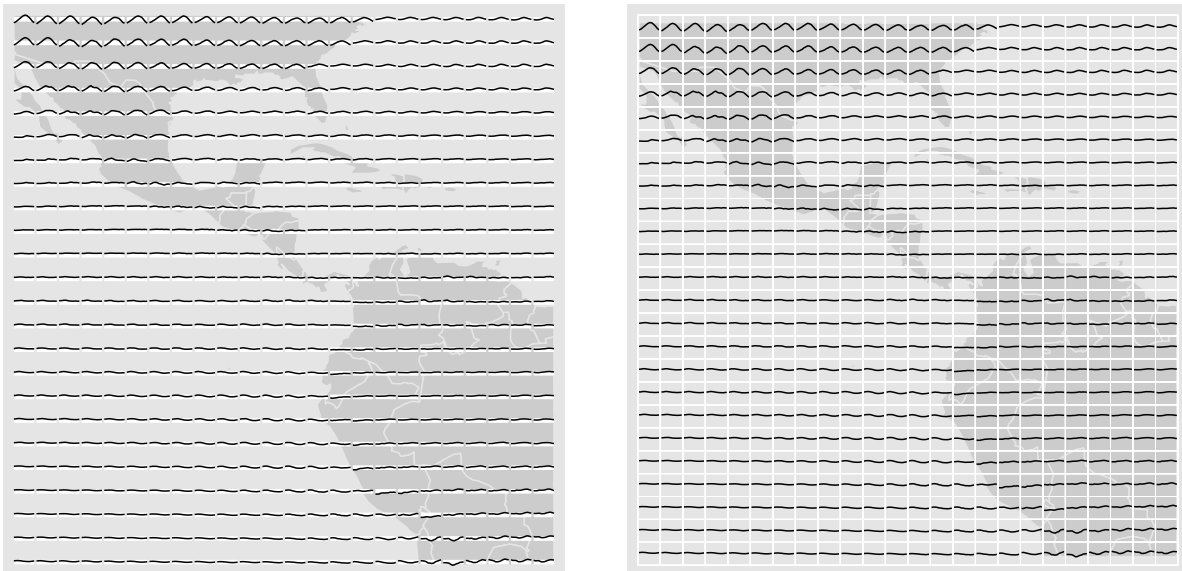


Figure 4: Glyph-maps of seasonal temperature patterns (averages for each month), using the data-expo data. Adding mid-range reference lines (left) and grid-cell reference boxes (right), makes it easier to see differences in glyph position, not just shape.

For star glyphs the reference frames are also useful and the equivalent of the reference line is a circle. Figures 3 and 7 show more examples.

### 3.2 Scaling

Different features of the data can be emphasized by varying the scale used with each cell. By default, a global scale is used, so the same position within each cell corresponds to the same value in all locations. This facilitates the comparison of absolute values, and draws attention to plots with large overall variation.

Alternatively, local scale can be used. In local scaling, the values within each location are scaled to range  $[0, 1]$  before plot construction. This makes it easier to compare shape (ignoring amplitude), and draws attention to locations with large relative variation. Figure 5 compares global scaling and local scaling for the smoothed temperature data from the data-expo data. On the global scale, the El Niño blip in temperature is just visible, along with variation in the average temperature at each location. Local scaling emphasizes the individual shapes: the impact of El Niño can be seen to be over a wider region, dramatic linear increases across the Andes, and decreases in the Gulf of Mexico.

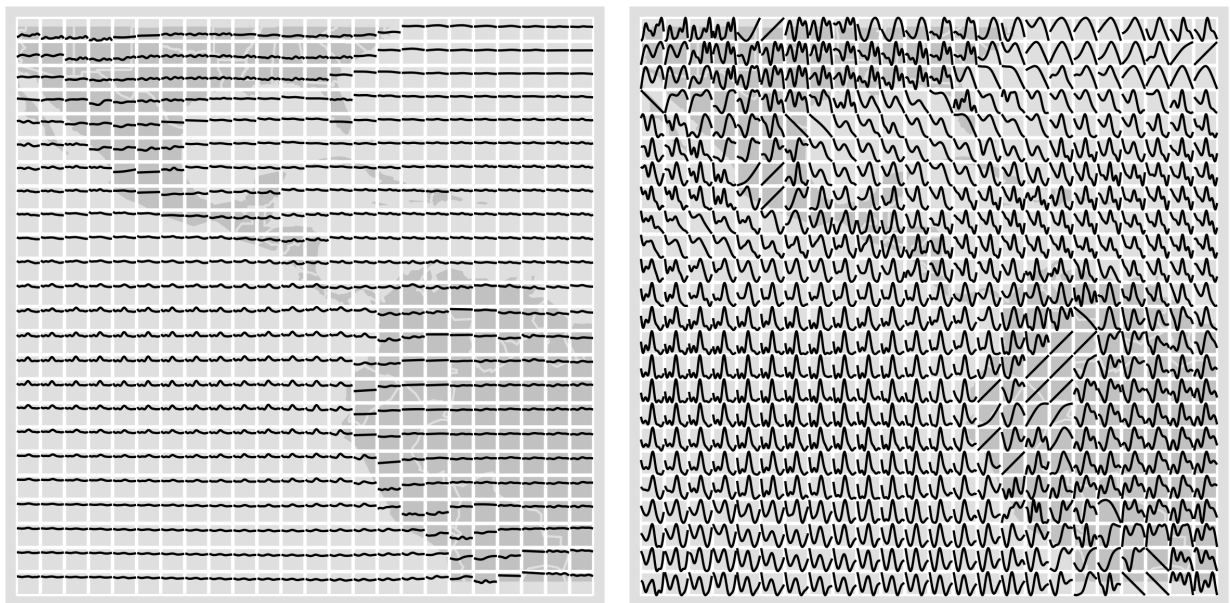


Figure 5: Smoothed de-seasonalized daily temperature glyph-map. (Left) globally scaled, (right) locally scaled.

Other types of shifting and scaling can be useful:

- The mean and standard deviation (or robust equivalents) could be used to standardize values to common moments.
- Scaling to maximum 1 (but leaving the minimum unchanged) can help emphasis the relative shape,

but not distort the range as much as rescaling to range  $[0, 1]$ .

- Shifting the values at each location to have mean zero places more focus on the trend. This is particularly important for the star glyphs, where differences in the average value lead to unintuitive patterns in the resulting glyphs.

Note that the use of local scales needs to be clearly marked. The viewer must realize that the scale is not relative and that big patterns in some locations might be just tiny effects. To indicate this on the locally scaled data plots, one can use an additional aesthetic, like color, to encode the range of original scale. This is illustrated in Figure 6.

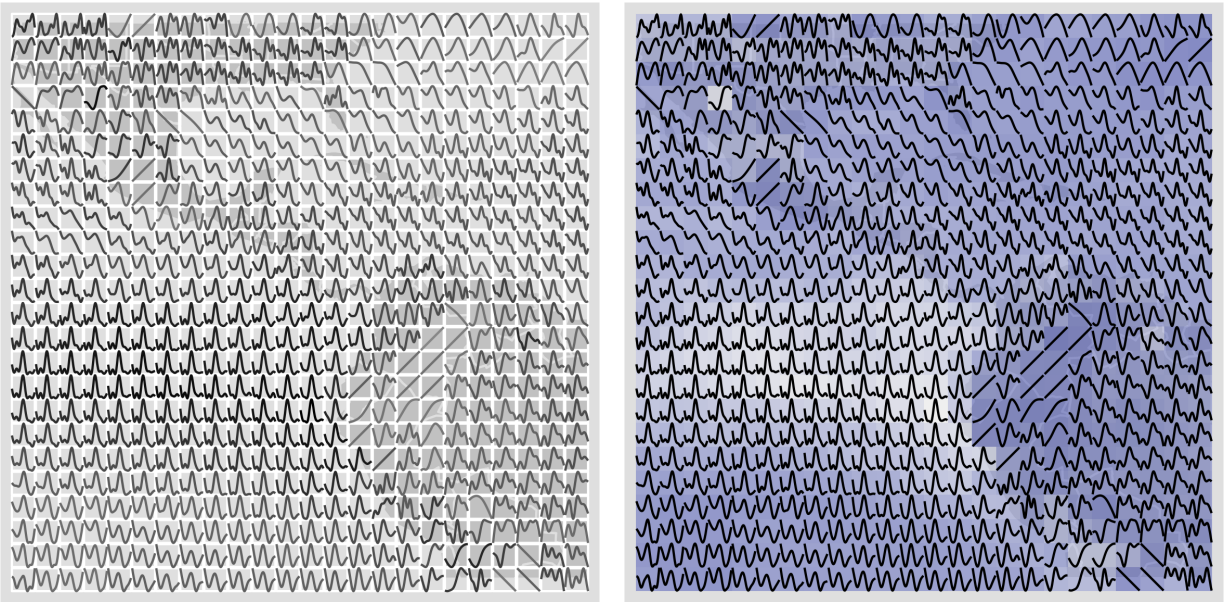


Figure 6: Adding intensity and color to scaled plots. Here each location has been scaled to range  $[0, 1]$  and color mapped to the range of the original predictions. (Left) Range mapped to intensity of line: pale locations have smaller range than darker locations. (Right) Range mapped to fill color of grid box: blue boxes have smaller ranges. Shapes at all locations are visible, but attention is drawn to locations with larger ranges.

One more important point: if a model summary is displayed in the glyph-map, as discussed in the following section, it may be important to force the scale to that of the original data. The range of predicted values is typically smaller than the raw data, so angles of slopes is increase.

## 4 Large data and models

Climate data is usually large: many locations in space, and many points in time. This makes climate data arduous to work with computationally, and to visualize. Careful use of models to decompose data into long-

term trend, seasonal effects and error is vital for to make the problem both computationally and visually tractable.

Figure 8 shows the a glyph-map for the large GISTEMP dataset. Glyphs at each location represent temperature from 1880–2011, starting with Jan 1880 at 12 o'clock and proceeding clockwise to 2011 at 11:59. The most important feature to notice is the pattern of missing values. Most glyphs do not complete the full circle, typically missing points between 12 and 1, suggesting missing values in the earliest part of the data. Another important feature is variability, shown by the thickness of the circle. Some locations, particularly in the north western mountain region, have large variations in temperature.

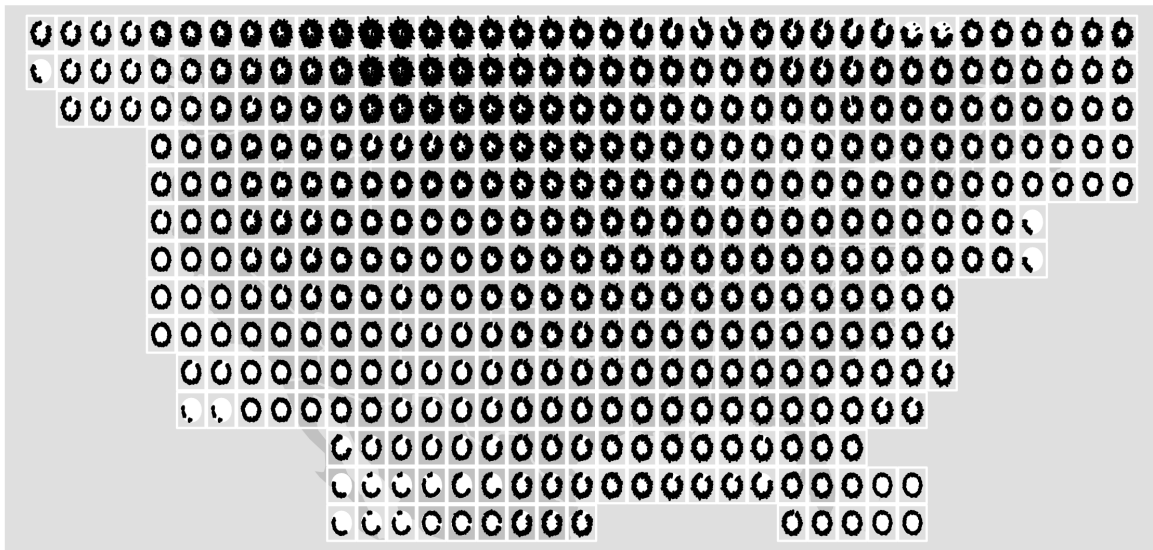


Figure 7: Glyph-map of raw temperature anomaly data (GISTEMP), from 1880-2011. Gaps in the full circle indicate missing values, and thickness of the circle indicates more variability in the measurements.

The raw data shows us missing values and gives some sense of variability, but gives no information about long-term trend. Trend can be examined by fitting a linear model to each location and displaying the predicted values. Figure 8 shows this for 1950–2010 (limited time period used due to missing data). Most locations have increasing temperature over this time period. Only a few locations show a decline. These locations tend to be different from their neighbors, which might suggest isolated data problems. The rate of increase differs across regions: steeper inclines in the north and western mountain regions and across the Great Lakes, and flatter inclines in the southeastern USA.

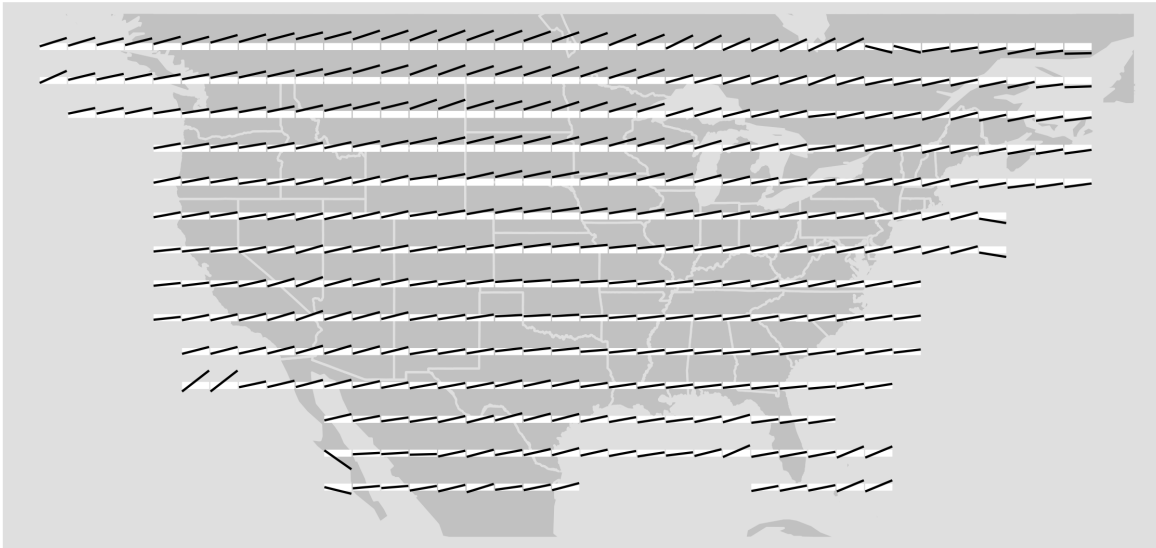


Figure 8: Glyph-map of predicted temperature anomaly data (GISTEMP), from 1950-2010. Increasing trends can be seen over most of the USA, with varying degrees of increase. Some isolated locations exhibit steep declines, and inclines, perhaps indicative of some isolated data problems. Note that, the reference lines are straight, but they look curved because of the Zöllner optical illusion.

## 5 Non-gridded data

Regularly gridded spatial data helps with the perception of structure, because it helps to provide the reference frames upon which to make comparisons among the icons. When spatial locations are not on a regular grid, some difficulties arise. In this situation it is not clear what the scale of icons should be, there is no regular spacing. Some locations may be close to each other which would generate icons that overlap.

### 5.1 Non-rectangular grids

Glyph-maps work best when displayed in the coordinate system in which the grid is rectangular and regular. For example, the GISTEMP grid is  $2^\circ$  square and is displayed treating longitude and latitude as cartesian coordinates in Figure 8. For regularly gridded data, this results in icons that are rectangles of equal size. It is often preferable to display maps as a projection of the geographical coordinates which preserves a property of interest, area for example. When a projection is desired a choice needs to be made between applying the projection to the completed glyph-map, or to just the spatial locations. When the completed glyph-map is projected the icons are no longer rectangles of the same size but they do retain their space filling property.

When applying a projection to the spatial locations and then constructing a glyph-map, the grid becomes irregular.

## 5.2 Irregular locations

Figure 9 shows linear trends for monthly average surface temperature at stations in the USHCN data. It illustrates three challenges in working with irregular locations: there is no natural choice for icon size, icons can potentially overlap, and comparisons rely more heavily on the reference guides. The icon sizes were chosen manually for these two plots, although the icons are displayed at about the same size, they correspond to  $1^\circ$  and  $0.5^\circ$  square respectively in the geographical coordinates. One solution to the overlapping problem

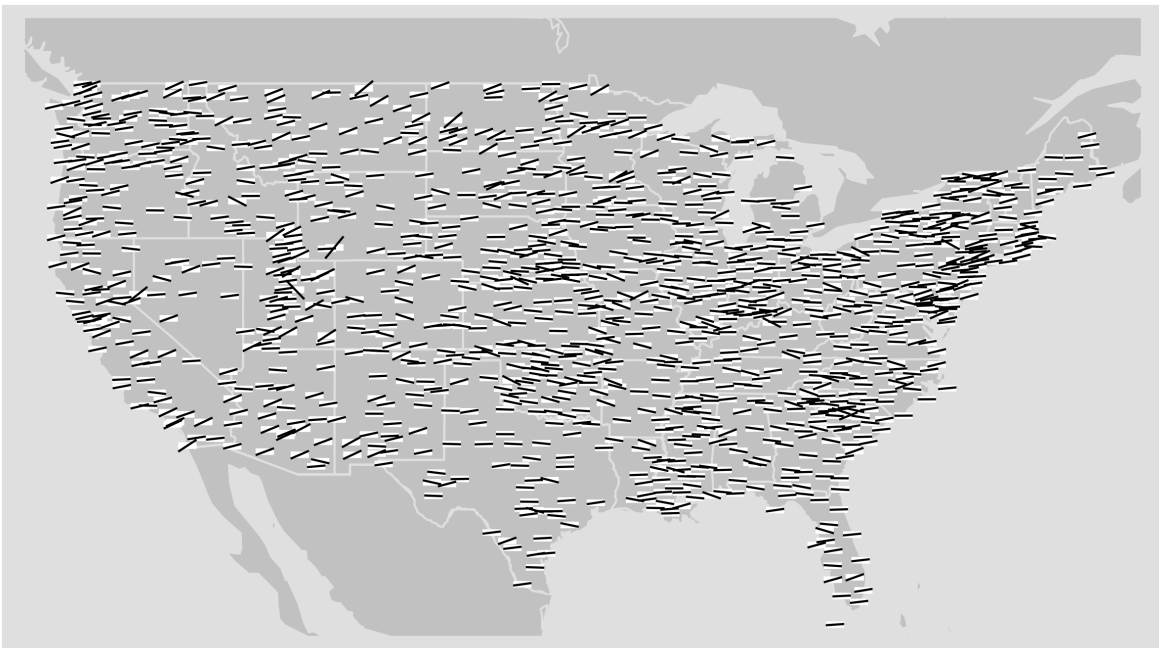


Figure 9: Glyph-maps of the linear trend in ground station temperature values, 1950–2010, for USHCN stations, illustrating three issues for irregularly gridded data: overlapping icons, no natural choice for icon size, and the disordered appearance making comparisons more difficult.

is to zoom in on smaller regions. There is an interplay between icon size (in geographical units), plotting area (in geographical units) and plot size (in display units). Icons need to be big enough to be readable on the display, but small enough to avoid overlapping. Readability can be maintained and overlapping reduced by decreasing the icon size and simultaneously increasing the display size or decreasing the plotting area. However, there are generally limitations in the maximum display size (or resolution) and the utility of plotting very small areas.



An alternative approach is to combine overlapping locations and plot a single summary icon. In Figure 10 nearby locations are collapsed to a single icon. In the top panel each station still has an individual glyph but locations close to each other share a plotting area. Not all glyphs are suited to this kind of space sharing, another option is to perform a summary of the locations that are collapsed and display that instead of the raw data. The bottom panel shows an example of this where the average trend for the collapsed locations is displayed. It is helpful to differentiate glyphs that involve summaries of more than one location, here a heavier line indicates at least two locations have been averaged.

In both examples of collapsing locations, the icon locations still represent real geographical locations of data, but contain data from up to an icon width away. An alternative way to combine stations is to round locations to the icon size, Figure 11. This produces a regular grid with the advantage of easier perception of structure. However, it hides the irregular nature of the locations and icons may end up centered over nonsense places (i.e. land measurements in the ocean).

Reference guides become especially important in the irregular location case. Comparisons are harder in this case, as the ability to easily compare icons along a common baseline, be it along a row or column of glyphs, is absent. Without reference it is hard to tell if short series are missing data at the start or end.

## 6 Generalizations

Glyphs can be more general than time series or stars. For multivariate data we may want to represent several variables in the glyphs. Figure 12 shows scatterplots of the data-expo data as icons: temperature values are displayed horizontally and high-cloud values are displayed vertically, both locally scaled. The bivariate relationship between the two variables is explored spatially, and it is clear that it is different across the region. The equatorial Pacific and Caribbean have positive association between temperature and high cloud, while the south American continent has negative association. The second plot displays a loess curve fit to the data instead of the individual points, sharpening the view of association. Another option is available in Gauguin (Gribov et al., 2006), a icon based on a histogram, giving small univariate distribution displays for each location.

## 7 Conclusions

In summary, glyph-maps enable the study of temporal patterns in multivariate spatiotemporal data. Climate change is focused on changes over time, so glyph-maps provide a way to study these changes directly. Glyph-

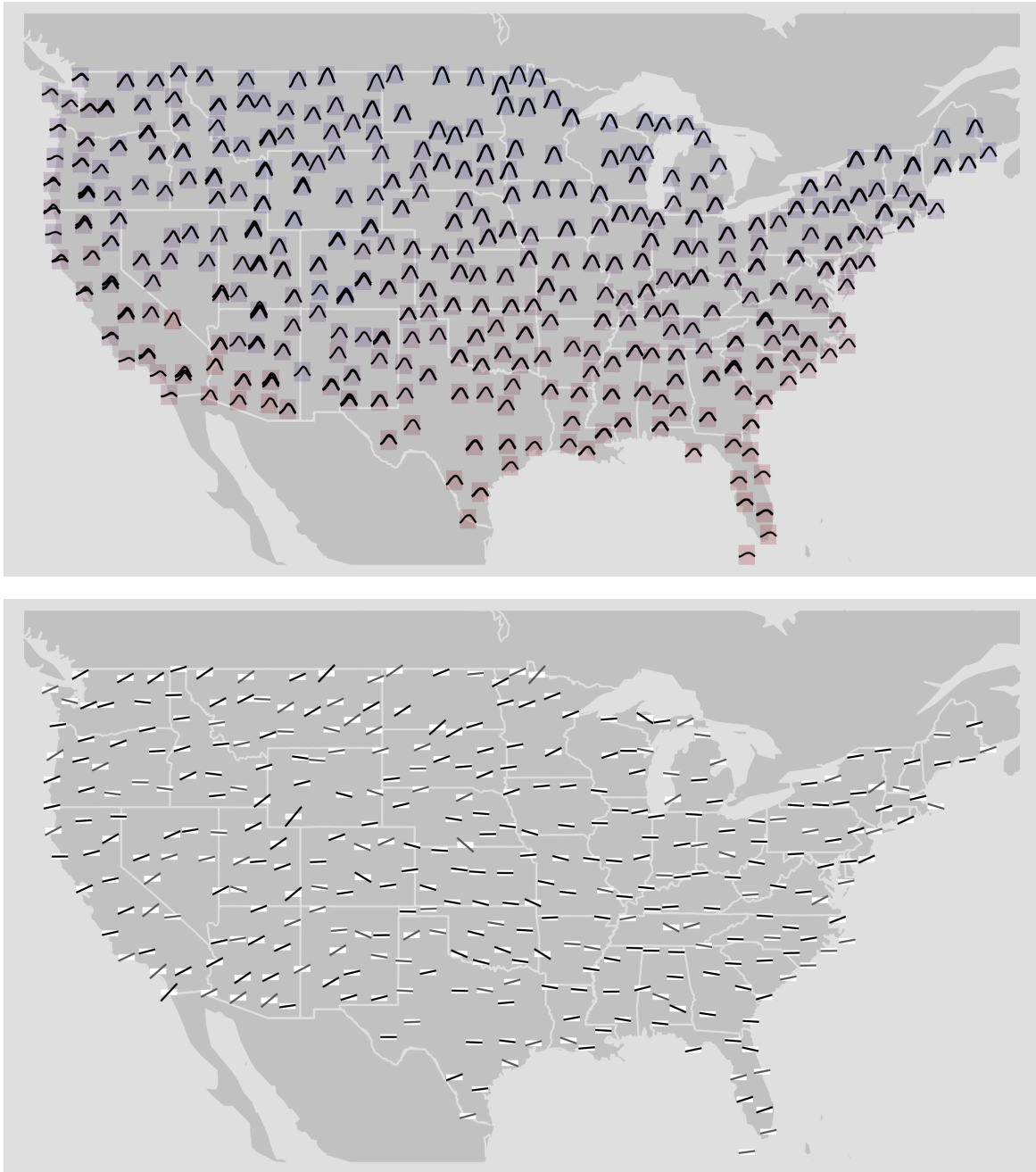


Figure 10: Glyph-maps of seasonal patterns (top) and linear trends (bottom) of USHCN stations, 1950–2010. Nearby stations whose icons would overlap have been collapsed. Each location has a glyph in the seasonal plot and color of the tiles corresponds to the average temperature of the locations (red=warm, blue=cool). Heavy lines in the linear trend plot correspond to average trends over more than one location.

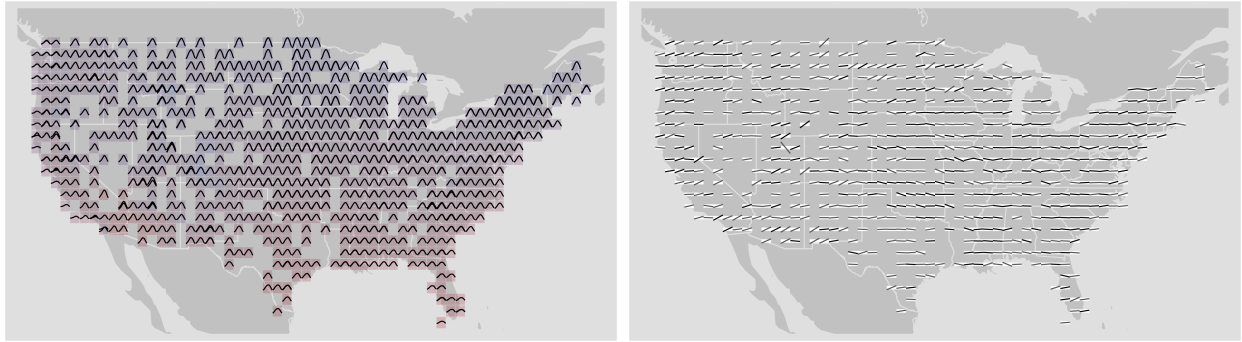


Figure 11: Glyph-maps of seasonal patterns (left) and linear trends (right) of USHCN stations, 1950–2010. Station locations have been rounded to the nearest degree. Each location has a glyph in the seasonal plot and color of the tiles corresponds to the average temperature of the locations (red=warm, blue=cool). Heavy lines in the linear trend plot correspond to average trends over more than one location.

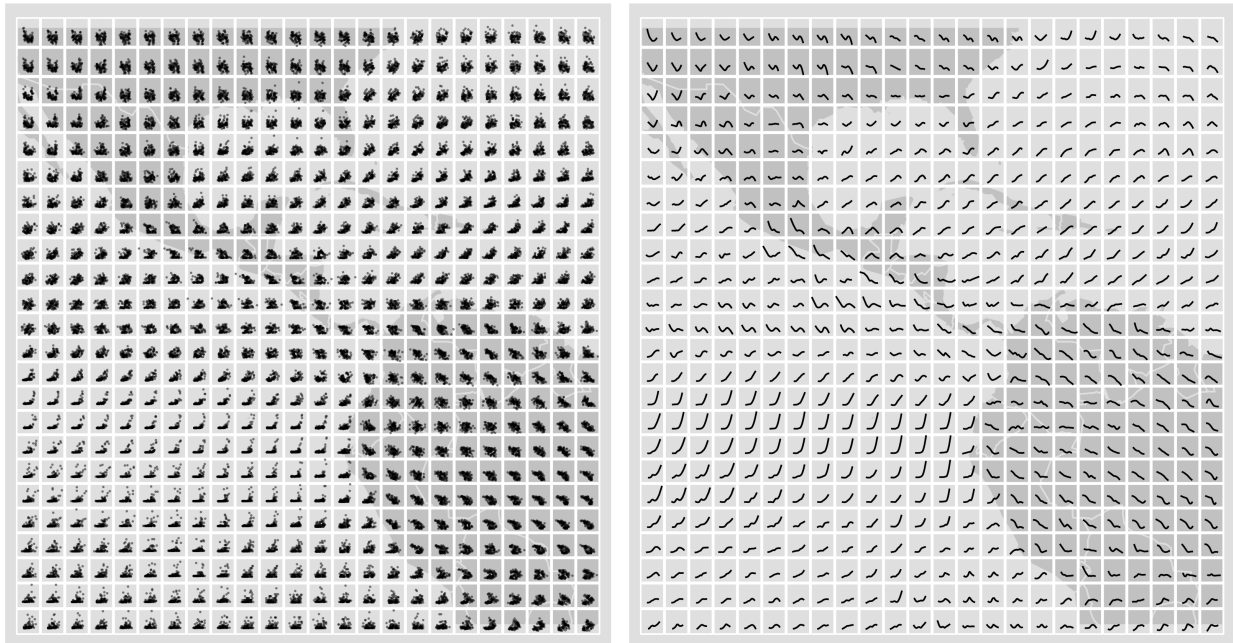


Figure 12: Glyph-map showing (left) scatterplot of temperature vs. high-cloud and (right) smoothed (loess) curve fit to each location. The relationship between these two variables varies considerably over the spatial domain.

maps enable different resolutions of the data to be examined: the raw data to discover data quality issues, early, global trend, seasonality, residuals, and multivariate dependencies. When the data is spatially gridded the glyphs are organized in a manner that makes reasonable comparisons. For irregularly gridded data the icon size needs to be chosen to minimize overlap but use as much display space as possible, or combined to produce close to gridded icons.

The next direction would be to examine the perceptual elements of these displays, in more detail. For example, it appears that periodic trends are easier to perceive in star-glyphs, and long-term trends in line-glyphs. Is this always the case? Are there certain types of periodicity that are particularly easy to spot? These questions need to be answered with rigorous perceptual studies to help guide the use of these plots.

## References

- E. Anderson. A Semigraphical Method for the Analysis of Complex Problems. *Technometrics*, 2(3):387–391, 1960.
- A. Buja, C. Hurley, and J. A. McDonald. A Data Viewer for Multivariate Data. In T. Boardman, editor, *Computing Science and Statistics: Proceedings of the 18th Symposium on the Interface*, pages 171–174, Washington, DC, 1986. American Statistical Association.
- A. Buja, D. Cook, and D. F. Swayne. Interactive High-Dimensional Data Visualization. *Journal of Computational and Graphical Statistics*, 5(1):78–99, 1996.
- T. Busey. Change blindness. <http://cognitron.psych.indiana.edu/CogsciSoftware/ChangeBlindness/>, 2011.
- H. Chernoff. The Use of Faces to Represent Points in K-Dimensional Space Graphically. *Journal of the American Statistical Association*, 68(342):361–368, 1973.
- W. Cleveland. A Model for Studying Display Methods of Statistical Graphics. *Journal of Computational and Graphical Statistics*, 2:323–364, 1993. URL <http://stat.bell-labs.com/doc/93.4.ps>.
- William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

- D. Cook, E.-K. Lee, A. Buja, and H. Wickham. Grand Tours, Projection Pursuit Guided Tours and Manual Controls. In Chunhouh Chen, Wolfgang Härdle, and Antony Unwin, editors, *Handbook of Data Visualization*, Springer Handbooks of Computational Statistics, chapter III.2, pages 295–314. Springer, 2008.
- Earth System Research Laboratory, Physical Sciences Division, National Oceanic and Atmospheric Administration. Monthly temperature anomaly. <http://www.esrl.noaa.gov/psd/>, 2011.
- A. Gribov, A. Unwin, and H. Hofmann. Tools for Multivariate Data Exploration. *ASA Statistical Computing and Graphics Newsletter*, 17(2):18–21, 2006.
- C. G. Healey. Perception in Visualization. <http://www.csc.ncsu.edu/faculty/healey/PP/index.html>, 2011.
- C. G. Healey and J. T Enns. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 2011.
- J. Hobbs, H. Wickham, H. Hofmann, and D. Cook. Glaciers Melt as Mountains Warm: A Graphical Case Study. *Computational Statistics*, 25(4):569–586, 2010.
- C. B. Hurley and R. W. Oldford. Pairwise Display of High-Dimensional Information via Eulerian Tours and Hamiltonian Decompositions. *Journal of Computational and Graphical Statistics*, 19(4):861–886, 2010.
- B. Kleiner and J. A. Hartigan. Representing Points in Many Dimensions by Trees and Castles. *Journal of the American Statistical Association*, 76(374):260–269, 1981.
- P. Murrell. The 2006 Data Expo of the American Statistical Association. *Computational Statistics*, 25:551–554, 2010.
- National Oceanic and Atmospheric Administration, National Climatic Data Center. Us historical climatology network. [http://cdiac.ornl.gov/ftp/ushcn\\_v2\\_monthly/](http://cdiac.ornl.gov/ftp/ushcn_v2_monthly/), 2011.
- R. M. Pickett and Georges G. Grinstein. Iconographics displays for visualizing multidimensional data. In *Proc. IEEE Conference on Systems, Man, and Cybernetics*, pages 514–19, 1988.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. URL <http://www.R-project.org/>. ISBN 3-900051-07-0.
- D. J. Simons, S. L. Franconeri, and R. L. Reimer. Change blindness in the absence of a visual disruption. *Perception*, 29:1143–1154, 2000.

- D. F. Swayne, A. Buja, and .N Hubbell. XGobi meets S: integrating software for data analysis. *Computing Science and Statistics*, 23:430–434, 1991.
- D. F. Swayne, D. Temple Lang, A. Buja, and D. Cook. GGobi: Evolving from XGobi into an Extensible Framework for Interactive Data Visualization. *Computational Statistics & Data Analysis*, 43:423–444, 2003.
- E. R. Tufte. *The Visual Display of Quantitative Information (Second Edition)*. Graphics Press, Cheshire, CT, 2001.
- G. von Mayr. Die Gesetzmässigkeit im Gesellschaftsleben. 1877.
- H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. useR. Springer, July 2009.
- H. Wickham. The Split-apply-combine Strategy for Data Analysis. *Journal of Statistical Software*, 40(1):1–29, 2011. URL <http://www.jstatsoft.org/v40/i01/>.
- F. Zöllner. Ueber eine neue Art von Pseudoskopie und ihre Beziehungen zu den von Plateau und Oppel beschriebenen Bewegungsphänomenen. *Annalen der Physik*, 186:500–523, 1860. doi: 10.1002/andp.18601860712.