

# Dynamic Small Multiples

Shannon Kao\*

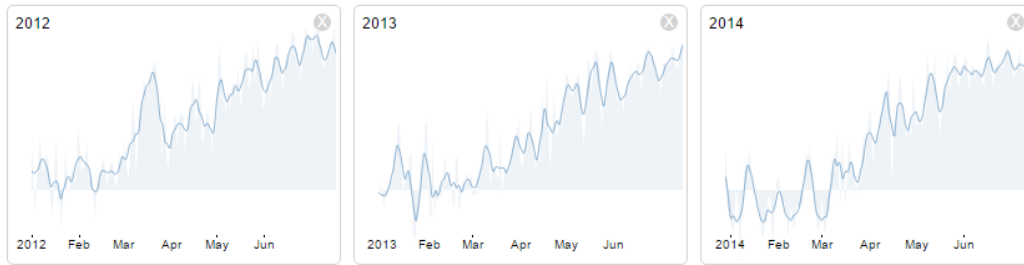


Figure 1: Small multiples displaying the maximum temperature in Detroit over the first six months of the year, for a three year span.

## 1 Introduction

Small multiples are a unique and efficient way to communicate information about a data set. They provide an easy avenue for comparison between distinct points, a data-rich display that highlights differences for trend analysis, and an intuitive interface for readers. However, small multiples are rarely used in interactive visualizations, due to their limitation in size. A small multiple display is generally focused on overarching trends, and does not offer much room for detail, or specific data values.

Significant research has been conducted on interactivity within small multiples. Research has found that small multiples and animated displays have different, often complementary strengths. However, there has been relatively little work that looks into combining these disparate visualizations in order to better analyze data.

The proposed Dynamic Small Multiples system allows users to rapidly generate and iterate over various small multiple displays, while simultaneously viewing the display in a large focus window. It combines the idea of an interactive focus-and-context view with a small multiples view of the data, augmenting the high-detail aspects of the focus chart with the communicative and exploratory powers of small multiples.

The implemented system uses a data set consisting of the daily temperature for six years. Using a large, highly cyclic data set showcases the strengths of being able to quickly generate multiples while exploring the data at a high level of precision.

### 1.1 Data

The weather data is a 6-year span of the daily maximum temperature in Detroit, pulled from the National Oceanic and Atmospheric Administration.

## 2 Related Work

The concept of dynamic small multiples touches on the idea of viewing the same data set at multiple scales. Fundamentally, the goal is to compare different intervals of the same continuous data set, and there is interesting work to consider in other techniques for visualizing multi-scale data. Additionally, there has been research on adding interactivity into small multiples themselves.

### 2.1 Multi-scale Data

Visualizing and analyzing multi-scale time data has been previously explored from a variety of different angles. Researchers have developed fascinating new visualization methods, from the ripple graph [Cho et al. 2014], to data compression [Reeves et al. 2009], to wavelet transforms [Woodring and Shen 2009], in order to better cluster and extract trends from data.

The system I propose does not seek to visualize the entirety of a large time-varying data set, but rather enable the user to interactively view and explore cyclic time-based data at various intervals.

### 2.2 Combined Visualization

Research has also been done in the area of combining different types of visualization, to form a more effective final product. There have been multiple studies on brushing and linking data [Doleisch et al. 2004], as well as in multi-window displays for plotting related graphs [Swayne et al. 2003].

Visualizations with different types of graph augmenting each other have been shown to be extremely effective at communicating information. The Dynamic Small Multiples system hopes to take advantage of this, using two separate types of visualization to augment the weakness of each.

### 2.3 Small Multiples

In terms of small multiples, various studies have shown that when comparing animation and small multiples, both have strengths. Animation offers more accuracy, while small multiples emphasize speed [Archambault et al. 2011], and, on the whole, participants enjoy animated visualizations more [Robertson et al. 2008]. Though Tversky [Tversky et al. 2002] questions the usefulness of animation, the proposed system is a highly interactive visualization, which Tversky admits has strengths, rather than the direct animation in place of static graphics that the paper critiques.

There have also been various systems proposed for correlated, interactive small multiples [Liu et al. 2015; Keefe et al. 2009], which allow a user to explore specific small multiples at a more granular level. However, none of these proposed methods allow the user to simultaneously explore the data set as a whole, while also viewing small multiples of the data at regular intervals.

\*e-mail:kaos@stanford.edu

### 3 Methods

Interactive data visualizations generally focus on providing a large window where users can explore the full data set. Within these systems, users interact primarily by filtering along different dimensions, or directly manipulating the visualization to pan and zoom. While implementing the Dynamic Small Multiples system, we hoped to preserve this functionality while augmenting it with the strengths of a small multiples display.

#### 3.1 Use Case

The particular use case for showcasing this technique had to be highly cyclic data with interesting variations between cycles. For this reason, time series data was chosen, as time is inherently both extremely hierarchical and cyclic. Weather data was used as something with both seasonal and month-to-month variations.

One could picture an application where a user wanted to draw out trends in temperature over the last six summers, or compare corresponding summer and winter temperatures over several years of time.

#### 3.2 Design Iteration

The initial system was comprised of a large, interactively animated chart and a workspace for generating small multiples. Users were able to slide between predetermined intervals and generate multiples based on sub-intervals within each frame of the animation. This prototype was shown to several users, then altered based on feedback and observed usage.

##### 3.2.1 Data frames versus full data set

Comments from users and classmates indicated that the structure of the data frames was not showcasing the strength of the system, which should allow users to explore, free-form, an entire time series, and generate small multiples from this series. By generating predetermined frames, the application essentially predefined the scope of each multiple, limiting the interactivity of the user.

With that in mind, we removed the data frames entirely, replacing it with a large, interactive plot of the entire data set, allowing the user to navigate the data set without boundaries.

##### 3.2.2 Animation

The original intent of the project was to combine an animated visualization with a small multiples display, as seen in the idea of data frames. However, as described above, the frames were found to be limiting rather than useful.

The animation properties were, instead, moved into the small multiple display itself. By clicking on a small multiple, users are able to see a larger, more detailed version of the view. Users can then click between different multiples to manually animate this data.

##### 3.2.3 Snapshot multiples versus automatic generation

An additional suggestion was to let the system automatically generate small multiples for the user. Initially, this was not a central feature, but after several tests it was clear that users were most interested in the automatic generation, rather than the single-multiple snapshot generation, which creates one multiple from the current view. The snapshot feature was instead integrated into the automatic generation process, which allows users to generate a single multiple, or a range of them.

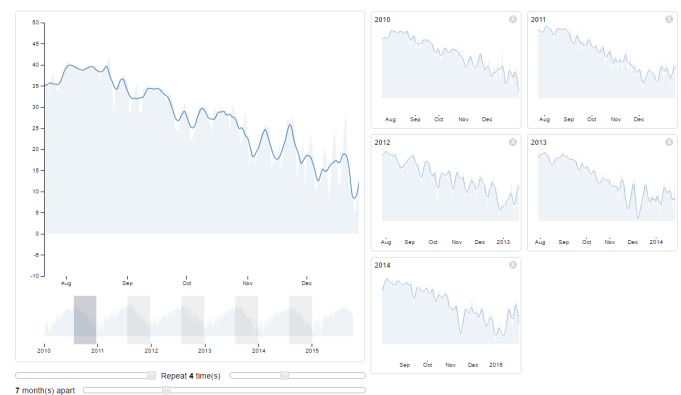
#### 3.2.4 Navigation controls

It was key that the system retain the ability to directly manipulate data that is offered by most interactive visualizations. These interactions are already familiar to the user, and provide an efficient way to explore details in a data set.

Initially, zoom scaled both the x- and the y-axis at the same rate. However, this meant that there was a significant amount of wasted space, as the data-rich portions of the graph often did not fit within the fixed aspect ratio of the frame at high zoom levels. After converting the system to an x-axis only zoom level, it became much easier to see and compare temperature levels between different frames.

#### 3.3 Approach

Dynamic Small Multiples first presents users with a large visualization of the data set and a context chart containing the full range of data, using the focus-plus-context interface, as seen in Figure 1. They can manipulate and explore this data set as they would on a typical interactive visualization, but they are also given the option to generate a set of small multiples based on input parameters. These small multiples may be rearranged or manipulated as the user continues to explore.



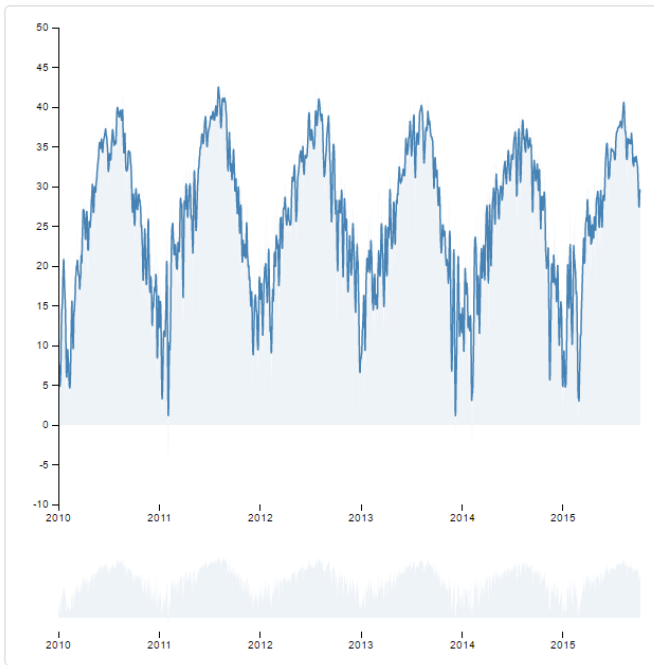
**Figure 2:** The left half of the page is composed of a focus and context visualization, with a large focus panel showing detailed information about the data set, and a narrow context panel displaying the entire data set, with the current view highlighted by a dark rectangle.

##### 3.3.1 Focus panel interactions

The user is initially presented with a zoomed-in view of one region of the data set (Figure 2). This region is also highlighted in the context chart below, to show the user which portion of the data they are currently viewing.

**Navigation** Within the focus panel, users navigate by dragging to pan left and right and scrolling with the middle mouse button to zoom. The range of the focus panel corresponds to the highlighted area in the context panel below.

**Hover** When a user hovers over the chart, a tooltip containing the y-axis data for that time point is displayed at the appropriate position on the chart. This gives the user both a high level of precision when looking at a close up view of the data, as well as a more general sense of data trends, by looking at the plotted line.



**Figure 3:** The focus panel has precisely numbered data, and can be zoomed or panned by direct manipulation. Below, the context panel is displayed.

**Adaptive Smoothing** The large size of the data set causes issues when viewing the entire time series. It becomes difficult to see trends in the noise of day-to-day fluctuations, which create a lot of meaningless jitter in the line when viewing the over a large time span. To counter this, the line chart dynamically smooths itself. At close range, it is a five-day rolling average between points, as the user zooms out, this rolling average slowly increases to a 20-day rolling average over the full five year span.

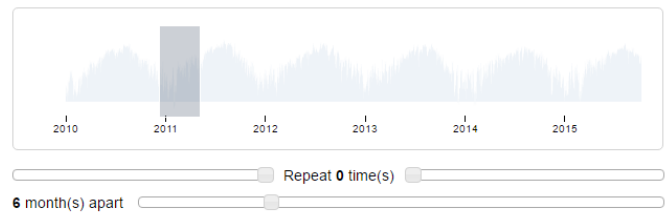
This adaptive smoothing makes it easier for users to see overarching trends in the data. At the same time, in order to preserve the accuracy of the data, the area fill is not adaptive. Rather, it is constrained to the precise values of the actual data, allowing the user to compare real data points using the hover-tooltip and area, as well as general trends, using the line.

### 3.3.2 Context panel interactions

The context panel consists of a high-level visualization of the full data set, providing the user with a context for the area they are currently viewing in the focus panel. The active area consists of the current domain of the focus panel, and is highlighted by a dark rectangle (Figure 3).

**Dragging** The active area can be directly manipulated by the user, allowing for intuitive navigation of the data set using the context panel. Users can drag to resize and drag to move, easily modifying their view with this interface.

**Small Multiple Preview** The sliders below the context panel allow users to specify the spacing and number of multiples that they wish to generate. As users adjust the sliders, live previews of the multiples are drawn onto the context menu to visualize the general range of each small multiple.



**Figure 4:** The context panel displays an overview of the entire data set. The range of data that is currently being viewed in the focus panel is marked by a dark rectangle overlay in the context graph. Below, there are two sets of sliders. The top slider allows users to specify the number of multiples to generate, while the bottom specifies the interval between multiples.

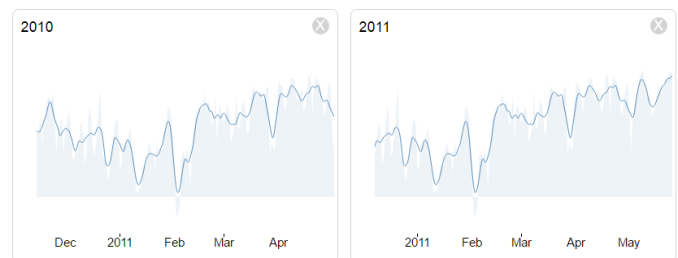
The context menu is meant to provide a very high-level view rather than specific detail, so the previews are not labeled with specific data. The goal of the application is rather to encourage users to generate small multiples to better inspect their data.

**Preview Manipulation** As the user resizes the active region in the context panel, the preview regions will also resize accordingly. The multiples also respond to zooming or panning in the focus panel.

Users may specify the spacing between each multiple using the slider. The granularity of this slider is set to months, a value that was decided experimentally based on the domain of this particular data set. In future iterations of this work, a more adaptive or intelligent means of specifying spacing could be explored.

### 3.3.3 Small multiple interactions

The goal of this application is to allow users to explore a data set by rapidly and efficiently generating small multiples. They can do so by clicking a button to generate an automatic range of small multiples, or, if no range is specified, a single multiple is generated from the segment of the graph currently in view.



**Figure 5:** Multiples are intelligently labelled. The label is the year of the starting date for the viewing region, unless that year takes up less than twenty percent of the graph area, in which case the label is the following year.

**Labelling and Sorting** When generated, multiples are automatically labelled with the year that takes up the majority of the multiple domain (Figure 4). They are also sorted chronologically by default, as a logical default axis for users to compare multiples across.

**Jump to Multiple** Once multiples have been generated, each indi-

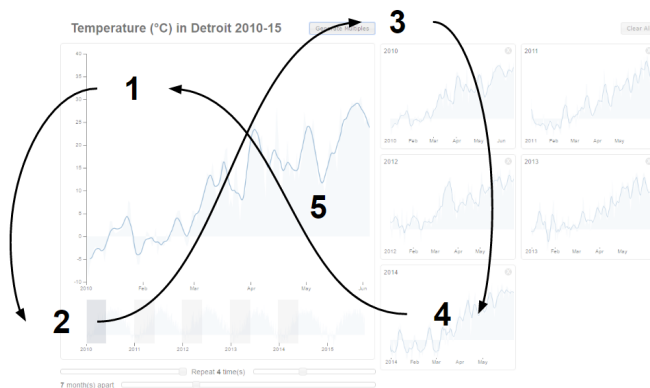
vidual window acts as a link to a portion of the focus graph. When the multiple is double-clicked, the focus panel will automatically switch to the view of this multiple, allowing users to examine the data within the multiple in more detail. Manipulation Users can drag to rearrange multiples, delete individual multiples, or *Clear All* to remove multiples from the workspace. This is a fast and straightforward way to rearrange and analyze the data from the small multiple visualizations.

### 3.4 Implementation

The system is implemented in JavaScript, using d3.js. While d3 is incredibly powerful and provides a rich set of existing features, it does limit the system to working within the specified bounds of the framework.

## 4 Results

### 4.1 Workflow



**Figure 6:** The user is encouraged to first discover interesting regions of data at a high level of detail focus panel. They can then explore similar regions by generating multiples across years. These multiples can be examined for trends. Finally, double-clicking on a multiple allows the user to return to the focus panel and continue data exploration.

The user workflow for this system is as follows (Figure 5):

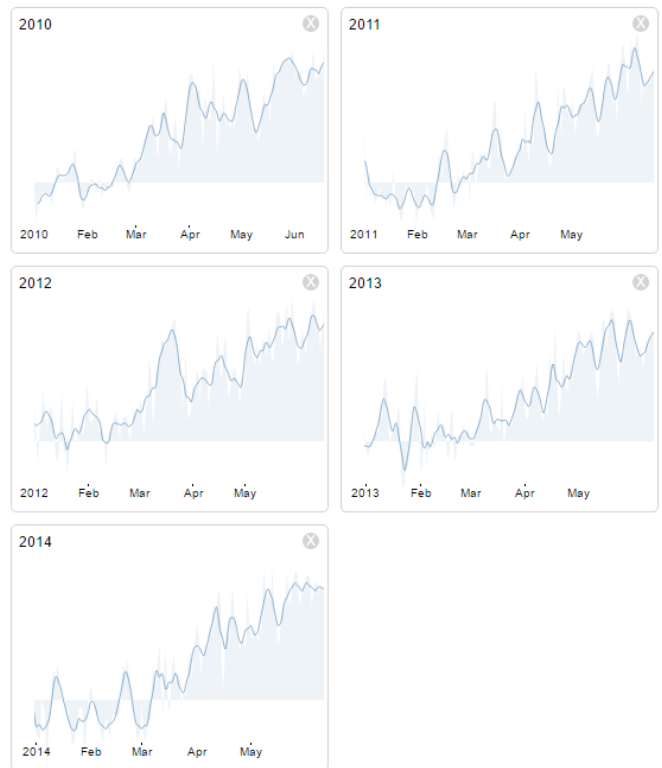
1. Users explore the data set in the focus window and find an interesting region
2. A range of multiples can be specified and previewed in the context window
3. Users then generate a series of multiples, and can rearrange or remove these multiples to better analyze the data
4. An individual multiple can be clicked to explore it in more detail in the focus window.
5. As users find further regions of interest in the focus panel, they can specify and generate more multiples to repeat the process

### 4.2 Usability

A series of informal usability studies were conducted by letting users unfamiliar with the system test and explore Dynamic Small Multiples.

When presented with Dynamic Small Multiples and Tableau and given the task of generating a series of small multiples, users generally found it more intuitive to use this system. This is an expected result, as the system is optimized for generating small multiples, and not nearly as feature-rich as Tableau, but does indicate that that the Dynamic Small Multiples system is fairly straight-forward to learn. On average, it only took users a few minutes to understand the system and begin exploring the data for interesting trends.

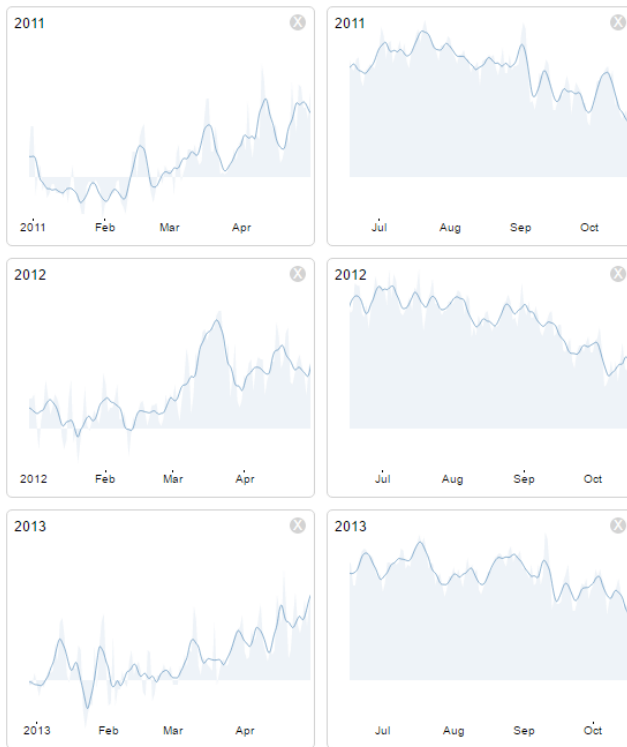
Users did wish for the ability to compare other data sets, or plot multiple data sets in one chart. Additionally, the ability to snap to certain time intervals (e.g. months) in the context panel would have been useful, as well as some preset lengths of time (day, week, month, year) to use when specifying intervals and multiple size. Some interesting trends are highlighted in Figure 6 and Figure 7 below.



**Figure 7:** Small multiples displaying the first six months of every year between 2010 and 2015. These charts plot the maximum temperature in Detroit. It's immediately obvious that the winter of 2012 was significantly warmer than any of the other years, and the winter of 2014 was, on average, colder than the rest. A user interested in exploring these trends further, could then double click on the 2012 multiple and examine the exact data points in the focus panel.

## 5 Discussion

To properly evaluate the system, a follow-up usability study could be to compare the usage of this Dynamic Small Multiples system with 1) a system that consists only of a large view of the full data set and 2) a system that consists primarily of small multiples that can be navigated or rescaled. We are particularly interested in whether the combination of a rapidly-generated, short term small multiples system and a large persistent view of the full data set allows users to more rapidly discover interesting trends in data, as compared to either of these visualizations separately.



**Figure 8:** Every winter from 2011 to 2013, plotted next to the corresponding summer. A user might generate this type of visualization to try and determine correlations between temperatures in the winter and temperatures in the summer. Within the dynamic small multiples system, it is fairly straightforward to specify an interval between two multiples, then generate appropriate multiples at this interval every year.

In terms of performance, users found the system workable. One drawback to small multiples is that generating many graphs very quickly involves a lot of processing power and a lot of memory usage. The d3 framework is somewhat limiting in that respect, and further optimizing the speed of the system is definitely desirable.

On the whole, users were interested in the idea of specifying intervals relative to a current view, as well as the concept of rapidly generating small multiples as a form of data exploration. A user study of some sort was definitely pinpointed as a crucial next step, in order to quantify if these methods were genuinely beneficial, or offered improvements in any specific areas (accuracy, speed, etc).

## 6 Future Work

### 6.1 Immediate Next Steps

There are several features that would immediately fit into the framework of this application and might provide new angles for users to look at data:

*Automatic sorting of multiples:* allowing users to sort multiples along various features. Currently, the sorting is only chronological, but would be useful to analyze the multiples and sort in an intelligent manner to extract trends. In the Detroit weather data set, for example, sorting by maximum temperature over the multiple, or maximum average temperature, could be useful.

*Automated animation:* though the animation system was removed from the final product, allowing users to select a region of multiples and generate an interactive graph consisting of the aforementioned data frames still seems worth exploring. Users would be able to slide forward and back in time within the region specified by their small multiples. The idea of dynamically generated animation is still intriguing.

*Linked annotation:* the ability to drop reference lines at particular x- or y-values, and have these lines be visible across all multiples would allow users to compare precise data values more easily

## 6.2 Future Directions

At a higher level, this visualization is a springboard for future systems that might integrate dynamically generated small multiples in different contexts. High-level questions to direct following applications include

*How would different data sets be integrated?* Weather from different cities, for example, or a plot of both maximum and minimum temperatures. This involves thinking about how to generate multiples in a multivariate system. Should users be allowed complete freedom or will the multiples be determined automatically? If automatic, how does the system decide when to compare trends across data sets, and when to plot several data sets within one multiple?

*How would one generate multiples along different dimensions?* The current system is limited to generating multiples along the x-axis, but allowing users to easily switch the axis or dimension they split the data along would be an extremely interesting feature. This relates to the question of integrating different data sets in that additional data sets are simply an additional dimension.

## 7 Conclusion

The Dynamic Small Multiples system is a fast, iterative framework to analyze data via small multiples, while simultaneously providing users with the ability to look at high-detail views of the data within a larger visualization of the full data set. This implemented system is a novel look at the potential uses of small multiples for the tasks of data exploration and visualization, and opens up several interesting channels for future work.

## References

- ARCHAMBAULT, D., PURCHASE, H., AND PINAUD, B. 2011. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Transactions on Visualization and Computer Graphics* 17, 4 (Apr.), 539–552.
- CHO, M., KIM, B., BAE, H. J., AND SEO, J. 2014. Stroscope: Multi-scale visualization of irregularly measured time-series data. *IEEE Transactions on Visualization and Computer Graphics* 20, 5 (May), 808–821.
- DOLEISCH, H., MAYER, M., GASSER, M., AND HAUSER, H. 2004. Case study: Visual analysis of complex, time-dependent simulation results of a diesel exhaust system. In *In Proc. of the 6th Joint IEEE TCVG - EUROGRAPHICS Symposium on Visualization (VisSym 2004)*, 91–96.
- KEEFE, D., EWERT, M., RIBARSKY, W., AND CHANG, R. 2009. Interactive coordinated multiple-view visualization of biomechanical motion data. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov.), 1383–1390.

- LIU, X., HU, Y., NORTH, S., AND SHEN, H.-W. 2015. Correlated multiples: Spatially coherent small multiples with constrained multi-dimensional scaling. *Computer Graphics Forum*, n/a–n/a.
- REEVES, G., LIU, J., NATH, S., AND ZHAO, F. 2009. Managing massive time series streams with multi-scale compressed trickles. *Proc. VLDB Endow.* 2, 1 (Aug.), 97–108.
- ROBERTSON, G., FERNANDEZ, R., FISHER, D., LEE, B., AND STASKO, J. 2008. Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (Nov.), 1325–1332.
- SWAYNE, D. F., LANG, D. T., BUJA, A., AND COOK, D. 2003. Ggobi: Evolving from xgobi into an extensible framework for interactive data visualization. *Comput. Stat. Data Anal.* 43, 4 (Aug.), 423–444.
- TVERSKY, B., MORRISON, J. B., AND BETRANCOURT, M. 2002. Animation: Can it facilitate? *Int. J. Hum.-Comput. Stud.* 57, 4 (Oct.), 247–262.
- WOODRING, J., AND SHEN, H. W. 2009. Multiscale time activity data exploration via temporal clustering visualization spreadsheet. *IEEE Transactions on Visualization and Computer Graphics* 15, 1 (Jan), 123–137.