Weatherbin: Visually Exploring Similar Days in Air Traffic Weather

Christopher Skeels, Kyle I. Murray, Kenneth Kuhn, and Akhil Shah; RAND Corporation; Santa Monica, California

Abstract

Interactive visualizations of complex datasets are an important tool for data exploration, but finding relationships between variables in highly multivariate environments often requires domain-tailored combinations of visualization techniques. We address this challenge in the domain of air traffic flow analysis with a tool designed to show how days with air traffic management initiatives are influenced by the weather on those days and to explore how various day cluster analyses may provide insight into relationships between the measured weather events and air traffic management. Our tool, called Weatherbin, provides both a broad overview of the day clusters as well as a detailed view of the weather conditions on any individual day, with interactive features to connect day details back to overall cluster averages.

Introduction

Data visualization is increasingly important in a world with more – and more varied – data sources becoming available. Well-studied areas have developed standard visualizations for important scenarios, but new combinations of data sources need to be integrated into analysis workflows with new tools and visual techniques. Visualizations of complex scenarios need to break down each scenario into comprehensible visual subcomponents, relying on people's existing knowledge of common visual techniques to let them synthesize an understanding of variable relationships in a broader and more complicated visual context.

Air traffic in the presence of adverse weather conditions is an especially complex and important scenario to understand, leading us to develop a novel, interactive tool for exploring clusters of multifaceted data. Our tool focuses on the sensemaking of our many-dimensional weather and air traffic datasets through coordinated multiple views of our weather data via parallel coordinates plots, calendar view, and weather radar animations.

Design

The design of Weatherbin is centered around supporting air traffic planners in exploring past weather patterns and associated traffic management initiatives used to cope with bad weather. Its novelty, in relation to existing tools, is the use of visual clustering algorithms to segment weather patterns into groups of similar characteristics and to further allow interactive drilldown, whereby cluster aggregated weather patterns can be explored in an iterative, user-directed fashion.

The Weatherbin interface is structured to give an initial broad overview of the selected dataset's clustering algorithm before letting users interactively explore relationships among the measured variables and clusters. Three primary columns separate the main feature areas of the visualization: 1) parallel coordinates plots of each of the measured variables separated into clusters, with the first showing overall average values and the second showing daily averages, 2) a set of year columns, each showing a year's calendar broken down into months and days showing how the clusters visually map onto each day, and 3) a day-detail column showing an interactive map of the day's 24 hour weather radar, the data measured for that day, and any air traffic management initiatives caused by weather on that day.

The Weatherbin interface supports switching between several kinds of datasets. These datasets may include different kinds or numbers of measured variables, as well as arbitrarily many-colored clusters. We validated Weatherbin with datasets deprived from typical traffic planning sources including Meteorological Aerodrome Report (METAR), Terminal Aerodrome Forecast (TAF), North American Regional Reanalysis (NARR), and Rapid Update Cycle/Rapid Refresh (RUC/RAP) data.

Cluster Average Exploration

The leftmost column focuses on two parallel coordinates plots that show the averages and range of each of the variables for each cluster. The top plot uses overall averages for the variables in each cluster while the bottom plot renders each day individually. Interactivity is key to using and understanding the relationships between the two plots. Hovering over one cluster line in the top plot will dim the rest of the lines representing the other clusters. This action results in the lower plot dimming all of the cluster lines except the ones representing individual days associated with the highlighted cluster. Lines belonging to the same cluster are color-coded for easy visual correlation as discussed in the next section.

Calendar Overview

Calendars are a familiar way to show the shape and progress of a year. We display a calendar where individual days have been color-coded and where days identified as similar by our clustering algorithm are painted the same color. Color coding increases our users' abilities to correlate features structured by the calendar framework with features exhibited by individual clusters. For example, because each day of the week always has the same horizontal position, a cluster which falls along a vertically-aligned set of days may suggest further investigation into a correlation between weekends and weather-related traffic management initiatives. Similarly, multi-month semi-homogenously-colored calendar segments can be associated with seasonal patterns that may be interesting depending on context.

Interactively, a primary feature of the calendar overview is to let users select a specific day to explore. Hovering over a day will de-emphasize days on the calendar overview not in the same cluster, will highlight only that day or cluster on the corresponding parallel coordinates plots, and will trigger the detail view in the right-hand column for that specific day. The cluster legend lets users refer to clusters by their identifying number, and hovering over a legend entry highlights elements of that cluster in all of the other columns.

Highlighted Day

Weather map radar animations are an accessible and intuitive way to summarize an important aspect of a day's weather: the

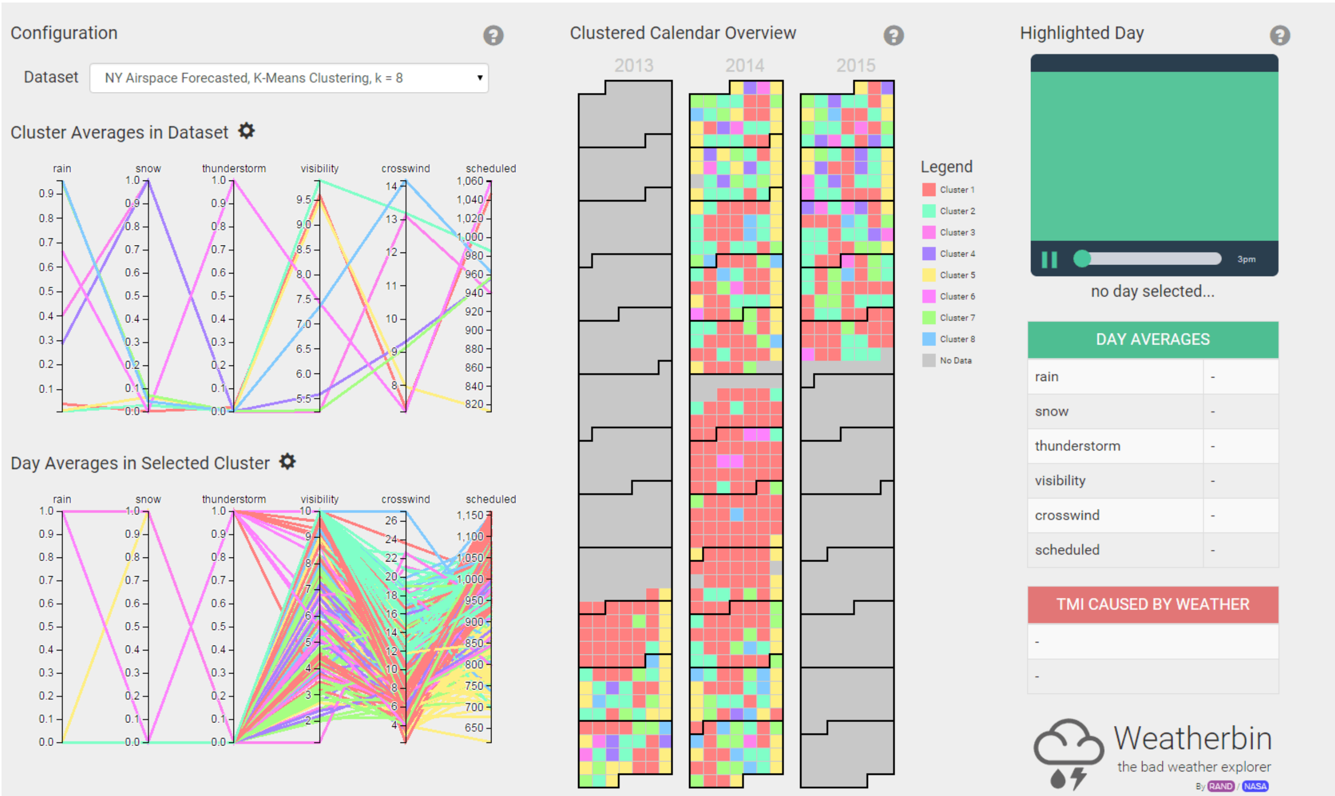

Figure 1. *The Weatherbin application showing the default overview perspective for the 2013-2015 New York Airspace dataset.*

presence of significant cloud coverage and precipitation. Along with a detailed readout of the specified day's measured variables, this section features a weather animation covering the entirety of the region of interest for the selected dataset. The animation can be paused or scrubbed backward and forward to let the user focus on specific portions of a given day's weather. The radar animation allows for visual confirmation that, for example, a given day's reported rain or snowfall measurements match the intuitive expectation generated by either an empty or cloud-filled radar image.

Implementation

We implemented Weatherbin as a web application built on top of D3 [1], Bootstrap, and jQuery. Weather radar maps were generated using weather data from the sources mentioned in section 2 and an R graphics script. Our novel contributions to the core D3 components include a new vertically-oriented, interactive calendar component and the unique data binding between the clusters and their ability to be highlighted simultaneously in multiple parts of the application at once. Weatherbin is designed to scale without forced horizontal cropping to any screen size, making it amenable for display on large-format displays or on small notebook displays.

Data Clustering

We apply two distinct clustering methodologies [2] in order to generate the results that our web-based application displays. The first involves k-means clustering using heuristic approaches to find good solutions to the computationally challenging problem of assigning all days to one of k clusters in such a way that the within-cluster sum of squares is minimized. This form of cluster analysis should be sufficient for many applications but does have a few disadvantages, such as requiring the data analyst to specify the number of clusters k and not being robust to outlying data points.

Future Work and Conclusion

Many-dimensional datasets are difficult to visualize succinctly and are often amenable to a coordinated multiple views approach to aid in the sensemaking of those datasets. We designed and implemented Weatherbin, and interactive visualization tool for identifying similar days in the airspace. Weatherbin uses an interconnected combination of parallel coordinates plots, a calendar grid interface for exploring individual days, and animated weather radar maps alongside a clustered dataset.

We plan to evaluate our tool among groups of both expert air traffic planners and non-experts to better understand the ease of comprehension of our tool and associated datasets as well as, in the non-expert case, to broaden the public understanding of the data science behind traffic management initiatives during everyday air travel.

We anticipate that Weatherbin will be beneficial to researchers developing more advanced visualizations in the future as well as serve as an aid to those developing algorithmic models for air traffic flow management initiative planning and analysis.

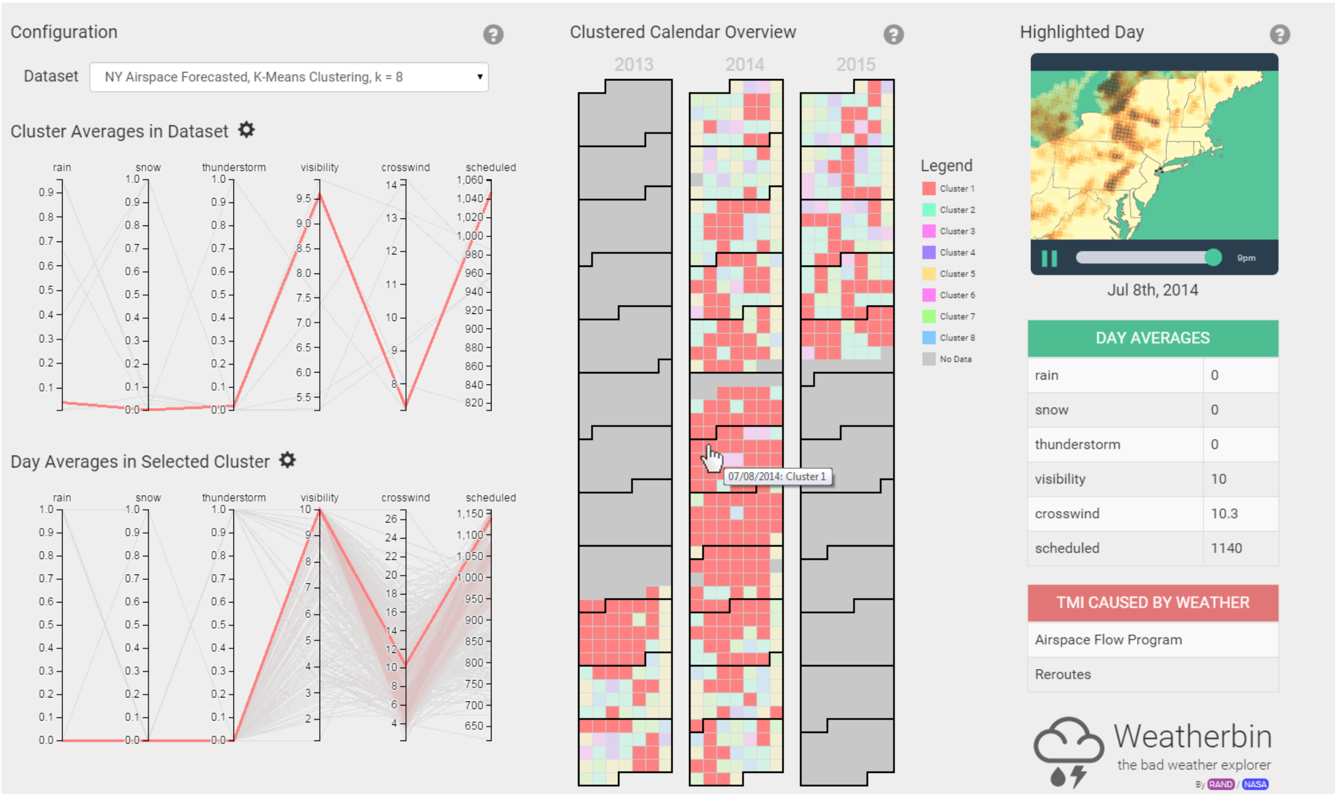

Figure 2. *The Weatherbin application after hovering over a specific day in the Clustered Calendar Overview section. The corresponding details for that day are displayed in the Highlighted Day section to right.*

References

[1] [1] M. Bostock, V. Ogievetsky, J. Heer. D3: Data-Driven Documents. *IEEE Trans. Visualization & Comp Graphics (Proc InfoVis)*. 2011.

[2] K. Kuhn. Characterizing and Classifying Historical Days Based on Aviation Weather and Air Traffic. *The 34th Digital Avionics System Conference*. 2015.

[3] A. Shah. Applying Causal Inference Models to Air Traffic Flow Management. *The 34th Digital Avionics System Conference*. 2015.

Author Biography

Christopher Skeels is Associate Director of Research Software Engineering at the RAND Corporation. His work has focused on the intersection of human computer interaction, machine learning, and modeling & simulation. He received an MS in Computer Science from the Georgia Institute of Technology (2005).