Cave Plots: A Graphical Technique for Comparing Time Series

Richard A. Becker, Linda A. Clark, Diane Lambert
AT&T Bell Laboratories Murray Hill, NJ 07974

Abstract

Cave plots are a good way to compare data from two or more time series, even if the two series are observed at irregular times. Cave plots can show thousands of measurements, but they are readily understood, preserve individual values, highlight short-term, long-term and periodic fluctuations, and make correlation (or lack of correlation) between two times series apparent. We have used a page of cave plots to show millions of measurements on network reliability and tens of thousands of measurements on air quality. The plots highlight features of the data that warrant a closer look.

1 Cave Plot Basics

Cave plots for two time series X and Y resemble the inside of a cave (see Figure 1). The values of the X series are plotted as stalagmites from the bottom up, and the values of the Y series are plotted as stalactites from the top down. The coordinate axes for the X series are shown by the left vertical axis and bottom horizontal time scale, while the coordinate axes for the Y series are shown by the right vertical axis and the top horizontal time scale. The X and Y horizontal axes are identical, but showing both makes it as easy to find time points for the Y series as it is to find time points for the X series. The X and Y vertical axes are different, though. The X vertical axis increases from bottom to top, as usual, but the Y vertical axis is *opposed* in the sense that it increases from top to bottom. The vertical axes also *overlap*, so that the X and Y series share part of the plotting region. In Figure 1, for example, the line $X = 50$ is identical to the line $Y = 100$. Because the extreme values from X and Y do not occur simultaneously, re-using the plotting region does not destroy information. It only puts the two series closer together, saving space and making them easier to compare visually.

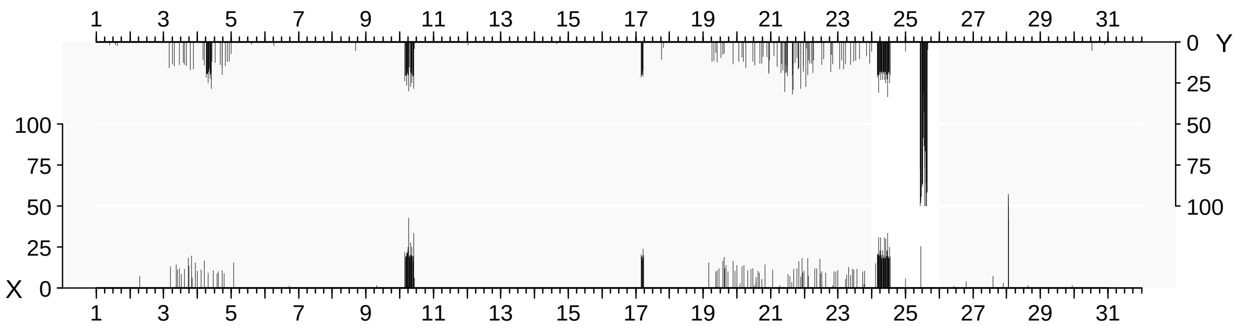

Figure 1: A typical cave plot showing network reliability data between two locations for one month. The horizontal axes show time (days). The vertical axes, which are opposed and overlapping, show percent of blocked calls for networks X and Y. Data values, which are 15 minute averages, are encoded by line segments starting from the horizontal axes. Notice the long intervals with no blocking, the symmetric blocking in both networks and the heavy blocking in network Y on day 25. The plotting region for days 24 and 25 is emphasized because of the sustained heavy blocking during that time.

Cave plots are related to three lane highways that reserve the center lane for turns by vehicles traveling in either direction. But collisions in cave plots are less serious than those on highways, so we accept infrequent collisions. Also, we can sometimes infer what happened in a cave plot collision by looking at the data before and after it. Making the plotting region do double duty and requiring one set of top and bottom margins rather than two also follows Tufte's heuristic (1983) of a high data density.

Cave plots do more than save space, though. A cave plot consists of two vertical line graphs, which can show thousands of individual values and highlight short-term, periodic and long-term fluctuations (Cleveland 1985). Line graphs also preserve raw data values and observation times, which is important for irregularly-spaced time series or time series with gaps. The vertical lines emphasize (add more ink to) large values, which is appropriate when large values correspond to important events. Finally, using a common horizontal axis for two series makes it easier to find similarities or lags in the series. But the real value of a graphical display is determined by its effectiveness for real sets of data. Section 2 presents two examples for large sets of data.

2 Two Examples For Large Data Sets

2.1 Network Data

We invented cave plots to visualize a database detailing hundreds of millions of observations on networks. Monitoring equipment reported the fraction of calls blocked between pairs of origination and destination switches roughly every fifteen minutes for more than a year for several networks. Each origination switch called several destinations and most destinations were called from more than one origination. We wanted to look at the data, but conventional plots took page after page to print.

Figure 2 shows how well cave plots summarized the data. In each cave plot in Figure 2, the fifteen minute blocking rates from a node to each of its destinations are shown as separate lines on the same plot for networks X and Y. For instance, when a node has a nonzero blocking rate to three destinations at times t_1, t_2, t_3, there are three line segments for the node at times t_1, t_2, t_3. This allows us to show the data for all destinations simultaneously. Positive blocking rates for nearly simultaneous intervals overlay each other, but the maximum blocking rate from a node is always visible at each time point. Since the larger the blocking rate the more serious the problem, this highlights the most important events. Plotting the blocking rate to each destination in a different color would preserve more information, of course. In all, about 2,250,000 blocking rates are shown in Figure 2.

Several features of the networks are obvious from Figure 2. The cave plots are not overly dense, suggesting that most of the time there is no blocking at any node on either network. Some of the stalactites and stalagmites are roughly symmetric for one node for several days (for example, Node F for days 19 through 24). This pattern hints at problems with a local origination switch that affected access to both networks equally. Symmetric stalactites and stalagmites that appear at more than one node (for example, Nodes D and E for days 5 through 6 and 29 through 31) suggest problems at a destination called by more than one node. Note that we can learn about the destinations even though the data are shown by origination. Finally, network Y has several periods of high blocking (see node A on day 14, node B on day 20, node C on day 11, and node H on day 25), while network X has persistent low blocking at node H.

The X and Y data shown in Figure 2 were usually not sampled at precisely the same times. Nev-

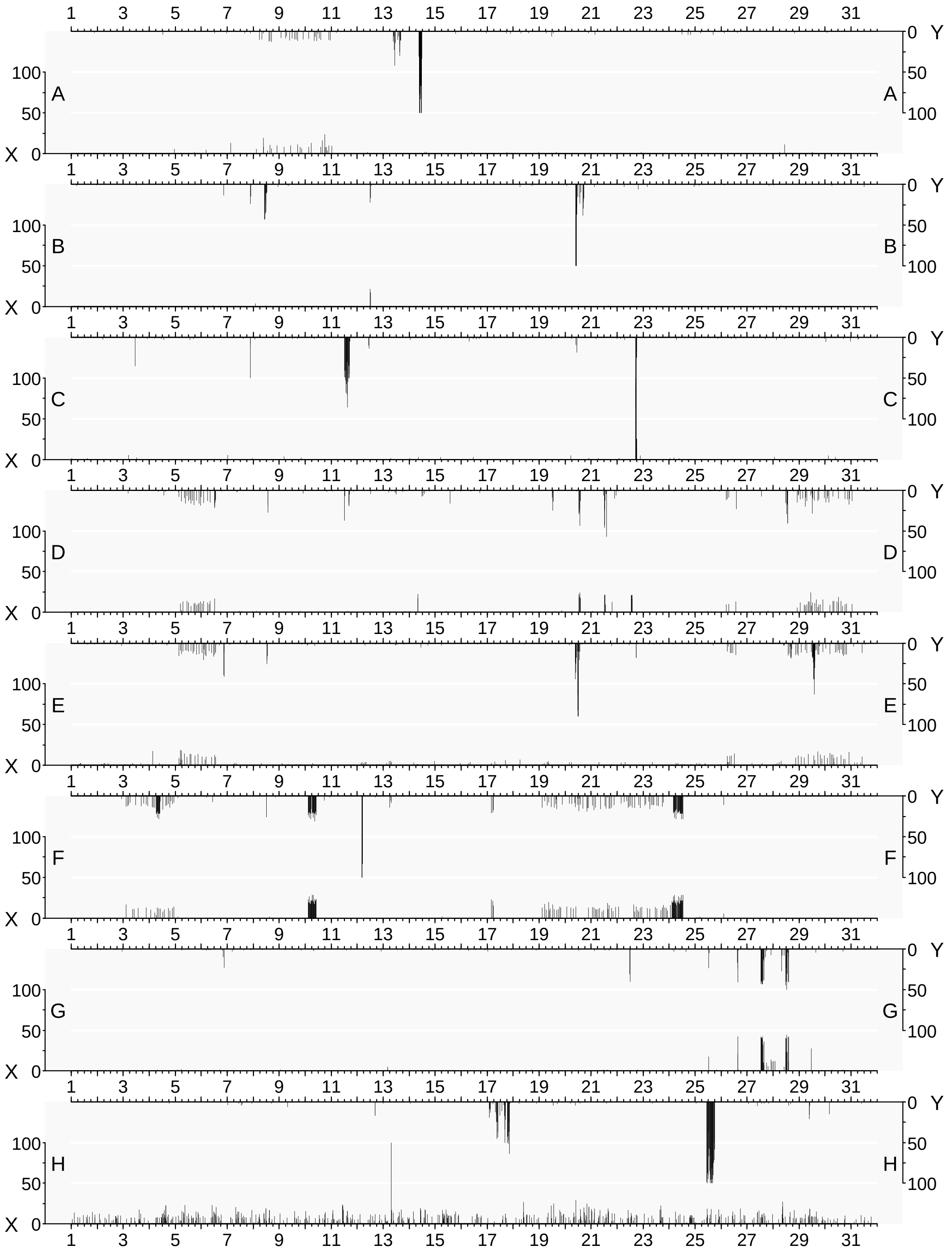

Figure 2: A cave plot comparing the fifteen minute blocking rates (in percents) for networks X and Y for one month for nodes labeled A–H.

ertheless, the cave plot allows the user to detect correlations between networks visually. Detecting those same correlations numerically would be more difficult.

Occasionally, there is a collision in Figure 2. Both networks experienced high blocking at node C on day 22, but the exact magnitude of the blocking cannot be determined from the plots. Nor is it possible to determine whether one network had more blocking than the other during the collision. The only conclusion that the viewer can draw is that the sum of the blocking on the two networks was at least 150%. (Disambiguation of the blocking rates would be possible if the two time series were plotted in different colors.) This is one of the few problems with cave plots, and it is a compromise between data density and resolution. If the vertical axes did not overlap, then ambiguity would be avoided. But then all the lines on the plots would be shorter and more subtle features would be obscured, or the plots would be fatter and more pages of plots would be needed. Of course, the time and origination node of the collision suggest which small subset of the raw data needs to be re-considered.

2.2 Clean Room Data

Ideally, a clean room has no particles in the air. To monitor air quality, air samplers report particle counts in factory air and in clean room air above work height, at work bench level, and near the floor every two minutes. Figure 3 shows the log_{10} counts for one clean room for four days. Here, “incoming” air refers to the air before it is cleaned and pumped into the clean room. (For more details, see Clark et al, 1992.) Counts are logged to allow low level trends to be seen even in the presence of extremely high counts.

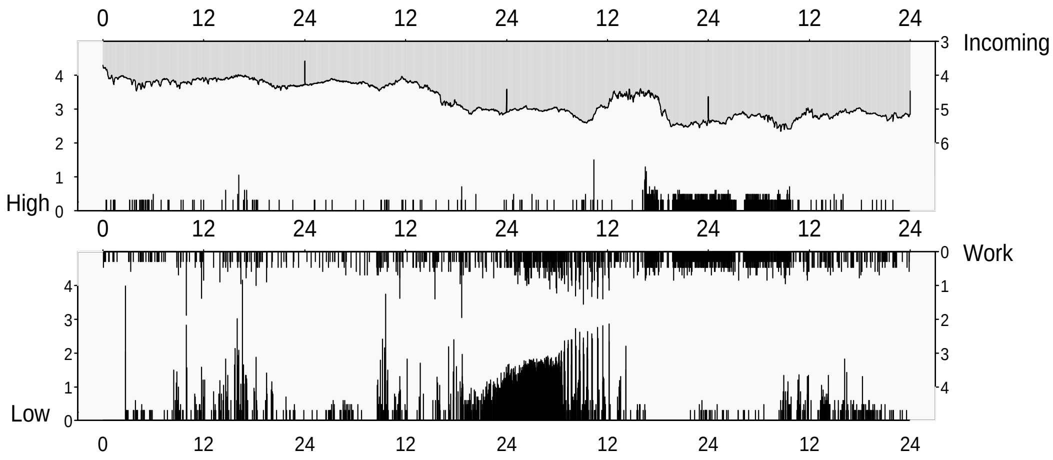

Figure 3: A cave plot comparing log_{10} air particle counts at four samplers for four days. Incoming air is monitored before being filtered and pumped into the clean room, high air is monitored above work height, workbench air is measured at work height, and low air is monitored near the floor.

Figure 3 shows the particle counts in the same sequence as the direction of the air flow: incoming air enters at the top of the clean room and moves down towards the clean room floor. Since there is little margin space between cave plots, it is possible to compare the counts at work and high levels, even though they do not appear in the same cave plot. Incoming air and clean room air are shown with different scales, because using the same scale would only emphasize that the incoming air is unfiltered. Similarly, incoming air counts are shown as a connected line with light shading rather than as a vertical line graph, because the line segments would form a thick black band and visually dominate the other data. Even though Figure 3 only has one vertical line graph instead of two, we still consider it a cave plot because of the opposed and overlapping axes. Here, the light shading lets us see that the air samplers report lower counts for a short period at midnight every day. Note that 20% of the interior of each cave plot is used for both vertical scales simultaneously. This avoids collisions, although there are a few near misses.

The incoming air counts seem to be unrelated to the clean room counts, which suggests that the air filters on the incoming air are effective. There is a disturbance in the counts for the low part of the clean room starting in the evening of the second day and continuing into the morning of the third day. The workbench measurements show a similar disturbance, but those at the high level do not. Because the incoming air counts are not higher during this period, the disturbance probably occurred in the cleanroom rather than in the incoming air. It is interesting to note that the low sampler begins to show its lowest counts when the disturbance ends, while the workbench and high air samplers show some of their highest counts then. We have been told by an air sampling expert that there were problems with the low air sampler during this period. Making such plots in real time could help identify problems that would otherwise go undetected.

Finally, it is not just the high counts that are interesting in Figure 3. Periodic patterns suggest problems that might be able to be diagnosed and corrected. For example, counts at the low sampler, and perhaps at the workbench sampler, too, seem to be highest between 8 am and 6 pm. This time–of–day effect is surprising, because the clean room is operated around the clock. Cave plots for longer periods would help to determine whether the pattern is real.

3 Further Considerations

We used the S language (Becker, Chambers and Wilks, 1988) to implement cave plots. S provided an environment for initial data manipulation, storage of the processed data, and analytical computations as well as graphics. Its flexible graphics routines enabled us to experiment with alternative graphical forms, finally arriving at cave plots.

Details in the implementation of cave plots are important. For example, we are careful to treat the two time series symmetrically whenever possible. The adjacent cave plots share a common time axis that labels the bottom axis of the top plot and the top axis of the bottom plot. All labels are horizontal and easy to read. In Figure 2, node names are symmetrically displayed at both ends of each plot.

The y-axis scaling on a cave plot is important. The maximum of each y-axis should be at least a little larger than the maximum data value to be plotted. If not, it is impossible to tell whether the maximum data value is associated with the top or bottom series. For our network data, where 100% blocking is achieved during periods of maximum stress, this means that the entire vertical scale should correspond to at least 105%. In Figures 1 and 2 the vertical scales reach 150%. Besides avoiding collision, this gives natural reference lines at 50% blocking, which some network engineers consider an important level of blocking.

To emphasize that each time series has a separate coordinate axes, we connect a vertical axis only to the horizontal axis from which the data values grow. This decision makes it easier to associate axes with sets of data, but it leads to a disconnected frame. Without the closed frame of an ordinary plot, it can be hard to tell which time series are paired in a page of cave plots. Lightly shading the interior of a cave plot helps us to pair the time series visually.

The order in which multiple cave plots are shown on a page is important. Alphabetical order is unlikely to be a good choice. Figure 2 shows alphabetical node labels, but the labels were attached after the plots were ordered by a multidimensional scaling. That is, we computed distances between origination nodes based on the number of destination nodes they had in common and used metric multidimensional scaling to select an appropriate linear ordering. This made it more likely that plots for originations with similar destinations would be placed near to one another so it would be easier to find problems at destination switches. In Figure 3, the cave plots are ordered according to the direction of air flow. Arranging cave plots by geography could be effective in other applications.

There is considerable flexibility in cave plot axes. Our figures show data in which the horizontal (time) axes are identical, but offsets could be used if needed. For example, if incoming air particle counts led clean room counts by several hours, the factory time scale axis could be offset by several hours to bring the series into alignment. The measurements in the time series can also be completely different, enabling comparisons between correlated variables.

How much information can be plotted using a cave plot? One measure is related to the size of a pixel. On a 300 dot-per-inch laser printer, an 8 inch wide plot occupies 2400 pixels and should be able to do something with about that many observations. When most data values are zero, many more than 2400 values can be displayed on one cave plot.

Finally, there are two obvious alternatives to cave plots: aligned time series and back–to–back time series. It would be more conventional to align two series, one above the other, so that they point in the same direction. But it would take more space and pairing the series correctly can be difficult when many aligned series are drawn on one page. Back–to–back time series plots are similar to the double dual histogram plots introduced by Dallal and Finseth (1977). These plots would not save space (except for margin space) since each time series would have its own, non-overlapping plotting region. Further, the horizontal axis labels would not be adjacent to the starts of the line segments, making it harder to find the times associated with events in Figures 2 and 3, for example.

Finally, more could be achieved with interactive cave plots. For example, more than two time series could be studied together if an interactive system allowed the series displayed on the top and bottom axes to be changed rapidly. The region of overlap could be chosen interactively if a slider controlled the amount of axis overlap. Color would also enhance cave plots. For example, in Figure 2 different colored backgrounds could distinguish high blocking caused by natural disasters such as earthquakes from high blocking caused by fiber cuts. Plot linking might enable highlighting on one plot to be controlled by events found on another plot.

4 Summary

Cave plots are easy to produce, readily understood, and capable of displaying large sets of data compactly. The key new idea is to display two time series in the same plotting region with overlapping and opposing vertical axes. Cave plots work best when the highest values of the two series seldom occur together, so collisions are rare. For example, cave plots work well when the data are nearly always zero but during rare "events" the data are higher. This describes failures of reliable systems, air particle counts in clean rooms, seismic activity, and unusual weather events, for example. These kinds of applications often generate massive amounts of monitoring data, and cave plots can be an important tool for exploring its structure.

ACKNOWLEDGEMENTS

We thank Carol Rigler and Camille Sobers for their help with the network data and Doug Sinclair for his help with the clean room data.

REFERENCES

Becker, Richard A., Chambers, John M., and Wilks, Allan R. (1988), *The New S Language*, Chapman and Hall, New York.

Clark, L. A., Hastie, T., Psota-Kelty, L. A., Sinclair, J. D. and Rauchut, J. (1992), "Sources of Particle Contamination in an IC Manufacturing Environment," *Aerosol Science and Technology*, 16, 43–50.

Cleveland, William S. (1985), *The Elements of Graphing Data*, Wadsworth, Belmont CA.

Dallal, G. and Finseth, K. (1977), "Double Dual Histograms," *The American Statistician*, 31, 39–41.

Tufte, Edward R. (1983), *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT.