



COMPLEX DATA MADE CLEAR

MARINE INFORMATION OFTEN CONTAINS CONSIDERABLE INFORMATION GATHERED OVER LONG PERIODS OF TIME. VISUALISING THAT DATA USING TRADITIONAL TECHNIQUES CAN BE HARD. BLAKELEE MILLS LOOKS AT A TECHNIQUE THAT MAKES IT EASY FOR BOTH EXPERTS AND LAYPEOPLE ALIKE TO UNDERSTAND THE INFORMATION AVAILABLE

Analysing the many variables in a marine environment can be a challenge. Tidal patterns affect maritime traffic, coastal communities and aquatic breeding periods; water temperatures affect algae growth, which in turn affects marine animals and the fishing industry; the list goes on.

To further complicate matters, an important aspect of acquiring marine data is time. Typically, data is gathered over time, as opposed to spatially, which results in large, multi-variate datasets that contain a wealth of information. However, the information is not very useful if it cannot be properly visualised. Traditional 2D and 3D visualisation techniques such as line/scatter or bar plots are adequate for basic analysis but fall short when one needs to analyse both large and small data patterns.

The 2D line graph in Figure 1 is the daily passage of adult Chinook salmon in Bonneville in the US. The X-axis represents the day of the year and the Y-axis represents the total number of salmon passing the gauge station. Each line plot represents a different year.

Major patterns, such as the peak passages, are easy to distinguish, but smaller patterns such as seasonal variability are difficult to identify. Additionally, the graph represents only five years of data. What if five years isn't enough? One can see how messy the graph would quickly become when adding another 10, 20, or 50 years of plots.

An alternative visualisation technique is to plot this time-series data as an image map, also known as a raster map (see Figure 2). The date is turned into two temporal coordinates where the X-axis is a short-term time step such as days or minutes, and the Y-axis is a

longer time step such as months or years. The Z-value is represented in the cell location.

Using this time map technique, the Chinook passage data can be plotted in Golden Software's contouring, gridding and 3D surface mapping software, Surfer (see Figure 3). In this map, an additional 71 years of data is displayed. Despite the extra information, it is quite easy to distinguish temporal patterns, both large and small, that would otherwise be missed in the 2D line graph. The smaller spring and summer run patterns are now easier to see in this display. Also notable is the steady increase of salmon in recent years during the August through September run.

Ocean data

Time maps are also useful when displaying ocean tide data. The Hawk Inlet on Admiralty Island in southeast Alaska, US, is one of two ports in the Alaska Panhandle where the arrival and departure of ships depends on the tides. The timing of tides and daylight hours are critical factors for safe navigation; therefore, it is important for schedulers to know when the tides are favourable for vessel travel.

In these traditional views (see Figure 4), high and low tides are distinguishable. But these displays are insufficient for planning when ships can safely navigate in and out of the port that day, week or month. A significant amount of information is lost and obtaining a clear understanding of tidal patterns is practically impossible with these views.

However, by using this time-map visualisation technique, a whole new picture appears (see Figure 5). Schedulers can quickly

pinpoint days and times when tide levels are conducive for ship travel. This view also makes it easier to add extra information, such as the time of sunrise and sunset, which is represented by light grey lines. Despite having over half a million data points, this plot is clear and understandable. Not only do port schedulers benefit from tidal information, so do coastal managers preparing for nuisance flooding, recreational fishermen, surfers and boaters, and scientists working on habitat restoration projects, to name a few.

Algae blooms

Multi-variate temporal analysis is another area in which time maps provide a clearer and more understandable visual representation of complex data. For example, several environmental factors must be considered when analysing algae blooms. Such factors include sea surface temperature, sea surface salinity, air temperature, precipitation, stream flows, tidal height difference, upwelling, and wind speed. Individually, these factors do not have much of an impact on whether or not an algae bloom will occur. Instead, the factors must be analysed as a whole which can be a troublesome task.

How can one recognise the windows of opportunity for these algae blooms? The first step is to display each environmental factor as a time map, seven different environmental time maps being created in all. The second step is to apply a binary filter to the data that highlights days with favourable conditions for an algae bloom. A value of one equals days with favourable conditions and a value of zero equals days that are unfavourable for an algae bloom. The binary filter is also displayed as a time map.

For example, one environmental factor is stream flow (see Figure 6). Flows less than 350m³/s are favourable for an algae bloom. By viewing the binary image map, one can quickly pinpoint when stream flows are less than 350m³/s. These typically

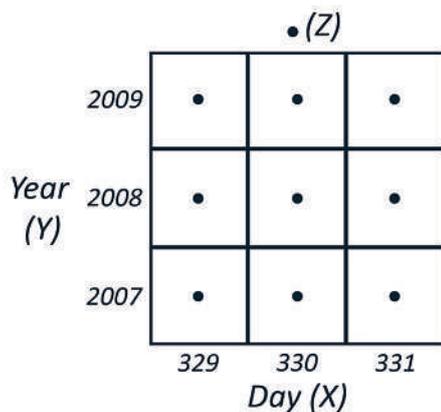


Figure 2. An example of the grid format used to generate time maps. Source: <http://arizona.openrepository.com/arizona/handle/10150/280516>

Adult Passage / Adult Chinook Bonneville, Chinook 2010-2014

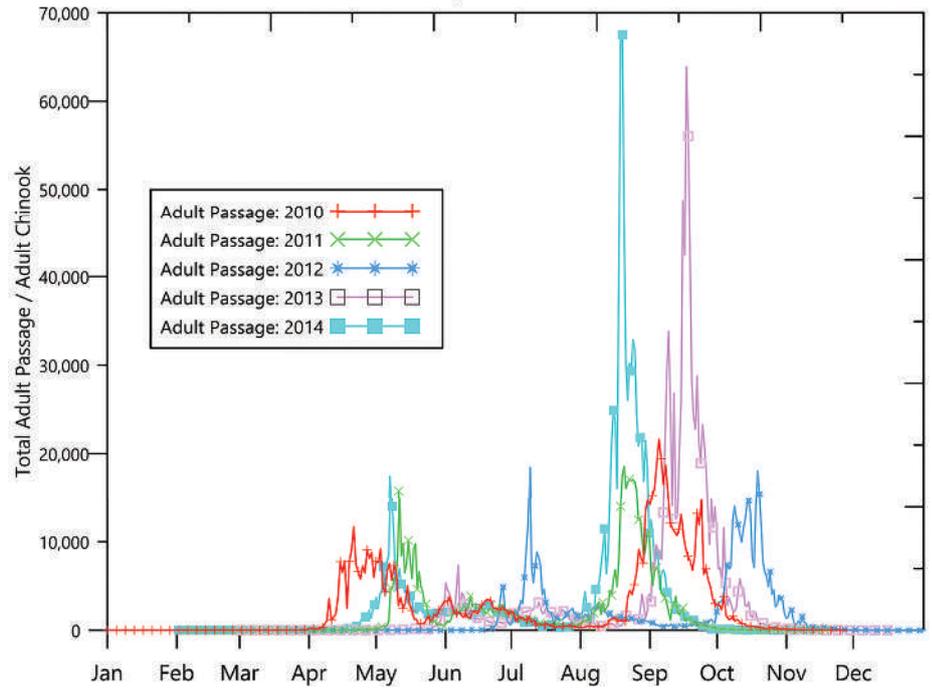


Figure 1. The migration of Chinook salmon is typically displayed as a series of line graphs. Major patterns such as peak passage times are easy to discern, but smaller patterns or seasonal variability is hard to identify. Source: www.fpc.org. Graph plotted in Golden Software's 2D and 3D graphing software, Grapher

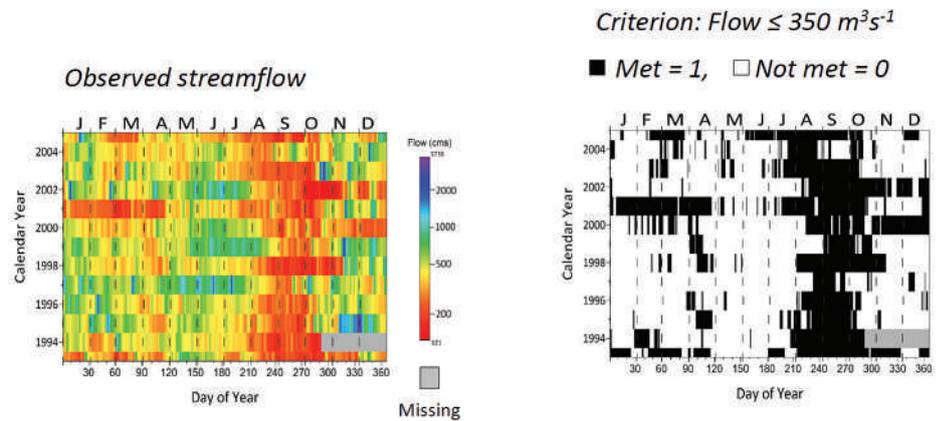


Figure 6. The time map on the left is the observed stream flow. The time map on the right is the binary filter highlighting days when stream flows are less than 350 cubic meters/second in black. Maps plotted in Surfer

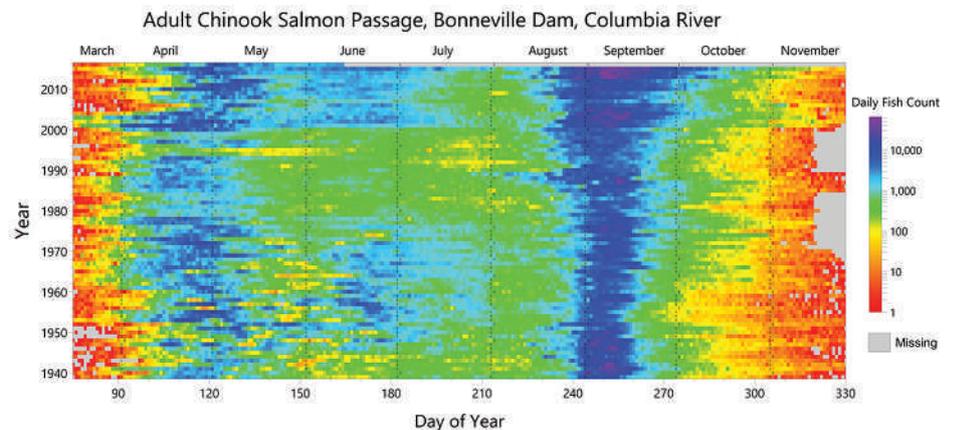


Figure 3. The spring run can be seen in late April, and the summer run is in June with an obvious low count from years 1975 to 2000. The fall run is late August through September with notable increases in recent years. Data source: www.nwd-wc.usace.army.mil. Map plotted in Surfer

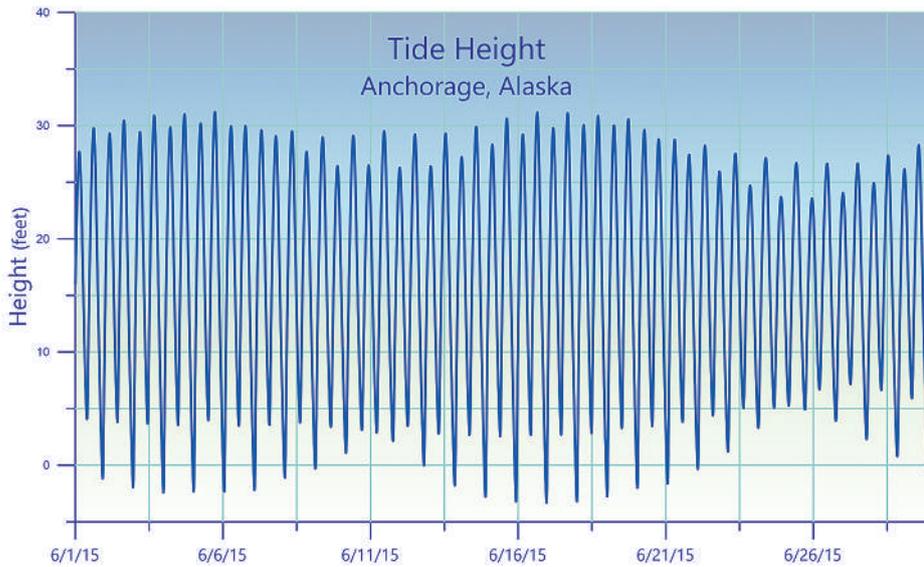


Figure 4. 2D line chart plotting tide heights over time. Large patterns are distinguishable in traditional plots, but smaller patterns are difficult to decipher. Graph plotted in Grapher

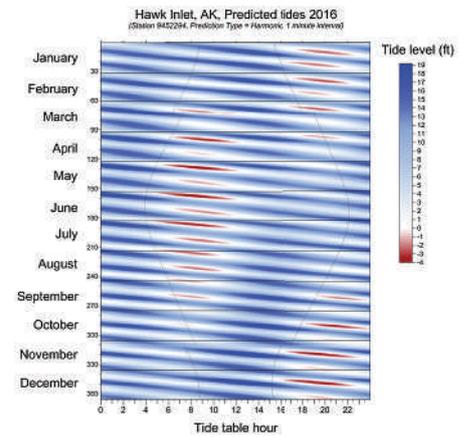


Figure 5. The timing of the tides, such as the lowest tides shown in red, vary throughout the year due to the ever-changing positions of the Earth, Moon, and Sun. Represented by the grey line is the daily sunrise and sunset times which is important for safe vessel movement. Map plotted in Surfer.

	A	B	C	D
1	Day(X)	Year(Y)	Value(Z)	Date
2	121	2015	1023	5/1/2015
3	122	2015	1104	5/2/2015
	D2-Date(Year(D2),1,1)+1	Year(D2)		
	D3-Date(Year(D3),1,1)+1	Year(D3)		

Figure 8. Convert your data to a format where you have three columns of data (X, Y and Z). The X column is the day of the calendar year, the Y column is the calendar year, and the Z column is the data variable

occur between August and October. One can also see a drought occurred in 2001, resulting in less than normal stream flows for the months of January through April. All days represented in black are favourable stream flow conditions for an algae bloom.

You then repeat these steps for all remaining environmental factors.

The final step is summing the binary data for the individual factors and displaying the resulting data as a time map. The summed data will contain values between zero and eight. Cells containing zero indicate on that particular day, none of the eight factors had conditions suitable for an algae bloom. Cells with a value of four indicate on that particular day, four out of the eight environmental factors had conditions suitable for an algae bloom. Cells with a value of eight indicate all eight environmental factors had conditions suitable for an algae bloom on that particular day. When displaying this combined data as a binary time map (see

Figure 7), values of eight means all criteria was met and values zero through seven mean at least one criterion was not met.

Of the 4,380 days between 1993 and 2005, there were 127 days where all eight environmental factors were favourable for an algae bloom.

The key to creating a time map is properly converting the typical date format, 5/9/2016, into two temporal coordinates, the day and year, within the input file. Using an Excel spreadsheet, Figure 8 shows the formulae using a calendar year. The X-axis is day and the Y-axis is year. The Z-value would be the time series such as salmon count, tidal height, or stream flow. Arrange your input data as column A = day (X), column B = year (Y), column C = value (Z).

Once the data is in the correct format, the time map can be created in Surfer by 'gridding' (interpolating) the data and creating an image map from the gridded data.

Analyses can be greatly improved

by incorporating data visualisation. Depending on the need, traditional 2D and 3D views are satisfactory; however, by creating new ways to visualise both univariate and multi-variate data, unique and actionable insights will often be found.

The purpose of any graphic is to turn data into information. A well-crafted plot can be incredibly useful in understanding data and properly conveying your point. Having the right tool is critical to achieving this goal.

A WELL-CRAFTED PLOT CAN BE INCREDIBLY USEFUL IN UNDERSTANDING DATA AND PROPERLY CONVEYING YOUR POINT

Blakelee Mills is the CEO of Golden Software (www.goldensoftware.com)

Leica
Geosystems

Be Captivated

Visit leica-geosystems.com/becaptivated
to find out more and request a demonstration.



Leica Viva GS16 Experience 3D innovation

Meet the self-learning GNSS, powered by RTKplus and SmartLink, seamlessly connecting with the Leica Captivate touch-technology software. While RTKplus automatically selects the optimal GNSS signals with a robust 555-channel engine, SmartLink uses precise point positioning technology to stay connected. Experience 3D innovation with the most accurate positions and complete the job from anywhere.



Leica Geosystems AG
leica-geosystems.com



© Copyright 2016 Leica Geosystems. All rights reserved.

Visit us at **INTERGEO 2016**

Hamburg, Germany 11- 13 OCTOBER Hall A3 Booth E3.021

- when it has to be **right**

Leica
Geosystems

PART OF
HEXAGON