White Paper

Environmental Data Visualization:
Turning Environmental Data Into Big Insights

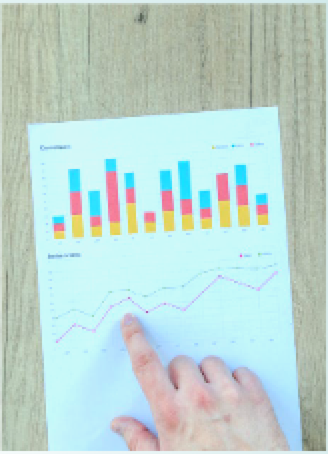
The Rise of MultiDimensional Visualization Tools for Data Discovery

Prepared by

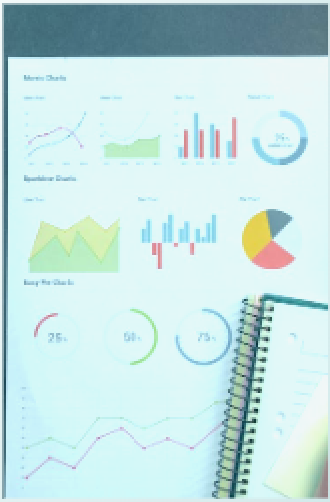
SustainableTech

Table of Contents

The Trend towards Environmental Data Visualization Tools



Studies on climate change are becoming more and more important today. In order to determine the effects of climate change on people and the environment, an ever-growing massive amount of environmental data is measured, stored, and examined. Online environmental data is made available by several organizations to support global climate change research. The National Oceanic and Atmospheric Administration (NOAA), the Western Regional Climate Centre (WRCC), Cal-Adapt, and the Nevada Climate Change Portal (NCCP) are a few of these institutions. Even though the majority of climate-related data is readily accessible on the Internet, it can be difficult and time-consuming to study very big datasets without the aid of visualization.



The need for visualization is expanding in a variety of fields, including business, engineering, education, physical sciences, social sciences, medical sciences, meteorology, hydrology, economics, and genetics, as graphical depictions of data convey concepts more effectively. Finding hidden patterns in raw data more quickly and effectively is made possible through visualization. It aids the researchers in narrowing their focus and moving forward more swiftly. There are numerous industry-standard visualization tools available right now, but each one was created with a particular objective in mind. A researcher faces a difficult decision when selecting a tool from the extensive array of tools accessible. The survey described in this research was carried out as visualization technologies were selected to create a visualization system for the NCCP.

The Struggle to Make Meaning Out of Environmental Data

There aren't many thorough studies on visualization tools for environmental data. But a survey quite similar to ours was done for biological network analysis visualization tools [5]. This study assessed several visualization programs, including Medusa, Cytoscape, BioLayout Express3D, ProViz, and Pajek. These tools' visualization abilities, effectiveness, interoperability with other tools, supported input data formats, user-friendliness, and application were evaluated. For each difficulty listed, such as the integration of heterogeneous data, scalability, interconnection, and pattern detection, the authors offer tools as a summary of the survey. Furthermore, although they are noted but not examined in the study, a number of the more specialized tools for visualizing biological networks are used. Additionally, the discussion and evaluation of common network file formats for biological network representation, including BioPAX, SBML, CellML, and RDF.

The report also highlights several difficulties encountered by biological network visualization techniques. For instance, one of these difficulties is usability when hundreds of nodes need to be shown (because most tools become less efficient and slower when operating with large datasets). Leveraging online services or altering the current techniques to outsource the computational workload is advised.

Representing heterogeneous data is another difficulty that visualization tools must overcome. Aigner et al. [6] conducted a comparison of information visualization technologies. The features and functions offered by the tools were categorized, and a matrix listing the features offered by each tool was presented.

"Information limitations are severely constraining our ability to identify and understand emerging environmental problems, devise interventions to address them, and evaluate whether our responses work. [3]

Tools like Tableau, Xmdv Tool, Spotfire, and ILOG Discovery were evaluated and compared, and basic user tasks involving one or two variables, and advanced tasks dealing with complex problems were tested. The surveyed visualization tools might benefit from embracing common input file formats and handling huge datasets, among other things.

Mozzafari and Seffah [7] provided a brief overview of the environmental data visualization tools that are now available, including IBM Data Explorer, OceanShare, ImmerseDesk, CAVE, and Infinity Wall. Although less user-friendly and unable to link multivariate data, these techniques were shown to be effective for integrating heterogeneous data. The toolset that the paper primarily introduces provides interaction and visualization features. The services include employing visuals to gain insight into isolated data, connecting visualized data using interaction approaches, and identifying hidden patterns and correlations in massive datasets.

Aigner et al. [8] conducted a survey of 101 visualization approaches that are widely employed for portraying time and time-oriented data. Data, time, mapping, and dimensionality were used to group the techniques into different categories. A table displaying the features supported by the approach was provided based on this categorization. According to the survey, none of the visualizations could handle data with cyclic time progression and various points of view. Even when the problem is identical to the original one, the majority of techniques cannot be used for another visualization issue. Therefore, a broader application is required for visualization techniques. The results of previous studies make clear that visualization tools and methodologies face numerous difficulties. We look at a number of current visualization tools in the section after this.

Environmental data gaps are obvious when such data exist, but often they are ad hoc.
Data on biodiversity, hydrology, productivity, and erosion are insufficient and disjointed because of limited expertise, and a lack of sustainable funding for long-term monitoring and coordination. [4]

Visualization Tools for Environmental Data

For analysis, a few of the current tools for visualizing environmental data were chosen. These tools are typically employed for specialized tasks. We offer a comprehensive assessment of each tool's features, usefulness, strengths, and weaknesses.

ArcGIS

ArcGIS is a software package that allows users to quickly create maps by using geographic knowledge [9]. It is a piece of exclusive software made available by the Environmental Systems Research Institute (ESRI), a pioneer in geographic information systems (GIS).

Features

ArcGIS is available in many versions. End users can utilize their desktop PCs, mobile phones, tablets, and the web to access maps, data, and applications. Using the ArcGIS API for JavaScript, Flex, Silverlight, and SharePoint, web developers can create apps that run on a variety of platforms. The online mapping aspect of the tool enables the sharing of information, maps, and apps created by other people. Predefined templates in the desktop edition of ArcGIS automate the production of maps, which are therefore freely shareable. Devices running Windows Mobile, smartphones, and tablets may all use the ArcGIS runtime SDK. Users can access a ready-to-use software environment with a wealth of resources through the ArcGIS online service.

Strength

- ArcGIS is easy to use.
- High quality maps can be created quickly by using automation for desktop users and no installation is needed for ArcGIS online users.
- Several input data formats are supported.

Drawbacks

- ArcGIS is expensive and there are additional costs for upgrades.
- Clip geoprocessing tool is slow and often produces inaccurate results [10].

AVS/Express

AVS/Express supports object-oriented development and is mainly used for visualization purposes by programmers and non-programmers [11].

Features

AVS/Express offers a number of visualization approaches for 2D and 3D settings and is available for a wide range of platforms. Additionally, it is possible to see multi-gigabyte datasets. Rapid application development is made possible by the tool's cross-platform compatibility and cross-platform GUI interface. AVS/Express provides a library with 900 modules for carrying out different data management and visualization activities. The tool's parallel processing features enable distributed computing via shared memory, producing extremely quick processing. Displaying visualizations across many displays is possible because of the tool's multi-channel output.

Strength

- Easily integrates modules from programming languages such as C, C++, and FORTRAN.
- Scales to very complex and large datasets.

Drawbacks

- It depends on virtual memory for sending results to users, which at times may result in low performance.
- As it supports multiple platforms, the number of available standard widgets and controls is relatively limited.

GrADS

Grid Analysis and Display System (GrADS) is a visualization tool used for data manipulation and visualization of Earth science data in a 5-dimensional space [12].

Features

GrADS offers a variety of powerful built-in functions and executes operations using command-line expressions that resemble FORTRAN. Furthermore, output graphics can be saved as pictures or PostScript, and users can add external functions written in any programming language.

Strength

- This toolset is open source.
- It supports several input data formats.

Drawbacks

- Only a desktop version is available.
- The users need to learn commands to draw and control the graphics, which could make the tool less intuitive.

Integrated Data Viewer (IDV)

IDV is a Java-based software framework for analyzing and visualizing geosciences data [13].

Features

IDV is available on multiple platforms. The IDV framework needs Java and Java3D to function. Apart from geoscience applications, its user interface can be tailored to carry out particular tasks. IDV can show multiple data types simultaneously and gives labels for the longitude and latitude axes.

Strength

- It is free.
- Provides high quality 3D visualizations.
- It can plot data from remote servers.
- Supports several data types.

Drawbacks

- It requires a lot of RAM, which can make it slow for large databases.

UV-CDAT

Ultrascale Visualization Climate Data Analysis Tools is a set of tools that support data analysis and visualization for large climate datasets [14].

Features

This software toolkit is based on Python and may be combined with additional tools like CDAT, VisTrails, and Paraview to enable researchers to create sophisticated scientific visualisations. This programme, which works across various platforms, offers many functions for climatic data, including mean, standard deviation, and linear regression. Additionally, it offers provenance features and supports VisTrails for data analysis.

Strength

- Provides a suitable, useful suite of tools for climate related data visualizations.
- It is open source.

Drawbacks

- Supports a limited number of operating systems.
- As it is a new product, it may still have bugs and need more testing.

VisTrails

VisTrails is an open-source system that provides support for scientific data workflow and visualization [15].

Features

The provenance feature of VisTrails allows the comparison of different workflows visually. The results of the workflows can be displayed on a broad range of displays, from large display walls to small iPod screens. VisTrails is a python-based tool to which existing code or new modules can be added. Specialized plugins allow for other tools such as VisIt and ParaView to be integrated. Among its educational purposes, VisTrails is used for teaching Scientific Visualization and Digital Media.

Strength

• Has a simple and easy to use user interface. • Benefits from a broad user community. • Contains good comparative visualization capabilities.

Drawbacks

• Sometimes it may hang up while updating large amounts of data from a remote site. • Has limited parallel computing capabilities.

VisIt

VisIt is an open-source tool that provides capabilities for the visualization of complex scientific data [16].

Features

VisIt can run on different platforms and supports C++, Python, and Java interfaces. Rich 2D and 3D visualizations of terascale data are achievable because of parallel and distributed architecture. By dynamically permitting plugins, extensibility is attained.

Strength

• Provides a good framework for customization.
• Has capabilities for interactive parallel visualizations.
• Accepts input data of different formats.
• Supports terascale data sets.

Drawbacks

Data movement could be challenging in future machines.

Visualization Toolkit (VTK)

VTK is open-source software that supports an object-oriented environment and consists of libraries written in C++. It is mainly used for image processing and visualization [17].

Features

VTK provides surface and volume rendering support. Through parallel processing, it makes massive, complex datasets visualisable. Additionally, it provides a large selection of 3D widgets and supports several platforms. The majority of the wrapper code used to bind Python, Java, and Tcl is automated.

Strength

- Manages and represent complex scientific data.
- Supports many visualization techniques.
- Has a large user base.

Drawbacks

Contains a limited number of modelling tools.

Discussion

20 visualization tools were surveyed, and it became clear that users had to transition between them in order to provide climate researchers with the outcomes they were looking for. Even while the majority of the data is easily and freely accessible to researchers over the Internet, managing huge streams of data can be challenging when the data is in various forms. Additionally, only a small number of visualization tools have their source code publicly available, allowing visualization researchers to swiftly implement their ideas without having to start from scratch.

Since high-quality computer displays are readily available, 3D/4D approaches are currently the main focus of visualization. Further research is required to determine the advantages of these methods, which take the place of the conventional 1D/2D methods that have been in use for many years. Additionally, the complexity of user engagement with visualizations is rising. For example, users can change the visualization by using functions like zoom in, zoom out, moving left/right/up/down, or selecting multiple views on a map - Google, for example, offers all these features in its visualizations. Users can also discover many details available in the visualization by rolling the mouse over the visualization.

What's next?

Creating applications that run on a variety of devices, such as desktop computers, mobile phones, display walls, and touch pads; support diverse operating systems, such as Windows, Mac, and UNIX; offer a variety of visualization options and capabilities so that users do not need additional tools; support various interaction and visualization techniques, and offer high-quality graphics with realistic shading are currently among the challenges facing data visualization designers.

Checking for missing data values from faulty equipment and the accuracy of the received data are both vital since the created visualizations play a significant part in decision-making. Additionally, the visualization tools would benefit from a common input format. We plan to create an all-encompassing visualization system for the NCCP in the future that has the aforementioned features.

Acknowledgement

This work was made possible through the support provided by the National Science Foundation under Cooperative Agreement No. EPS-0814372 and No. EPS-0919123.

End notes

[1] Likhitha Ravi, Q., & Sergiu M. Dascalu, F. (n.d.). A Survey of Visualization Techniques and Tools for Environmental Data.

[2] (MARCH 2013). Big Data Visualization: Turning Big Data Into Big Insights. Intel IT center.

[3] ANNE S. MARSH, CHRISTINE NEGRA, & ROBIN O'MALLEY. (2022, July 14). Closing the environmental data gap. Issues in Science and Technology. Retrieved November 19, 2022, from https://issues.org/omalley/

[4] Environmental data. Environmental Data - an overview | ScienceDirect Topics. (n.d.). Retrieved November 19, 2022, from https://www.sciencedirect.com/topics/earth-and-planetary-sciences/environmental-data#:~:text=Environmental%20data%20gaps%20are%20obvious,long%2Dterm%20monitoring%20and%20coordination

[5] Pavlopoulos G. A., Wegener A., and Schneider R., "A survey of visualization tools for biological network analysis", BioData Mining, November 2008.

[6] Aigner W., Bertone A., and Miksch S., "Comparing Information Visualization Tools Focusing on the Temporal Dimensions," 12th International Conference on Information Visualisation, pp. 69 - 74, July 2008.

[7] Mozzafari E. and Seffah A., "From Visualization to Visual Mining: Application to Environmental Data", IEEE Conference on Advances in Computer-Human Interaction, pp. 143-148, February 2008.

[8] Aigner W., Miksch S., Schumann H., and Tominski C., Visualization of Time-Oriented Data, Springer, May 2011.

[9] "ArcGIS - Mapping and Spatial Analysis for Understanding Our World", ESRI, Accessed December 3, 2012, available at: http://www.esri.com/software/arcgis

[10] "ArcGIS QGIS Faceoff", blog.donmeltz.com, Accessed: December 3, 2012, available at: http://donmeltz.com/blog/index.php/2011/06/10/arcgisqgis-faceoff/

[11] "AVS/Express Data Visualization Software", AVS/Express, accessed December 4, 2012, http://www.avs.com/products/avs-express/index.html.

[12] "GrADS Home Page", Grid Analysis and Display System, Accessed December 4, 2012, http://www.iges.org/grads/grads.html

[13] "Unidata | IDV", Unidata, Accessed December 5, 2012, available at. http://www.unidata.ucar.edu/software/idv/

[14] "UV-CDAT", UV-CDAT, Accessed December 5, 2012, accessed. http://uv-cdat.llnl.gov/

[15] "VisTrailsWiki", VisTrailsWiki, Accessed December 3, 2012, available at: http://www.vistrails.org/index.php/Main_Page#VisTrails_Overview

[16] "VisIt Visualization Tool", visIt, Accessed December 7, 2012, available at. https://wci.llnl.gov/codes/visit/home.html

[17] "VTK - The Visualization Toolkit", Visualization Toolkit, Accessed December 7, 2012, available at http://www.vtk.org/