

Visual Analytics of Paleoceanographic Conditions

Roberto Theron*

Departamento de Informática y Automática
Universidad de Salamanca

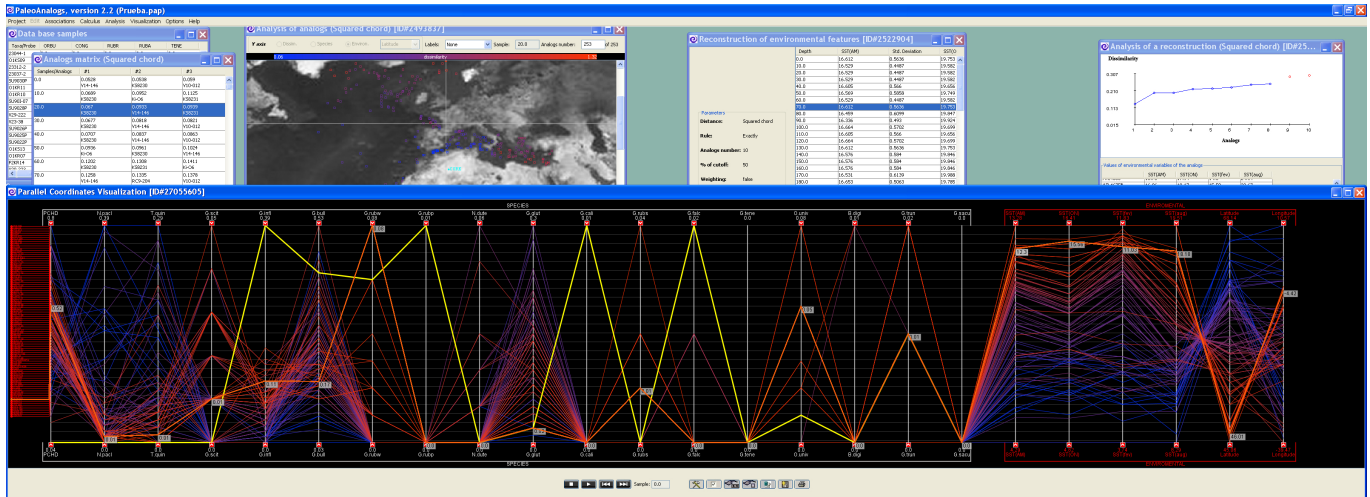


Figure 1: The PaleoAnalogs Visual Analytics Environment

ABSTRACT

Decade scale oceanic phenomena like El Niño are correlated with weather anomalies all over the globe. Only by understanding the events that produced the climatic conditions in the past it will be possible to forecast abrupt climate changes and prevent disastrous consequences for human beings and their environment. Paleoceanography research is a collaborative effort that requires the analysis of paleo time-series, obtained from a number of independent techniques and instruments, produced by several researchers and/or laboratories. Analytical reasoning techniques that combine the judgment of paleoceanographers with automated reasoning techniques are needed to gain deep insights about these phenomena. This paper presents an interactive visual analysis environment for paleoceanography that enables the discovery of unexpected relationships and supports the reconstruction of climatic conditions of the past.

CR Categories: H.1.2 [User/Machine Systems]: Human information processing—Visual Analytics; J.2 [Physical Sciences and Engineering]: Earth and atmospheric sciences—Applications

Keywords: infovis, parallel coordinates, multiple linked views, exploratory analysis

1 INTRODUCTION

In the last few years, the world has suffered from some of the most catastrophic natural disasters in recent history. While some of them

are of geologic origin, such as the Sumatra-Andaman earthquake (2004), that triggered the single worst tsunami in history, most of them are related to the weather. Examples of the later case are Hurricane Katrina (2005), one of the costliest and deadliest hurricanes in American history; or the big El Niño (El Niño-Southern Oscillation, ENSO) in 1997-98, that cost hundreds of lives and caused \$34bn in damage worldwide, partly through flooding to Chile, Ecuador and Bolivia and partly through failing harvests in Australia, the Philippines and Indonesia. A more recent, milder one in 2002-03 caused the worst Australian drought in a century¹.

Other phenomena, less understood than ENSO, include the Arctic Oscillation (AO), the Pacific Decadal Oscillation (PDO), and the North Atlantic Oscillation (NAO). Oceanic features like sea surface temperature (SST) variations associated with these phenomena can significantly impact local, regional, and global climate. Furthermore, fossil evidence has demonstrated that Earth's climate can change within a decade, and those new established patterns can persist for decades or centuries.

While the need to foresee abrupt climatic changes is an urgent challenge for the society, paleoclimate research has shown that the causes and effects of these changes are very different, with extremely rapid variations even on one-year basis. Computers have played a key role in our understanding of the climatic dynamics. Nowadays, the improvement of data acquisition methods offer us the opportunity to gain the needed depth of information to diagnose and prevent any natural disaster. By means of an analysis of such data, paleoceanographers are expected to assess (understand the past) and forecast (estimate the future). Although massive amounts of data are available, the development of new tools and new methodologies is necessary to help the expert extract the relevant information. This is the approach of Visual Analytics [21] as

*e-mail: theron@usal.es

¹After Ivan. *Prepare for Return of el Niño*, by Geoffrey Lean, Environment Editor, The Independent - Sept 12, 2004

the science of analytical reasoning supported by highly interactive visual interfaces.

If very high precision physical or chemical measurements are necessary to reconstruct paleoenvironments, they often need to be accompanied by sophisticated statistical analysis methods ([23], [22]). But, to be useful, these mathematical or software tools must not remain only in the hands of specialists in statistics, but must also be usable by the larger community of paleoclimatologists. It is therefore necessary to foster an optimal use of these mathematical tools, by establishing methodological choices among the most relevant and the most recent statistical methods, and to conceive a user interface adapted to the specificities of their use in paleoclimatology.

The data registered over thousands of years (mainly in ice and sediment cores) is an impressive source of information that, for instance, help us to model earth and oceans dynamics [18], first step to make climatic predictions. When looking for historic climatic data with durations exceeding decades, the largest and oldest record is found in the oceans. Palaeoceanographers need to manipulate, integrate and analyze time-series that are obtained from a number of independent techniques (such as ocean drilling, ocean tracers, AMC 14C datings, astronomic curves, etc.), which, moreover, are usually produced by different researchers and/or laboratories. This work is done with the aid of proper tools such as PaleoPlot [19] and AnalySeries [12].

Some of these data needed to understand paleoclimate are time-series of specific attributes related to the oceans. Thus, one problem scientists must face is how to know environmental parameters, such as sea surface temperature (SST), at each given past moment. For the reconstruction of this features, isotope measurements ($\delta^{18}\text{O}$) or biomarkers (U_{57}^k index) have been used. On the other hand, for the quantitative reconstruction of environmental conditions of the past, currently the *Modern Analog Technique* (MAT, actually a nearest neighbor prediction) [5], is one of the most commonly used techniques in paleoclimatology.

Although software tools for MAT have been developed [15], and some improvements have arisen such as SIMMAX [13], RAM [22] and artificial neural networks [10], they all have a main drawback: once developed they are black boxes. Paleoclimatologists can use them but no knowledge acquisition is involved; they just trust in the reconstructions obtained, they cannot know if the data used is valid from a geologic point of view. Furthermore, the classic MAT method inherently produces reconstructions whose precision is very difficult to estimate [11].

Visualization provides insight through images and can be considered as a collection of application of specific mappings from the problem domain to a visual range [7]. These visual representations combined with interaction techniques that take advantage of the human eye's broad bandwidth pathway to the mind allow experts to see, explore, and understand large amounts of information at once [21]. Thus, this paper presents an interactive visual analysis environment for paleoceanography that, through the combination of techniques coming from statistics, information theory, information visualization and visual data mining, enables the discovery of unexpected relationships and supports the reconstruction of climatic conditions of the past.

The rest of this text is organised as follows: in the second section it is explained how raw sediment core data is transformed into modern analogs (MAT), more appropriate for the analytical task of paleoclimatologists. Section 3 is devoted to explain how visual representations and interaction technologies enable knowledge discovery and provide a means of analysis of spatial and temporal data, the way to facilitate expert-driven accurate reconstructions. To finalize, the main conclusions and future work are described.

2 MODERN ANALOGS: DATA TRANSFORMATION

The National Visualization and Analytics Center, in its Research and Development Agenda for Visual Analytics [21] highlighted the key aspect of data representation and transformation as a way of support visualization and analysis.

This section describes how PaleoAnalog, a Java based program, makes use of the modern analog technique (MAT) [5] in order to provide faster and more accurate reconstructions of climatic conditions of the past. Initially, PaleoAnalog used this k-nearest neighbor prediction as a stable machine learning classification method [2], that provided an automated tool for reconstruction of paleoenvironmental features such as sea surface temperatures (SST). Being MAT a black-box method, where paleoclimatologists just collected the results of a mathematical algorithm that sometimes were not valid from a geologic point of view, a timid advance was introduced [20]: a better understanding of the reconstruction process by means of interactive visual analysis (colored scatter plots) of the analogs (neighbors) found by the algorithm. However, spatial and temporal dimensions were not considered in the interactive analysis.

This has been accomplished, by means of a combination of new visual representations and interaction techniques, for these transformed data. Both will be explained in section 3. Before that, how this transformation works is explained.

In k-nearest neighbor prediction, a database is used to predict the value of a variable of interest for each member of a *target* data set. The structure of the data is that there is a variable of interest (SST, for instance), and a number of additional predictor variables (e.g., microfossil species abundances). The k-nearest neighbor algorithm can be summarized as follows: a) for each case in the target data set, locate the k closest members (the k nearest neighbors) of the database. A distance measure is used to calculate how close each member of the training set is to the target row that is being examined. b) estimate the unknown variable of interest for that particular case as average of the variables of interest for its k nearest neighbors.

At this point, in paleoclimatology there is problem with the data coming from different sources, that prevent the automated transformation and the subsequent analysis. It is assumed that the user has faunal census estimates of one or more fossil samples, the core file; and one or more sets of faunal data from modern samples with the related environmental features, the database file. Furthermore, the user must understand the taxonomic categories represented in the data sets, and be able to recognize taxa that are or may be considered equivalent in the analysis.

With PaleoAnalog, the process begins after the selection of the core and database files; in general, these files will contain different taxa (figure 2.a and figure 2.b), both because different taxa are prevalent in different regions and because data providers use varying taxonomic categories (species and subspecies), names, and abbreviations. MAT requires that corresponding variables in different data sets be recognizable as such, otherwise it would be impossible to calculate the distance measures. With the help of the taxa association wizard (figure 2.c) this problem is easily worked out, allowing the user to determine which taxa from both the modern and fossil data files are compared, calculate proportions if needed, and identify the environmental features to be reconstructed.

Once the database and the core data are transformed to have the same number and equivalent taxa (figure 2.d and figure 2.f), each sample in the core is compared with each sample in the database using a dissimilarity coefficient.

Finally, using the distance measure selected by the user, a dissimilarity matrix is built. For each core sample N dissimilarity values are given, being N the number of samples in the modern database; these values are ordered increasingly so that each row of the matrix contains, left-to-right, the list of the N best analogs, that is, the

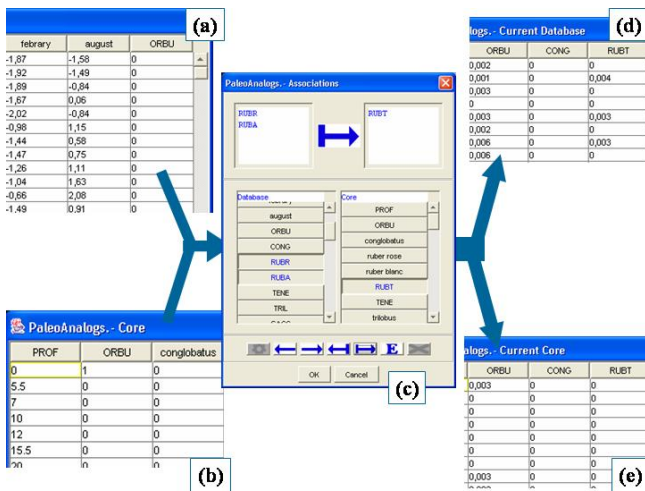


Figure 2: Automated data transformation with the taxa association wizard

database samples ordered by their likeness to that particular core pattern.

3 VISUAL ANALYTICS: VISUAL REPRESENTATIONS AND INTERACTION

This section explains how the use of proper interactive visual representations foster an analytical discourse (a dialogue between the analysts and the information) [21]. Using the original and transformed data, it is possible to automatically find patterns in information, and represent such information in ways that are meant to be revealing to the analyst. On the other hand, by interacting with these representations, using their expert knowledge, it is possible to refine and organize the information more appropriately. This way, it is possible, not only to reconstruct paleoenvironmental features, such as SST or salinity, but to visualize what information is being used to estimate these variables, and help the paleoclimatologists to decide upon using particular data or not, according to their field experience.

Following, a collection of visual representations and interaction techniques that enable the analytical reasoning in paleoceanography are described.

3.1 Improving reconstructions with visual insight

The next step in the classical MAT is to reconstruct the environmental conditions of each core sample based on the environmental data of a number of best analogs (generally ten). This can be done by calculating the average value or by weighting the analogs. However, this is somehow very strict, because some of the used analogs could not be valid from a geologic point of view and should be eliminated.

It must be noted that the classical MAT technique only permitted to select the number K of analogs (neighbors), normally 10, that are used to calculate the averaged variables. But consider the case where only 3 of these 10 modern analogs were actually similar to the sample being reconstructed, while the 7 remainders were only the following most similar; therefore, it would be a mistake to use them for the reconstruction from a geological point of view.

Figure 3 shows the PaleoAnalogs interactive reconstruction tool. It is possible to see which sites have been the closest analogs (neighbors) for each sample in the core. For the example in the figure,

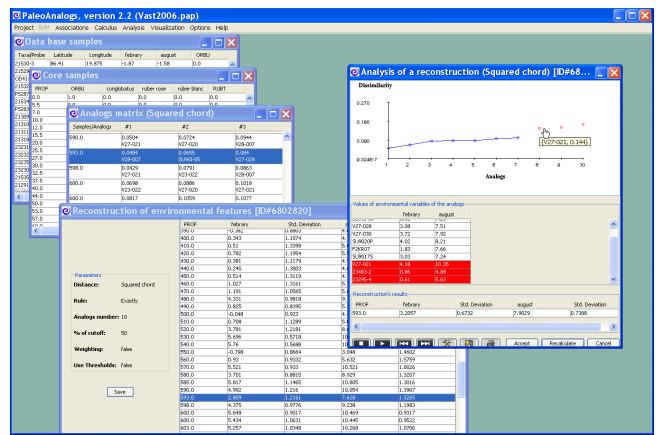


Figure 3: Reconstruction tuning

centimeter 593.0 of the core (which is a particular year in the past, depending on its sedimentation rate) is being studied. In this case the user has considered that the difference between the dissimilarity values of the seventh and eighth analogs is not acceptable; interactively, by clicking on each point of the plot, the analog is deselected (red circles), and the average recalculated using only the remaining selected analogs. On the other hand, the information of each analog is exposed, e.g. seventh analog is site V27-021, so the paleoclimatologist may decide upon rejecting any particular analog using geological criteria that may suggest not to use the data from that site. Note that after the recalculation, the temperatures of february and august have changed from 2.809 and 7.618, respectively (blue highlighted row in the *Reconstruction of environmental features window*), when the 10 best analogs were used, to 3.2057 (february) and 7.0029 (august), using only 7 analogs. Also, standard deviations are reduced 50%, approximately.

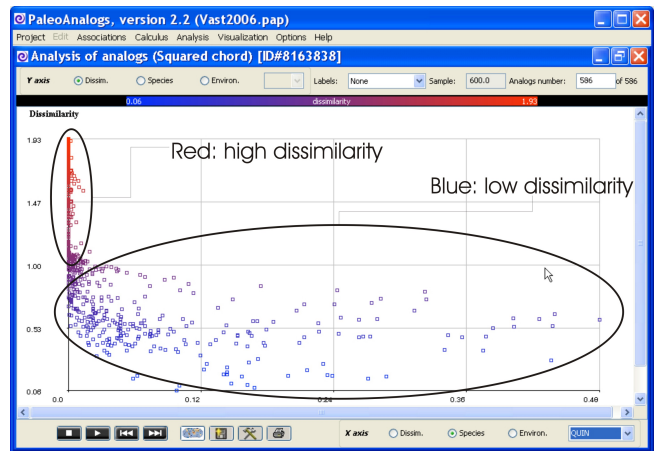


Figure 4: QUIN Species as an indicator of analog dissimilarity

Another and very valuable example of knowledge discovery for paleoclimatology, (figure 4) would be that a particular species (QUIN, for instance) is valid as an analog indicator, since the dissimilarity for the sample is very high when the proportion of this species is close to zero (all analogs in the x axis are red and have a high dissimilarity value, while the rest of analogs are getting more and more blue, i.e., more similar, as the proportion of the species

grows). This kind of analysis can be done for any pair of variables, such as any fossil species or environmental variables (latitude, longitude, and so forth). The dissimilarity is always present in these scatter plots as a third, color coded dimension.

3.2 Reasoning about space and time

Since going down in the core means going backwards in time, another important step that has been added is to provide a way to leverage the ability of paleoclimatologists of reasoning about time. On the other hand, all the analysis is done for geospatial information. Thus, both space and time are considered in the visual representations, in order to gain insight about the core at hand.

Actually, time is a variable that can be considered in the previous examples. Both the interactive reconstruction tool and the scatter plots can be animated. This way, the experts can analyze the whole history of the core. In the first case, it is possible to go down in the core, and visualize the dissimilarity plots of the analogs used for the reconstruction of each sample (depth/age) of the core. This way, visually, a different dissimilarity pattern can be easily discovered. In the second case, the relationship between QUIN species, for instance, and dissimilarity through time can be inspected, and learn if this species behaves in the same way through the geological history of that particular site.

Although the MAT method is very useful for paleoclimatic reconstruction there is much more information that can be provided than a mere neighbor distance calculation. Thus, before proceeding with the algorithmic reconstruction further knowledge can be easily discovered from the calculated set of analogs.

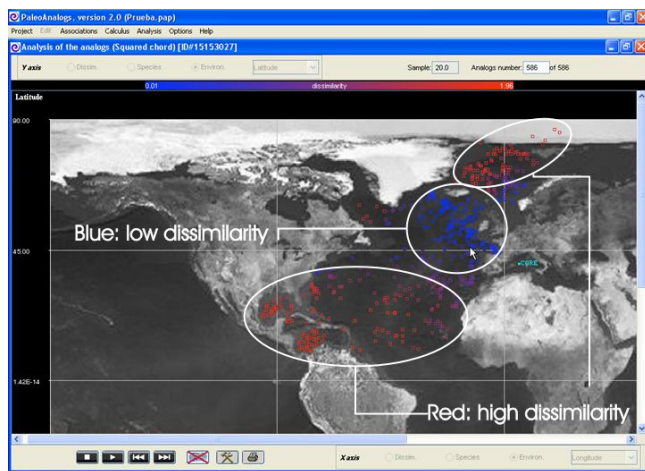


Figure 5: Analogs geographic distribution

As it has been stated in the previous sections, the problem is that paleontologists can obtain reconstructions as outputs from techniques such as nearest neighbor prediction, but no ways of knowledge acquisition are given. However, if the particular case of paleoclimatology is considered, ad hoc visualization tools may be developed, that will indeed provide insight in that forest of numeric data.

Geologists are trained for geographic visualization, and they basically face a problem of evolution through time, so we can design an interactive visual interface that takes advantage of both location and time.

Let's consider the following situation (figure 5): a paleoclimatologist is studying the data obtained from a particular point in the Mediterranean Sea (the label CORE in the figure shows that point)

and the modern data comes from the North Atlantic ocean. Instead of just calculate the temperature reconstruction, he/she can analyze first how the analogs for a given sample (depth/age) are distributed geographically. This can be seen in figure 5, which is the three dimensional (longitude(x axis), latitude(y axis) and dissimilarity(color)) representation for the sample at 20 cm². Thus, the expert would easily discover that the studied site, t kiloyears ago had temperatures much more similar to those of cold sites of today (blue zone of analogs in the picture) than those in warm or polar latitudes (red zone of analogs).

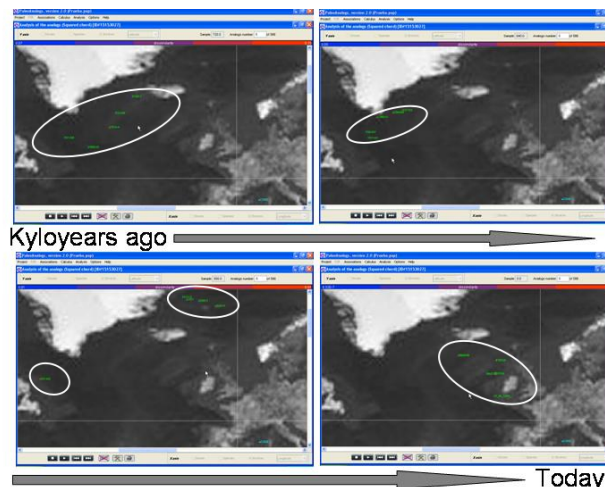


Figure 6: Best analogs trough time

Furthermore, this representation can be done for a selected number of analogs (just the number of analogs that will be used for the reconstruction, for instance, or for those with dissimilarity values smaller than a cutoff). Analog can be labeled with the associated database sample name so the expert might decide that a particular analog is not valid for the reconstruction due to a geological reason.

A combination of visualization approaches may discover a lot of information: choosing to show only 5 analogs, labeling each analog, zooming in, and animating the evolution of the whole core, i.e., visualizing the analog evolution through time, we may arrive to some interesting conclusions. Thus, in figure 6 we could start with the deepest sample in the core, i.e., t kiloyears ago: since at that age the planet was covered with ice, we can see that the best 5 analogs are distributed within a wide range of latitudes. As the animation is showing the evolution, we can see that best analogs are grouped, which is the typical distribution we should expect. The snapshot on the bottom-left shows a particular interesting situation: four of the five analogs are grouped up north (note the blue color), while the fifth one is located at a much warmer latitude (note the red color). This analog distribution should warn the paleoclimatologist, the most probable reason is that the outsider is only the fifth closest neighbor, but not a *real neighbor*, so that particular site should not be considered in the reconstruction for that sample.

3.3 A novel method for paleoenvironmental reconstructions: interactive parallel coordinates plots

Scatter plots, maps and animations are common methods for geovisualization that have a long history in cartography and information

²Depending on the particular age model this depth will be a number of kiloyears in the past.

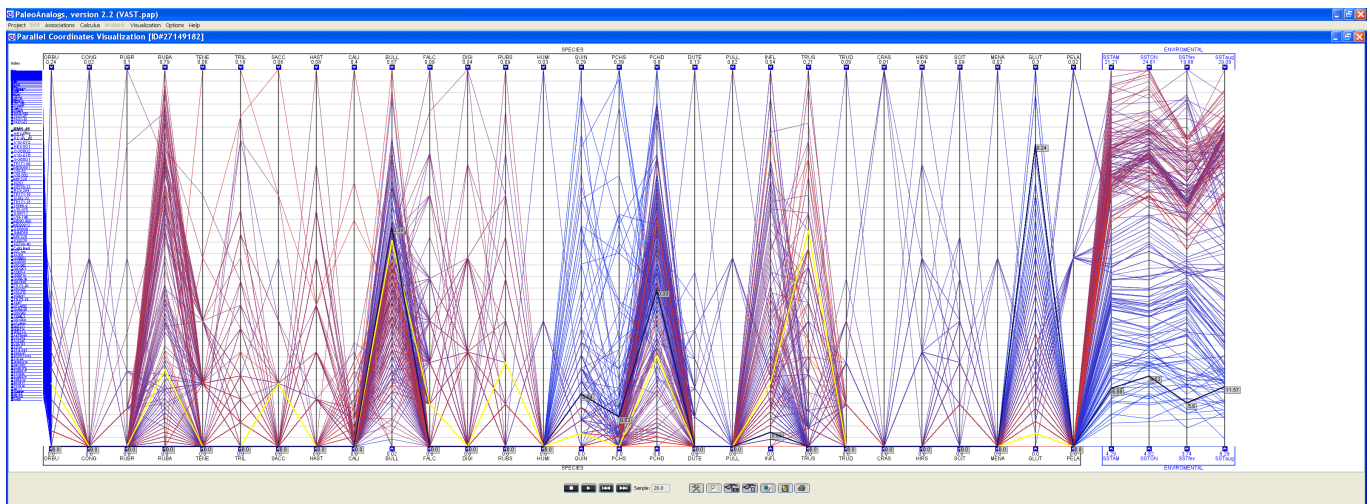


Figure 7: Reconstruction visually driven by Parallel Coordinates

visualization. Parallel Coordinates Plots (PCP) [6] are also a common method of information visualization, used for the representation of multidimensional data, and is an emerging practice in geo-visualization [9][14]. In the PaleoAnalog framework, the use of PCP (see figure 7) as part of the interactive visual analysis of multidimensional data can provide paleoclimatic knowledge discovery.

PaleoAnalog presents a unique use of interactive PCP; instead of just use it as another way of visualizing the data, it is used as a highly interactive tool that permits both gaining insight about the paleodata and visually reconstruct the paleoenvironmental features.

The data transformations (MAT) described above provides a mechanism for extracting patterns from the data. The output of this process is depicted using interactive PCP to facilitate the exploration of relationships among attributes. That is (see figure 7), each site of the database is drawn as a polyline passing through parallel axes, which represent the species, and the environmental variables that we want to reconstruct (last four axes on the right, in the example of the figure). The polyline corresponding to a particular sample (20 cm of depth in this example) in the core is represented as a yellow polyline. Note that, since the core only have the species data and we want to reconstruct the environmental variable for each sample, there are no yellow segments in the environmental axes. Each polyline of the database is color coded and the MAT technique is used for that purpose, i.e., the more red the polyline is, the more dissimilar is to that particular sample of the core.

This static picture is already showing many things that were hidden in the previous approach. For instance, it can be observed that the most similar sites for sample 20, are clustered in the low temperatures. This means that sample 20 corresponds to a cold period or a cold site.

However, several interaction techniques [17][8] have been integrated with this PCP to allow brushing [1], linking, animation, focus + context, etc., for exploratory analysis and knowledge discovery purposes enhance exploration.

3.3.1 focus+context

In the figure 7, a focus + context [17] technique can be seen for the labels of each site in the database. On the left side, each polyline is connected outside the axes with its label, allowing easy identification of each site. The labels are ordered top down, depending on different criteria chosen by the user: alphabetically, dissimilarity

index, latitude or no-crossing (the labels are ordered to avoid that the lines connecting with the values on the first axis produce any crossing). This way the expert can easily select the polyline of a particular site, that is highlighted in black and the values for each axis are shown. Since the context is maintained, the expert can access the label faster, depending on the chosen ordering criterium and the position of the current focus. Since there is a space shortage of space for such a number of labels, a fisheye approach [4], a powerful technique for organizing the suppression of irrelevant data, was developed.

3.3.2 Filtering and axis interaction

Another powerful feature in PaleoAnalog is dynamic filtering. PaleoAnalog provides a dynamic query on the PCP in the form of axis filtering [3][16]. The range of an attribute can be specified by moving the handles at the top and bottom of a range slider. The range sliders are embedded within the PCP.

To prevent users from losing global context during dynamic filtering, all the polylines are maintained on the background. Users can see the position of a polyline. Also labels are maintained on the background. Figure 8 shows a reconstruction already computed. After filtering the sites (the current ranges are shown in the handles) that were too dissimilar (maintained in dark gray on the background), the expert decided to reconstruct the SSTs for sample 20. Note that now the yellow segments of the polyline for that sample also occupy the environmental axis. As expected, the values are an average of the values of the blue polylines. On the left hand side, only the interesting site labels are highlighted. In the snapshot, the expert is comparing a site (orange polyline) with the reconstructed core sample.

Also note that in figure 8, all axis have the same scale (a percentage) in order to compare the relative abundances of the species, and help discover relationships between species and climatic features.

Another feature that helps in the interactive visual analysis is that any axis can be dragged and dropped, so the order of the axis is changed. This way the shape of the polyline also changes, helping to reveal hidden patterns and making analysis easier. In 9, axis order have been altered slightly so the PCP is uncluttered. The elimination of selected variables (axis) can also be helpful.

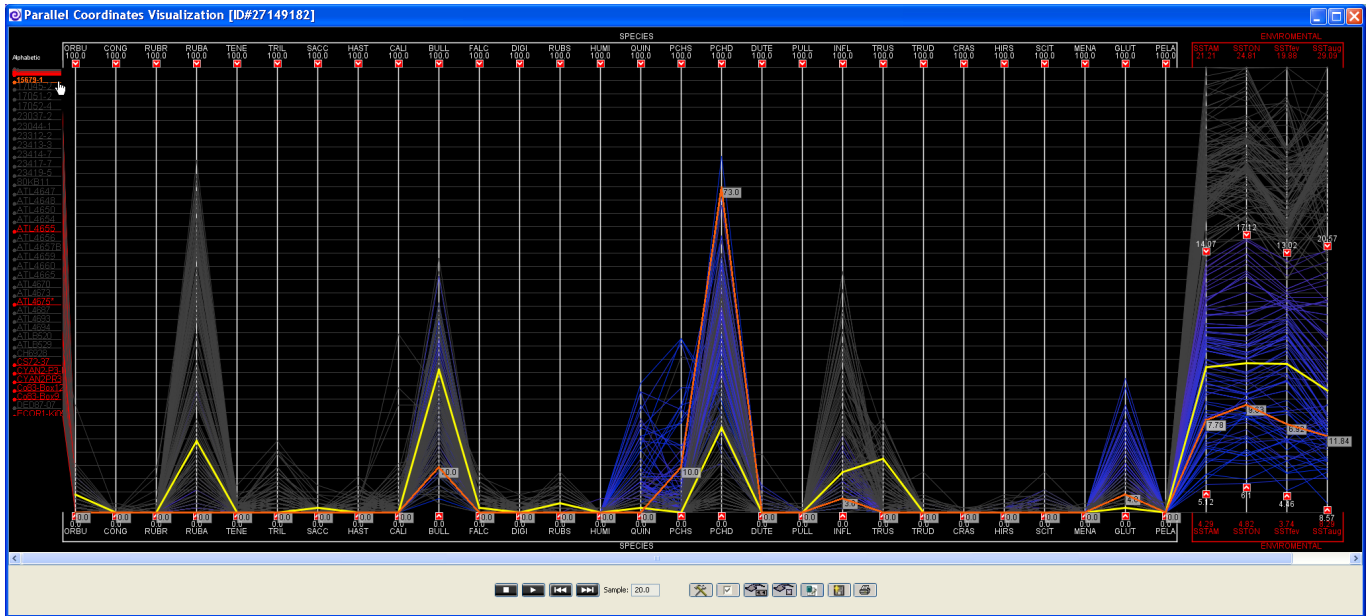


Figure 8: Analytical reasoning and reconstruction by means of interaction

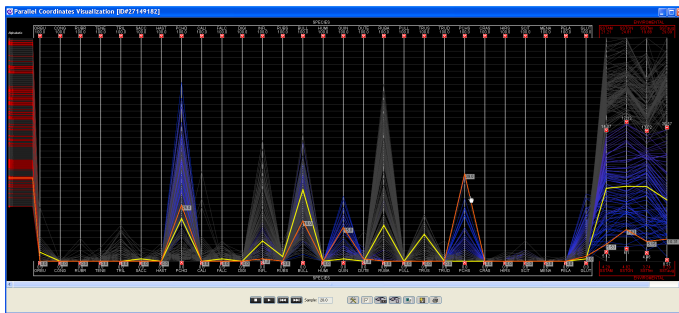


Figure 9: Uncluttering Parallel Coordinates by changing the axis order

3.3.3 Animation

On the contrary to other approaches such as [14], time is not represented on an axis. As in the previous cases, time is one of the most relevant variables in the analysis. PaleoAnalog can show an animation of the PCP. By doing this, it is possible to visualize the evolution of different species through geological time, and its relationships, both among themselves and between some species and SST, for instance.

3.3.4 Brushing and multiple linked views

Geolocation plays a key role in paleoceanography analysis. One possibility is to represent the latitude and/or longitude values for each database site. This can be interesting in order to highlight how temperature varies with time for the same latitude, for example.

In previous sections the benefits of using interactive maps in PaleoAnalog have been described. A common coordination technique is brushing and linking [1], where users can select objects in one view and the corresponding objects in all the other views are also automatically selected. This technique is the natural approach

for the problem at hand. This way, all the benefits explained above can be put together in order to provide the paleoceanographers with the best interactive visual tool to discover knowledge and support decisions about climatic reconstructions.

In figure 10 the brushing and multiple linked views approach of PaleoAnalog can be seen. As in [3], three modes of brushing and linking interaction that are coordinated among all the views described in the previous sections are possible:

- **probing:** this mode is used to view more details about an object (e.g. site labels and dissimilarity values) and to get an understanding of the relationships between the different views. Probing is a transient operation. Moving the mouse pointer over an object, highlight the object (e.g. a polyline) and as the mouse pointer is moved away, the highlighting disappears.
- **selecting:** this mode is used to mark objects that are of short-term interest, in order to further examine or perform operations on them (e.g. see the values on every axis of a selected polyline). Clicking on an object selects it and marks it. If a selected object is filtered, then it becomes deselected.
- **painting:** this mode is used to mark objects that are of long-term interest, in order to use them as references for comparisons (e.g. compare two polylines of two sites in the database). Objects remain painted until they are reset explicitly.

3.4 Presentation and Dissemination

As part of a future effort on providing a way to capture the analytic assessment of the paleoceanographers, it is possible to work on two different sets of colors, one for visual interaction and one for printing and image exportation (see figure 7). Also, an analysis is saved in a project, and the different stages can be independently saved and recovered. Also summaries, such as the associations used for one particular study are automatically produced.

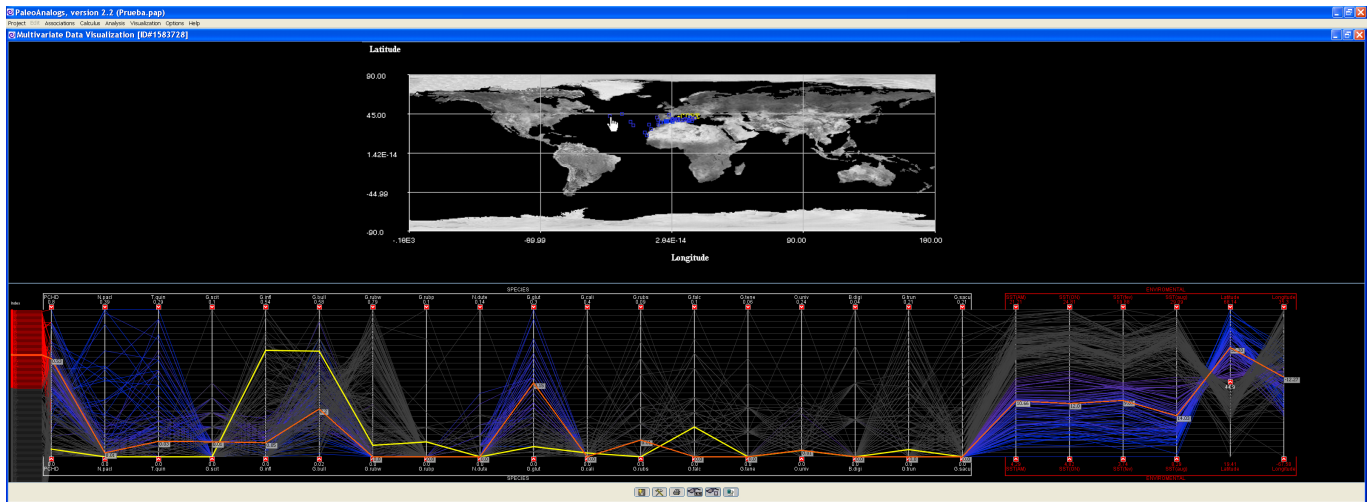


Figure 10: Multiple linked views in PaleoAnalog

4 CONCLUSIONS AND FUTURE WORK

This work is an example of how interactive analysis can help knowledge discovery in the paleoclimatology field. We have shown how a well-known standard technique of the field (MAT) can be greatly improved so the reconstructions of paleoenvironmental conditions can be more accurate. This is accomplished by fostering a user driven reconstruction procedure where the expert get more insight from the data and can decide on the validity of potentials reconstructions. Finally, we can add that more complex interactive analysis can be designed that will help to gain a deeper knowledge about the climatic evolution of a given area.

ACKNOWLEDGEMENTS

This work was supported by the MCyT of Spain under Integrated Action (Spain-France) HF2004-0277 and by the Junta de Castilla y León under project SA042/02. The author would like to acknowledge Ana Mayordomo García for her assistance in the implementation of this work.

REFERENCES

- [1] Richard A. Becker and William S. Cleveland. Brushing scatterplots. *Technometrics*, 29(2):127–142, 1987.
- [2] Leo Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
- [3] D. Brodbeck and L. Girardin. Design study: using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In *Proceedings of the International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pages 104–111, 2003.
- [4] George W. Furnas. Generalized fisheye views. In *Human Factors in Computing Systems CHI '86 Conference Proceedings*, pages 16–23, 1986.
- [5] W. H. Hutson. The agulhas current during the late pleistocene: Analysis of modern faunal analogs. *Science*, 207:64–66, 1980.
- [6] Alfred Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1:69–91, 1985.
- [7] Alfred Inselberg. Conflict detection and planar resolution for air traffic control. In *Intelligent Transportation Systems*, 2001.
- [8] Daniel A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.
- [9] E.L. Koua and M.-J. Kraak. A usability framework for the design and evaluation of an exploratory geovisualization environment. In *Proceedings. Eighth International Conference on Information Visualisation*, pages 153–158, 2004.
- [10] B. A. Malmgren, M. Kucera, J. Nyber, and C. Waelbroeck. Comparison of statistical and artificial neural network techniques for estimating past sea surface temperatures from planktonic foraminifer census data. *Paleoceanography*, 16(5):520–530, 2001.
- [11] H. Mannila, H. Toivonen, A. Korhola, and H. Olander. Learning, mining or modeling? a case study from paleoecology. In *Discovery Science*, pages 12–24, 1998.
- [12] D. Paillard, L. Labeyrie, and P. Yiou. Macintosh program performs time-series analysis. *Eos, Transactions, American Geophysical Union*, 77:379, 1996.
- [13] U. Pflaumann, J. Duprat, C. Pujol, and L. Labeyrie. Simmax: A modern analog technique to deduce atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments. *Paleoceanography*, 11:15–35, 1996.
- [14] A. C. Robinson, J. Chen, E. J. Meyer, and A. M. MacEachren. Combining usability techniques to design geovisualization tools for epidemiology. *Cartography and Geographic Information Science*, 32:243–255, 2005.
- [15] P. N. Schweitzer. Analog: A program for estimating paleoclimate parameters using the method of modern analogs. Technical Report 94-645, U. S. Geological Survey Open-File, 1994.
- [16] Jinwook Seo and Ben Shneiderman. Interactively exploring hierarchical clustering results. *IEEE Computer*, 35(7):80–86, 2002.
- [17] Robert Spence. *Information Visualization*. Addison-Wesley, 1 edition, 2001.
- [18] R. Theron, J. A. Flores, F. J. Sierro, C. Pelejero, J. Grimalt, and M. Vaquero. Using data mining and visualization techniques for the reconstruction of ocean paleodynamics. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume IV, pages 2382–2384, 2002.
- [19] R. Theron, J. A. Flores, F. J. Sierro, M. Vaquero, and F. Barbero. Paleoplot: A tool for the analysis, integration and manipulation of time-series paleorecords. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume VI, pages 3528–3530, 2002.
- [20] R. Theron, D. Paillard, E. Cortijo, J. A. Flores, M. Vaquero, F. J. Sierro, and C. Waelbroeck. Rapid reconstruction of paleoenvironmental features using a new multiplatform program. *Micropaleontology*,

50:391–395, 2004.

- [21] National Visualization and Analytics Center. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, 2005. IEEE Press, 2005.
- [22] C. Waelbroeck, L. Labeyrie, J.C. Deplessy, J. Guoit, M. Labracherie, H. Leclaire, and J. Duprat. Improving past sea surface temperature estimates based on planktonic faunas. *Paleoceanography*, 13:272–283, 1998.
- [23] P. Yiou, E. Baert, and M. F. Loutre. Spectral analysis of climate data. *Surveys of Geophysics*, 17(6):619–663, 1996.