

**"Multivariate Geovisualization of archaeological
events using the Space-Time Cube"**

Irvin T. Feliciano Santiago
March, 2008

"Multivariate Geovisualization of archaeological events using the Space-Time Cube"

by

Irvin T. Feliciano Santiago

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: (fill in the name of the specialisation)

Thesis Assessment Board

Chairman: Ms Dr. J.E. Stoter

External Examiner : Prof.Dr. M.J. Disjt

Supervisor : Dr. O. Huisman

Second Supervisor : Prof. Dr. M.J. Kraak



**INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS**

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

Geospatial data is often characterise by his accelerate production and complexity that involves. Geographical, spatio-temporal and multivariate data composed the most typical dataset in use. A large extent of software exists with the ability to use this kind of information in order to answers questions about locations, time and attributes domains. The field of Geovisualization concerns itself with providing the necessary tools and the adaptability to handle such geospatial data. On the other hand, when more dynamic information is accesses a serious challenge is posse to manipulate and convert the data into meaningful information. Multivariate Spatio-temporal data are the case of such example. Multivariate Spatio-Temporal data is characterized by changes in the characteristics of the object, time, and location and the relationship between them in order to explain the event or phenomena under study. A Geovisualization and visualization solution exits to handle the multivariate spatio-temporal characteristic of the data. But their limitations (as they are built for specifics purposes) not always present a reliable solution to explore such rich data sets when relationships between objects are desired. This research aims on developing a new multivariate spatio-temporal environment with the main objective to address the location, time and attribute domain as well relationship between the objects using archaeological sites as the case study. Combining the Space-Time and the Graphic Drawings visualization, the Semantic Multivariate-Time Cube is built. The prototype uses uDig platform and nourished using the main options, tools and functions of the Space-Time Cube and Graphic Drawings techniques. Moreover, the prototype is evaluated and conclusions are made.

Keywords: *Multivariate spatio-temporal data, Geovisualization, Semantic Multivariate-Time Cube, Space-Time Cube, Graphics Drawings.*

Acknowledgements

Dedicated to Tomas Feliciano Berrios and Elba I. Santiago Rosario for sacrifice the impossible for me and give me the opportunity to realize my dreams. To my brothers, sister and all my family who always support me.

My Special thanks to Farah Dwita to make the pillars form where I stand in this journey.

My gratitude's to my supervisor Dr. Otto Huisman and Prof. Menno-Jan Kraak for their effort and supervision.

To the Council for the Protection of Terrestrial Archaeological Heritage of Puerto Rico for encourage me to keep moving forward to archaeological issues and give me the first chance to test my abilities in the geoinformation field. My co-workers and supervisors Marisol Rodriguez and Belford Matias as well my dearest friend Erica del Valle and Dr. Jose Oliver.

Special Thanks to my model and example to follow MSc. Victor M. Cuadrado Landrau, to teach me the ways of geoinformatics and encourage me in this adventure.

To all the Natural Resource Department for helping me to acquire the information needed for this research. Thanks Felix!

My great gratitude to MSc. Bass Retsios, this research could never be done without his help, support and dedication.

Last, but not less to Nino Kheladze...forever.

Table of contents

1. Introduction	9
1.1. Background and motivation	9
1.2. Related work	10
1.3. Problem statement	11
1.4. Research objectives	12
1.4.1. Main objective	12
1.4.2. Sub-objectives	12
1.5. Resarch questions	12
1.6. Research hypothesis	12
1.7. Methodology	13
1.7.1.1. Literature review	13
1.7.1.2. Preparation and preliminary data analysis	13
2. Literaure review	15
2.1. Introduction	15
2.1.1. Space and Time components	15
2.1.1.1. Time and issues in history	15
2.1.1.2. Properties of Time and Space	16
2.1.1.3. Time geography and spatio-temporal phenomena	17
2.1.2. Methods, techniques, and concepts of spatio-temporal and multivariable data.	19
2.1.2.1. Representational views and frameworks for spatio-temporal data.	19
2.1.2.2. Spatio-temporal Maps	22
2.1.2.2.1. Cartographic Time Mapping	22
2.1.2.3. The Space-Time Cube	26
2.1.2.4. Multivariate data Visualizations	28
2.1.2.4.1. Traditional Graphs	29
2.1.2.4.2. Parallel Coordinate Plot (PCP)	29
2.1.2.4.3. Chernoff Faces	31
2.1.2.4.4. Lexis Pencil	32
2.1.2.4.5. Circular Axes-Base Visualizations	34
2.1.3. Methods, Techniques, and Concepts of Graphic Representations	37
2.1.3.1. Introduction	37
2.1.3.2. Readability parameters	37
2.1.3.3. Hierarchical Drawing Methods	37
2.1.3.4. Planar Layout Techniques	38
2.1.3.5. Force Directed Layout Techniques	39
2.1.3.6. Optimizing Aesthetic Criteria	40
2.1.3.7. Dynamic Graph Drawing	41
2.1.3.8. Drawing Graphs in 3D	42
2.2. Summary	44
3. Model design	47
3.1. Introduction	47
3.2. Space-Time Cube	47

3.2.1.	Justification of Space-Time Cube	47
3.2.2.	Space-Time Cube Visualization for Archaeological Sites.	51
3.3.	Semantic Network Graph	54
3.3.1.	Justification of Semantic Graph	54
3.3.2.	Graphic Drawings Visualization for Archaeological Sites.	57
3.4.	Semantic Multivariate Time Cube	59
3.4.1.	Justification of Semantic Multivariate Time Cube	59
3.4.2.	Semantic Multivariate-Time Cube Visualization for Archaeological Sites	63
3.5.	Multivariate Temporal Framework	65
3.6.	Combined Tools and Options Evaluation	65
3.7.	Prototype Stages	65
3.7.1.	Stage 1 (Java Code Matrix Implementation)	67
3.7.2.	Stage 2 (Java Code Footprint Data Mining Implementation)	68
3.7.3.	Stage 3 (Additional Options)	69
3.8.	Summary	69
4.	Prototype: case study archaeology	71
4.1.	Introduction	71
4.2.	Geography of Puerto Rico (Area of Study)	71
4.3.	Case Study: Archaeology Case PR	72
4.4.	Data for the Case	72
4.5.	Cluster Analysis	73
4.5.1.	Introduction	73
4.5.2.	Data Visualization Techniques	74
4.5.3.	Optimization of the Selected Clusters	74
4.5.3.1.	Questions to be Answered	74
4.5.3.2.	Cluster Optimization Techniques	74
4.5.4.	Visual Exploration	76
4.5.5.	Cluster Optimization	77
4.5.5.1.	North East Cluster	77
4.5.5.2.	North Center Cluster	78
4.5.5.3.	Center Cluster	80
4.5.5.4.	South West Cluster	81
4.5.5.5.	South East Cluster	83
4.5.6.	Cluster Analysis Conclusions	84
4.6.	GIS Analysis Model (attribute transfer)	84
4.7.	Independent and dependent Variables	85
4.8.	Semantic Multivariate Cube Implementation	86
4.8.1.	Spatio-Temporal Questions	86
4.8.2.	Visualizations	86
4.8.3.	Limitations	98
4.8.4.	Summary	98
5.	Evaluation	101
5.1.	Introduction to the focus group evaluation	101
5.2.	Discussion session and group members	102
5.3.	Discussion session objectives and questions	103

5.4.	Focus group results	104
5.4.1.	General results by theme	104
5.4.2.	Synopsis	106
5.5.	Summary	106
6.	Conclusions	107
6.1.	Introduction	107
6.2.	MSTC (Multivariate Semantic-Time Cube)	107
6.3.	Further Work	108
7.	Bibliography	111

List of Figures

Fig. 2-1. Illustration of changes in temporal and spatial domain (Blok, 2005).....	16
Fig. 2-2. Basic appearances of a timeline: a) the straight line (linear time - years); b) the circle (cyclic time - seasons); c) spiral (skewed time - geology) (Kraak, 2005).....	17
Fig. 2-3. The Space-Time Cube concept model and some of the most important time geography elements (Moore et al., 2003).....	18
Fig. 2-4. The Dual representational Framework (Peuquet, 1988).....	19
Fig. 2-5. Ozone concentration map California (Arc.GIS 9.2 Geostatistical Analyst Tutorial).....	20
Fig. 2-6. The Triad Representational Framework (Peuquet, 1994).	20
Fig. 2-7. Triad Framework of spatio-temporal data.	21
Fig. 2-8. The operational task topology, (Andrienko et al, 2003).....	22
Fig. 2-9. Color tints representing ages of built-up areas (De By, 2004).	23
Fig. 2-10. Napoleon Russian campaign - 1812-1813(Kraak, 2003).....	23
Fig. 2-11. Sequence of maps depicting the urban growth: Maastricht city, The Netherlands (Principles of Geographic Information Systems book).	23
Fig. 2-12. Map series representing the “paleogeography” of North America over the last 550 million years of geologic history (Ron Blakey).....	24
Fig. 2-13. Simulation of an animated map (Principles of Geographic Information Systems book). ...	24
Fig. 2-14. Space-Time Cube: Napoleon’s Russian campaign (Kraak, 2003).	26
Fig. 2-15. Human walkthrough visualization in Valla Campus: Linkoping University. Different colors represent different persons (Kristensson, et al. 2007).....	27
Fig. 2-16. The Black Death epidemic case: a base map of Europe serving as geographic reference for the disease locations (Kraak and Madzudzo, 2005).....	27
Fig. 2-17. Different time lines representing archaeological periods (Kraak, 2005).	27
Fig. 2-18. Graphs for time-dependent data: left: Sector Graph, middle: Stacked Bar Chart, right: Circle Graph (after Harris, 1996), taken from (Muller and Schumann, 2003).	29
Fig. 2-19. Seven dimension PCP. Seven attributes (4 geographic and 3 spectral) are used to explore geographic phenomena, (Gahegan et al, 2001).	30
Fig. 2-20. Eight universal expressions: happiness, elation, anger, sadness, disgust, determination, fear, surprise (Ekman and Friesen, 1975; Ekman, 2003).	31
Fig. 2-21. Chernoff faces. Different data variables are mapped to the sizes and shapes of different facial features (Ware, C., 2004).	32
Fig. 2-22. A classical Lexis Diagram. The x-axis depicting calendar time and y-axis representing age or time occupied by the events (solid sloping line) (Pressat, 1961).....	33
Fig. 2-23. The 3D Lexis Pencil. Each pencil stand for an object or events, the pencil height describes time and colors variables (Muller and Schumann, 2003).....	33
Fig. 2-24. The 2D TimeWheel compounded by six variables (circle set) and time variable (central placement) (Tominsk, et al, 2004).	35
Fig. 2-25. Two variants of the 2D MultiComb (Tominsk, et al, 2003).	35
Fig. 2-26. Variants of the 3D Axes-Base representations. Upper left: 3D Time Wheel, Upper Right: TimeWheel “folder style”, Down Left: 3D MultiComb, and Down Right: 3D Kivat tube, (Tominsk, et al, 2003).....	36
Fig. 2-27. Example of improvements of a hierarchical drawing. (Sugiyama et al., 1981).	38

Fig. 2-28. Example graph implementing the SA algorithm base on user define aesthetics (R. Davison and D. Harel, 1993). (a) Input nonplanar graph. (b) Output planar graph map.	39
Fig. 2-29. Example graph implementing barycenter approach (W. Ahmada and A. Khokhar, 2007)..	40
Fig. 2-30. Example graph implementing criteria base method (Ulrik and Dotothea, 1998).....	41
Fig. 2-31. Example of a Dynamic Hierarchical Radial Layout. Node A (left image) is selected to become the new focus. Node A's current parent is node B; the new layout (right image) is oriented so that the edge AB maintains the same direction (Yee et al., 2001).....	42
Fig. 2-32. Interpolation in rectangular coordinates (left) can yield a confusing animation. Interpolation in polar coordinates (right) works better for radial layouts (Yee et al., 2001).	42
Fig. 2-33. Example of 3D Cone Tree graph (G. Robertson et al., 1991).	43
Fig. 2-34. Example of 3D Hyperbolic Space graph (Munzner and Burchard, 1995).	44
Fig. 3-1. Steps for a TGS to design a geovisualization environment and understand spatio-temporal phenomena (Source: Author).	48
Fig. 3-2. The Relationship between the Triad Framework (Peuquet, 1994) and the STC environment (Source: Author).....	48

List of tables

Table 2-1. A table graph describing the relation of maps displays, and its components (Koussoulakou and Kraak, 1992).	25
Table 3-1. Attribute table depicting similarities between objects. Dashed lines: as differences in attributes values (Source: Author).	58
Table 3-2. A Multivariate Temporal Framework (Source: Author).....	65
Table 3-3. Tools and Options Evaluation (Source: Author).	66
Table 4-1. Data set in use. Orange color is given to the independent variable (archaeological sites).	85
Table 4-2. Attributes information (archaeological sites clusters).	86

1. Introduction

1.1. Background and motivation

Rapidly growing software and hardware technology allows us to store, retrieve, and analyze highly diversified information. Often, the magnitude and/or complexity of geospatial data pose a serious challenge on how to convert the data into meaningful information and ultimately transform it into knowledge. The field of Geovisualization concerns itself with providing the necessary tools and the adaptability to handle the complexity of such geospatial data. It is important for geovisualizers to reassess the ways to explore, understand and satisfy the user's need in order to develop new applications that provide more meaningful results. This is well described by Keller and Keller (1992), who in their approach to the visualization process suggest removing mental blocks and taking some distance from the discipline in order to reduce the effects of traditional constraints (Keller and Keller, 1992 in Kraak & Ormeling, 2003).

Archaeology is a discipline that could benefit from dynamic geovisualization techniques, since it is concerned with the study of material remains (architecture, artifacts, etc), locations of discoveries (where artifacts were found and when) and environmental data (landscapes) to help understand and reconstruct past societies. In archaeology, the main objective is to locate archaeological sites in the landscape in combination with temporal information. Apart from the most common attributes i.e. geographic locations and time period, it is also important to collect attribute data about the cultural characteristics and the surrounding geographical environment. Archaeological sites therefore present a rich environment of information attributes.

With regard to cultural variables, the most typical variables are related to ceramic styles (that target the period), types of artifacts (bones, stones, and shells), foods remaining etc. Considering the geographical context, some of the key variables that are of interest to the researcher are; proximity to water supplies, elevation, slope, geology, soil type and cluster identifications. This is probably because these variables have an impact on human survival. All these characteristics when analyzed together with time could provide a better understanding of the context in which the artifacts was found. As time passes, new archaeological sites will appear and others disappear and therefore, the set of variables related to geography will also change.

As mentioned before geospatial data is often characterized by the complexity of related items. Geovisualizers often need to represent data that consist of items related together. Such data sets can be abstracted to a mathematical structure, i.e. the graph. A graph contains nodes and edges where the nodes represent the items or concepts of interest, and the edges connect two nodes together according to some association scheme (Peter Rogers, 1995). A good graphic representation allows the users to investigate the data in more meaningful manner. For example, identifying the clusters of related observations or finds (represented as nodes), discovering general patters in data, and exploring represented relationship of the nodes will aid in further understanding of individual observations.

Several issues are responsible for the lack of existing research and development of Geovisualisation tools for Archaeology. These include the lack of data, data quality issues, database inconsistency. The second is that the most archaeologists are currently interested in contract archaeology leaving little or no room for scientific research. The third is that many existing Geovisualisation approaches and techniques have a number of limitations which makes them unsuitable for complex applications, such as Archaeology:

- Static Maps can present only limited information to the user. The primary task of static maps is a focus on communication of a particular message, and not on the user's ideas or insights.
- A series of static maps can by definition only present a limited number of images. The time can only be incorporated through sequences of maps or images, and it is generally difficult to deal with long series. There is often also no means of controlling these.
- Animated maps provide some dynamics, but it is difficult for the viewers to compare between two 'slices' that have a large temporal separation.
- Simple Graphs are limited to describing/displaying only few variables, and it is often difficult to compare between variables, especially as the amount of data increases.
- Parallel Coordinate Plots (PCP) are another common Geovisualisation technique. Large amounts of variables can easily overload the PCP representation making difficult to detect patterns on the data. Also, ranges and the order of the variables can influence understanding/pattern discovery.

The Space-Time Cube (STC) has demonstrated the power to manipulate complex relations of particular events across time. The concept was introduced by Hagerstrand (1970). Hagerstrand developed a graphic view on time as an additional spatial dimension. He suggested a three-dimensional diagram, the so called space-time cube, to show the life history of peoples and how people interact in space and time (Gatalsky et al, 2004). Within the STC, temporal variables can be represented along the Z axis and either the time point or the continued duration can be represented. X and Y axis can be used to represent spatial data, and the emphasis placed on spatial and temporal relationships (Xia, 2005). Features like the Space Time Prism, Space Time Path, Potential Path Space and Potential Path Area are concepts used to investigate and understand the movements of events through space and time.

1.2. Related work

Over the years researchers have been interesting in the applicability of the Space-Time cube to address different problems. For example: Moylan (2001) explore the different landscapes changes in a cell base cube model. Kraak, (2003) evaluate the visual opportunities of the Space-Time Cube in different applications, i.e., sports, Google Earth, diseases, battle events etc. Gatalsky et al (2004) implement the visual environment of the Space-Time Cube to demonstrate how can be used to detect spatio-temporal cluster of events using earthquakes data. Xia, (2005) design a PCP Time Cube to explore the temporal, spatial and multivariable characteristics for medical cases.

An archaeological application was made by Kraak and Koussoulakou (2004) in which the spatio-temporal characteristics of the archaeological sites was explored. "Complex relations" between artifacts excavated at particular times found from different historical periods at different locations could be made visible in the cube. The location of a find can be represented by a vertical line that has a color/thickness at the relevant time period. For example, to indicate the find of a particular type of

artifact (color), and with a particular value (thickness) for the duration of the time period (Kraak and Koussoulakou, 2004). Developments in geovisualization such as linking and brushing allow one to link the space time cube views to other alternative graphics (Kraak and Koussoulakou, 2004).

To summarise, the current state of the STC is that's not possible yet to visualize events base on some kind of association schema to explore the spatio-temporal and multivariable characteristics of the events. Different cultural attributes can be analyzed to gain insight into the spatio-temporal distribution of sites while taking into consideration the relationship between these attributes.

1.3. Problem statement

This research aims to demonstrate what Geovisualization and the Space-Time Cube (STC) can offer to Archaeology (and related fields), in order to improve understanding and insight, through improving and extending visualization methods and techniques. As noted, the Space-Time Cube offers a visual environment to handle spatio-temporal events. Additionally, the STC posses the flexibility to link to others views like: a static map, internet links, Parallel Coordinate Plots (PCP), graphs charts, animations, and so on. But is also true that the saturation of multiple linked views can cause the user to get confused while exploring the data.

As discussed above, the Space-Time Cube has the ability to represent relationship of objects and events in space and time, but also has a number of limitations:

- Difficulty in search for specific themes, not for specific period of time.
- The amount of the events that can be displayed and interpreted effectively using time line is limited.
- There no way to represent the relationship of temporal events at different locations.
- The heights of the 3D objects present a challenge when trying to combine with other kind of height data in the same display.
- Difficulty to deal with uncertainty of time.

In order to overcome such difficulties and to satisfy users' needs, geovisualizers have to think in better ways to provide more functionality within the STC environment. Specifically, to find ways to explore the relationship between attribute data of archaeological events. One possible way of doing this is by combining graphic drawing techniques base on similarities of geographical variables, including:

- elevation
- slope
- geology
- soil type.

The archaeologist is primarily interested in the timing of particular events existence (archaeological sites periods) in order to establish the belonging type of culture and to determine possible culture or clash interactions. It is necessarily for the exploration of these kind of events to have a particular interface that allows the time (culture) filter and temporal ordering in order to investigate these things.

Once it is determined that a possible interaction have taken place, attention can be directed on discovering how events are related (geographical or cultural) and the degree of this relationship. In

here archaeologist focus on distinguish what kind of similar attributes (artifacts or geographic variables) the events posses. For that reason a visual graphic representation based on some kind of network that address the similarities of the archaeological events plays an important role to visualize the relationship. It is also important to be able to design an approach that is capable to illustrate the degree of the relationship. A weighting table in where the similarity of attributes could be evaluated using simple computations could provide the necessarily information to address the degree of relationship.

These proposed extensions will aid in some extent one of the basic principles of the archaeological field that is to reconstruct the social interaction of past cultures. In so doing, it will provide an extended visual environment to the STC that was designed in the first place to explore and visualize how people interact in space and time.

1.4. Research objectives

1.4.1. Main objective

The main aim of this research is to provide an extended Geovisualization environment to explore the spatial, temporal and attribute characteristics of archaeological events.

1.4.2. Sub-objectives

- 1) To investigate the linkage between significant archaeological sites and certain geographical contexts.
- 2) To design an interactive interface in which the events can be addressed using their temporal domain (cultural period).
- 3) To implement and evaluate the ability of specific graphic representations to help visualize the relationship of the attributes of archaeological sites inside the Space-Time Cube.

1.5. Resarch questions

- 1) What graphics representations techniques are suitable for the display of the attributes of archaeological sites and what are their advantages and disadvantages?
- 2) How can the graphic representation technique be integrated into the Space-Time Cube?
- 3) What functionality should be included in the Space-Time Cube to deal with time period and the visual representation of the graphic representation techniques?
- 4) How effective is the extended STC environment for Archaeological data? E.g. What are the relationships among archaeological sites? Can we detect archaeological sites clusters of similar geographic relationship? Do the archaeological clusters change over time?

1.6. Research hypothesis

- 1) Combining the Space-Time Cube and extended graphical representation techniques can result in a more meaningful visualization environment for the interpretation of archaeological data.
- 2) An interactive manipulation of the Temporal dimension can provide clues for better understanding of archaeological sites (e.g. by grouping, filtering and selection).

1.7. Methodology

1.7.1.1. Literature review

Using the knowledge obtained from the previous studies done by researchers as regards to geovisualization analytical capabilities (basic display, query, animation, box plots, multiple dynamically linked views, etc.), this research will provide an extended visualization environment to help investigate the relationship between archaeological sites using both cultural and geographical variables. This will require literature reviews of the following:

- 1) Methods, techniques, and concepts of spatio-temporal and multivariable data.
- 2) Methods, techniques, and concepts of graphic representations and techniques that best fit the data.
- 3) Methods and technique to manipulate and transform qualitative to quantitative information).

1.7.1.2. Preparation and preliminary data analysis

- 1) Clean up and preparation of the archaeological sites database:
 - i. Choosing those sites that contain the most reliable temporal information
 - ii. Combination of the temporal characteristics of the archaeological sites with the selected geographical characteristics
- 2) Determine attributes of interest (elevation, slope, distance to water supplies) which best describe the Archaeological sites using the Parallel Coordinate Plot (PCP)
- 3) Generate the graphic representation to form a networking relationship for the archaeological sites based on the degree of attribute relationship(s)
- 4) Evaluate and interpret the results and compare the temporal changes of the geographical variables.

2. Literature review

2.1. Introduction

Once the problem, questions and objectives have been set, a literature review to address the most important issues in this research will be performed. In order to expand the knowledge to meet the necessary requirements to incursion this survey the following topics to explore are: concepts, techniques and visualization methods of spatio-temporal, multivariate and graphic drawing techniques. The following sections will discuss these topics in detail as well others related issues.

2.1.1. Space and Time components

2.1.1.1. Time and issues in history

Time in human history has been a major subject of science, art philosophy etc. Usually time is perceived via the changes that occur to objects in space (i.e. transformations and movements). Conversely, using time as a measuring unit, geographical phenomena (humidity, air pressure, heat, etc) that lack physical properties can be addressed. From these two perspectives we can imply that both time and space are intertwined, and should be examined together for the study of 'spatial' processes. Whatever the process under scrutiny, inclusion of the temporal element in the data is required so that changes can be represented, cause and effect relationship can be derived from the observational data, and ultimately understanding for the nature and structure for the processes involved can be advanced (Gregory, 1982), taken from Peuquet, 1994. We must be careful when we talk about cause and effect relationships. The perception of causation is another unfinished battle between physicists, philosophers and social scientist. For example Newton's ideas concerning force, activity and motion. He stated that "It's inconceivable that inanimate brute matter should, without the mediation of something else, which is not material, operate upon and affect other matter without mutual contact."

But if we think more like a social scientist (for example, an archaeologist), a phenomenon that involves changes does not necessarily have to experience physical contact. A human can be influenced into executing an action according to some magical or religious belief. For example, an inanimate object can possess the ability to affect a society in their psyche. In Puerto Rico one of the most important indigenous settlements (Caguana) is situated around a mountain formation. The belief that this object (mountain) posses some magical-religious power, influenced them to built and change the landscape around. We can say that for this particular case, the cause (mountain) use a non-physical medium (psyche) to affect the individuals and make them act.

When we deal with spatio-temporal phenomena we perceive changes according to the sudden break of the behavior of an object in time-space. In here the time and space components serve as a medium in which the objects interact. In this regard, we can detect changes only to factors that affect our senses and the instruments we use. A description of the dynamics of geospatial phenomena (changes in temporal and spatial domain) are discussed by Blok (2005). Temporal changes refer to: a moment in time, peace, sequence, duration and frequency. On the other hand spatial changes refer to: appearance or disappearance, movement, and mutation (see Fig. 2.1).

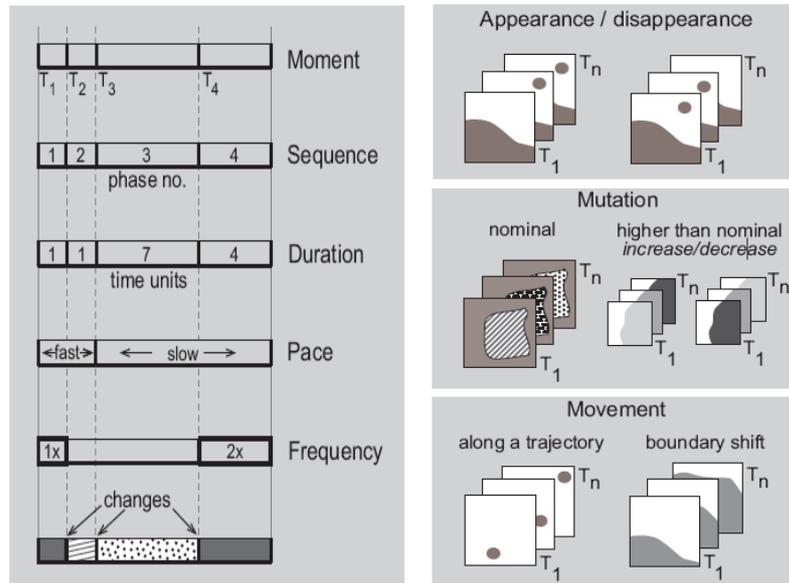


Fig. 2-1. Illustration of changes in temporal and spatial domain (Blok, 2005).

2.1.1.2. Properties of Time and Space

As Newton stated, absolute space-time trims down everything to loose bodies in space making the space a container in where these bodies exist. As this view focuses on space-time as a subject mater, it is considered to be objective. The objective view assumes an unbreakable solid-geometric structure in which the objects may or may not occur. In here every process is linear, evolutionary, and irreversible.

Conversely, in the relative space-time described by Einstein, both space and time are positional qualities that are attached to each object. This subjective view concentrates on objects themselves as a subject mater. It assumes a more flexible structure than the objective view in the sense of association between objects in where time and space are strictly inseparable. Hence interpretation of process and the flux of changing patters and process within specific phenomenological contexts is the target of the relative view.

Depending on the purpose of the investigation, and the amount and type of data obtained, a researcher trends to adopt one of the fundamental representation of time (absolute or relative) that fit with their immediate needs. For example, change detection scientists are interesting in monitoring changes in the land cover over time and understand the potential causes. Change detection is a technique used in remote sensing to determine the changes in a particular object of study between two or more time periods. The evaluation of a series of satellites images according some cyclic time period is necessarily in order to perceive the changes. For that reason the adoption of a relative view of time is usually taken.

Absolute dating is a process that most of the archaeologists use in order to state a more specific sequence for the archaeological events. For example, the technique of carbon-14 is used to calculate the degree of carbon isotope loss into the conversion of hydrogen-14 (Mitchels, 1972). By measuring

this proportion archaeologist can estimate the “precise” date of archaeological remains. Hence, for this type of procedure an absolute view of time is preferably adopted. In order deal with analysis of spatio-temporal process is necessarily to combine these perspectives, as in many situations they are complementary not contradictory (Peuquet, 1994).

According to the basic property of time taken and the data type in use, different measurement units are applied. As described by (Kraak, 2005) these measuring units or “granularities” are related to time resolution, uncertainty, description or classification of time (see Fig. 2.2.).

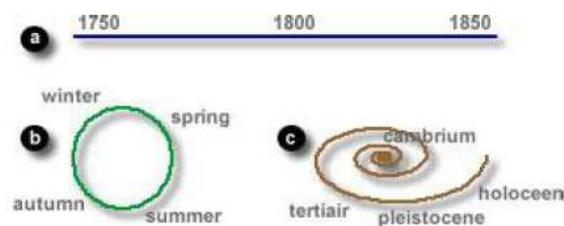


Fig. 2-2. Basic appearances of a timeline: a) the straight line (linear time - years); b) the circle (cyclic time - seasons); c) spiral (skewed time - geology) (Kraak, 2005).

2.1.1.3. Time geography and spatio-temporal phenomena

Perhaps one of the first pioneers in joining the spatial and temporal properties was Torsten Hägerstrand (1970). The proposed model of Hägerstrand (1967; 1970) focussed on exploring the natural or cultural phenomena patterns in space and time. Explicit in the paper of “What about people in regional science?” a 3D method the “space time aquarium” is introduced in order to represent the people in space and time. In his model space and time are seen as inseparable. It could be said that a relative view of time is adopted as a framework of the spatio-temporal process under study. The model environment takes form of cube geometry, in where two of the horizontal axes are used to represent the X, Y geographic coordinates and a vertical axis (t) used to represent the time dimension (see Fig. 2.3).

The so called Space-Time Cube (STC) is oriented to deal with the following time geography element: lifeline, prism, bundle, station, and domain:

- **Lifeline:** paths traveled by the object from one point to another. The slope of the line represents movement speed.
- **Prism:** indicates the locations that can be reached by an object in a certain time.
- **Bundle:** a grouping of several path locations depicting possible interaction.
- **Station:** vertical line represent stationary object usually seen as a vertical tube.
- **Domain:** a control area that serve as containers to restrict objects inside or outside the perimeter.

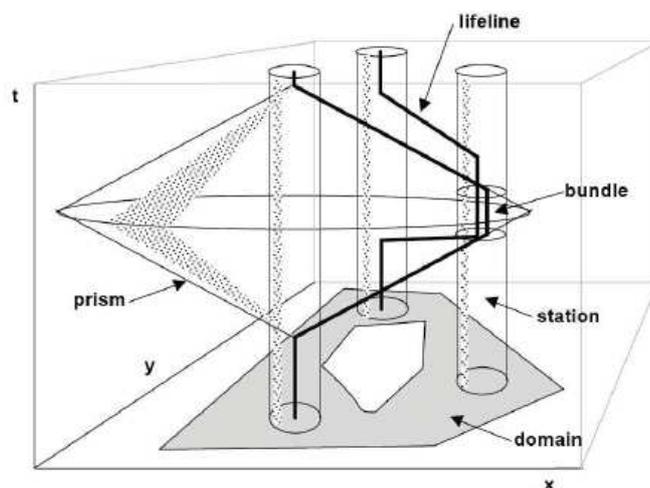


Fig. 2-3. The Space-Time Cube concept model and some of the most important time geography elements (Moore et al., 2003).

A more detail description of the STC model and functions is discus in section (2.2.4).

Hagerstrand's original ideas developed into the field now known as Time Geography (TG). TG deals with complex space time phenomena at the individual level, observing paths of individuals through space and time and their interactions. When these original concepts were introduced, technology in computers (particularly visualization) was in its infancy. In order to communicate the message the individual was limited to manually draw and redraw new version of the cube involving a hard work and time consuming drawing activity. In that regard development in "digital cartography" help TG take it's from by taking the advantages of technological development (computers, software's, databases, retrieval data instruments, etc.) towards more dynamic visual representations. Now days, new versions of the cube can be automatically created with a matter of seconds and others graphic views can be linked to support the exploration of such complex representations (Koussoulakou and Kraak, 1992). Much of the functionalities of "digital cartography", known as today as Geographic Information Systems (GIS), can be found lying on software that claims to be in proper time geography applications (those that include basics time functionalities and are intentionally commercialized as specialize applications). In order to facilitate the exploration and analysis of the geospatial data, manipulation, analysis operation, database support, statistical analysis, and the combination of non-geospatial data help to answer some of the basic questions (what, where, and when) involving geographical as well as spatio-temporal phenomena. This fusion helps to overcome the deficiencies of the traditional GIS (place-base) perspective and redirected time-geography into a more peapole base-perspective (Koussoulakou and Kraak, 1992).

2.1.2. Methods, techniques, and concepts of spatio-temporal and multivariable data.

2.1.2.1. Representational views and frameworks for spatio-temporal data.

In the early days of GIS, scientist usually tends to represent the geographic phenomena, purely oriented into a traditional vector or grid models. Thus, both views served well according the geographic phenomena under investigation, much more meaningful interpretations can be addressed combining the two approaches. A dual representation framework was presented by Peuquet (1988), in where the object based model is related to vectors and the location based to raster, (see Fig. 2.4). In her model, objects are seen as purely geometric (e.g., point, lines, and polygons) structures possessing spatial and non-spatial information (attributes). On the other hand, locations are perceived as “point with no area or volume or as an areal cell or pixel with a gridded data structure” (Peuquet, 1994). Moving from one model into another a more synthesized view of the geographic phenomena can be represented.

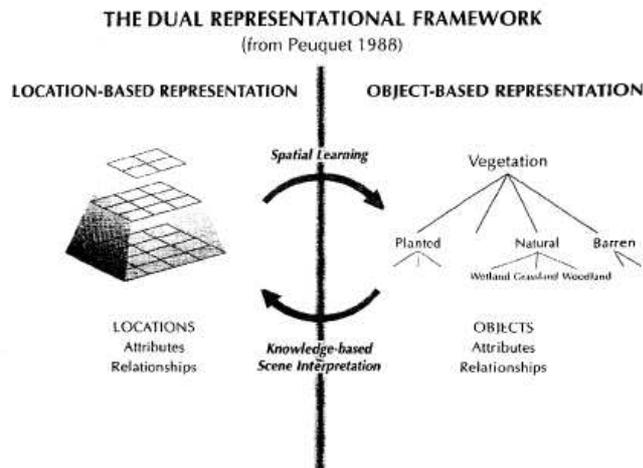


Fig. 2-4. The Dual representational Framework (Peuquet, 1988).

For example, both views can be combined in order to make predictions in locations where no data is available. In Fig. 2.5, a real case scenario for monitoring ozone concentrations in California is presented. In here points represent stations in where ozone concentrations are known. An interpolation method is used in order to predict those locations of insufficient information. The result is an object based representation view compounded by points and lines serving as reference related information within the location based view (interpolated raster).

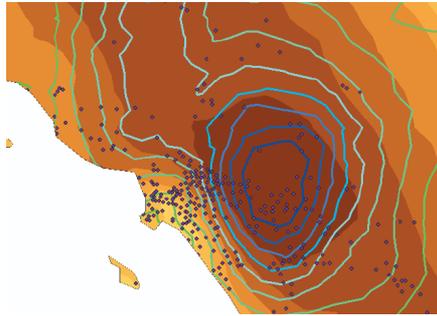


Fig. 2-5. Ozone concentration map California (Arc.GIS 9.2 Geostatistical Analyst Tutorial).

But what about when the geographic phenomena under study require the inclusion of an additional dimension (time)? In that regard The Dual Representation Framework lacks the ability to represent efficiently spatio-temporal phenomena.

Subsequently, a Triad Framework was developed by (Peuquet, 1994) in which the three basic components of spatio-temporal data are acknowledged. This model presents an extension of the dual (location “where” and object “what” based view) spatial representation framework, (see Fig. 2.6.). The inclusion of the time (time “when” base view) component form the triad framework (location, time and objects) permits the user to pose there basic core questions to deal with spatio-temporal process:

- **when + where → what:** describe the particular or set of objects (what) that are presents at given locations or set of location (where) at given time or set of times.
- **when + what → where:** describe the locations or set of locations (where) occupied by a given object or set of objects (what) at given time or set of time (when).
- **where + what → when:** describe the times or sets of times (when) that a given object or set of objects (what) occupied a given location or set of locations (where).

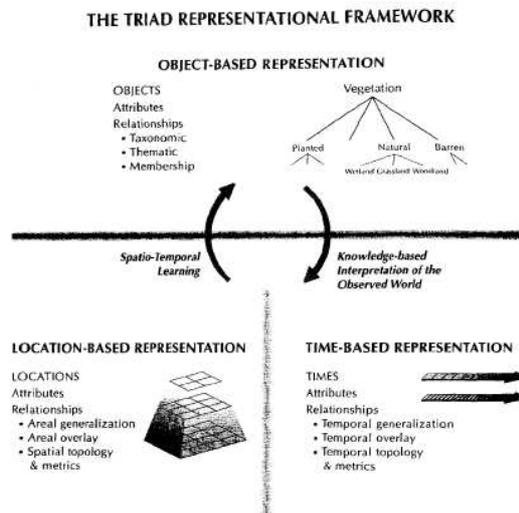


Fig. 2-6. The Triad Representational Framework (Peuquet, 1994).

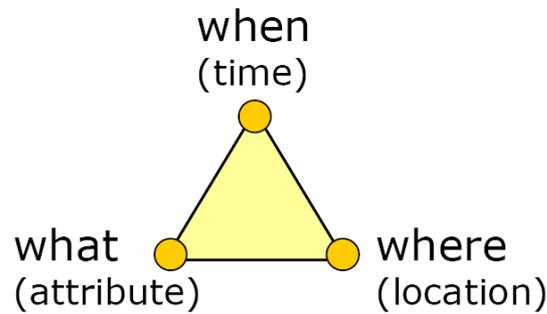


Fig. 2-7. Triad Framework of spatio-temporal data.

Additionally to the Triad Framework, Bertin's framework (1983), based on types of questions and readings levels rest. He mentioned that the amount possible questions are strictly related to the amount of variables in use, and for each type of question three readings levels (elementary, intermediary, and overall) exist. These readings levels focus on the data elements (single, group, or whole) found inside the question to be answered. The advantages of considering both frameworks for spatio-temporal phenomena is that they can help to anticipate what kinds of questions may arise and the proper tools to answer them (Andrienko et al, 2003).

An interesting approach combining the geographic variables of Bertin and the notions of reading levels for spatio-temporal data was discussed by Kraak and Koussoulakou (1992). In the sense of the posed questions regarding spatio-temporal process the TG scientist should be able to decompose the question according to its spatial, object and temporal components. For example, the question "What is the trend of changing's values at location 1?" belongs to the elementary level in relation to the **spatial** component and to the overall level with the respect to the **temporal** component (Andrienko et al, 2003). The use of the graphic variables considering the readings levels allows the cartographer to choose and simplify the map display in order provide a more clear and legible cartographic message.

On the other hand Andrienko (2003), adapt the readings levels of Bertin's, the approach of Kraak and Koussoulakou, and the identification and comparison components by Blok (2000) in order to define a operational task topology framework to review techniques and tools for spatio-temporal data, (see Fig. 2.8.). In this framework, the main objective is to concentrate on exploratory tools and task. For that reason, a conversion of readings levels into a search levels is given, resulting in four categories:

- **elementary "when" and elementary "what + where"**: describe characteristics of this object (locations) at the given moments.
- **elementary "when" and general "what + where"**: describe the situations at the given moment.
- **general "when" and elementary "what + where"**: describe the dynamics of characteristics of this object (at this location) over time.
- **general "when" and general "what + where"**: describe the evolution of the overall situation over time.

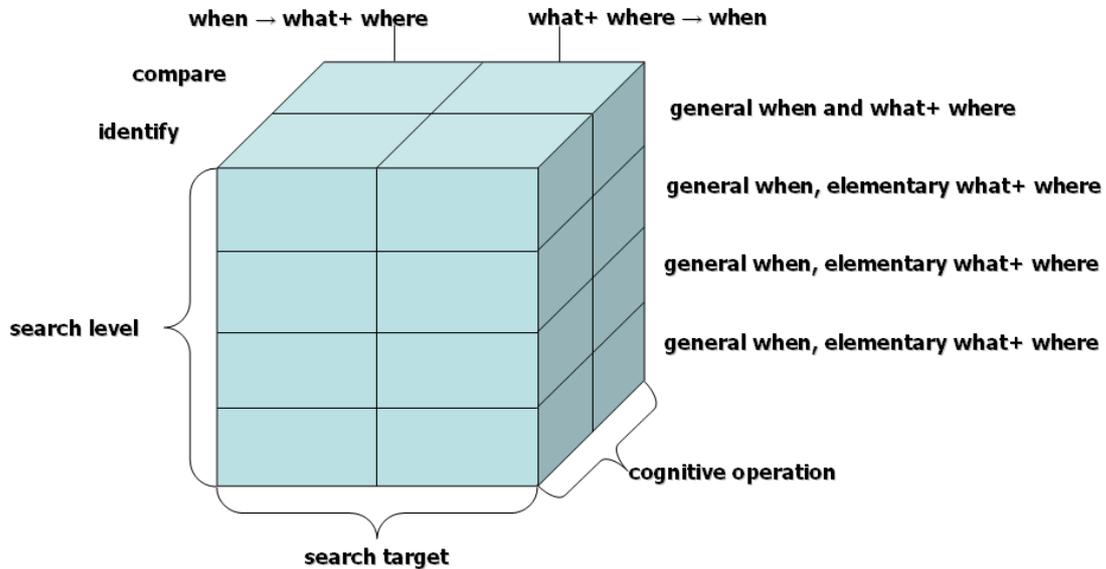


Fig. 2-8. The operational task topology, (Andrienko et al, 2003).

A detailed description of tools and techniques for spatio-temporal data is discussed by Andrienko, (2003). Is not the intention here to rephrase what they have reviewed, but better to say that a compressive framework has been proposed that helps software designers to target the dynamics, options, interfaces and suitable tools that an exploratory software environment must posses in order to be effective. This, in conjunction with a clear understanding of the nature of the data, the spatio-temporal phenomena, user community, and the desire exploratory levels.

2.1.2.2. Spatio-temporal Maps

The following section introduces the most used cartographic approaches to deal with spatio-temporal data. Their ability to depict changes in temporal domain is discus in order to detect the inefficiencies and state the need for more sophisticated methods.

2.1.2.2.1. Cartographic Time Mapping

As we mention before TG scientist in order to study spatio-temporal process much attention is required regarding the basic notions of space and time. Likewise an intelligent choice concerning the cartographic design and the graphic variables will provide the necessarily legibility degree for the map to be understandable. These temporal cartographic depiction models are discussed by Kraak 2003:

Single Static maps: Specific graphic variables and symbols are used to indicate change or represent an event.



Fig. 2-9. Color tints representing ages of built-up areas (De By, 2004).

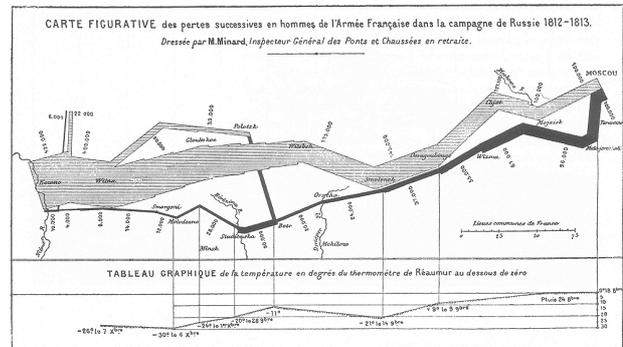


Fig. 2-10. Napoleon Russian campaign - 1812-1813(Kraak, 2003).

Fig. 2.9, shows the urban growth of the city of Maastricht, The Netherland. The application of color tints shows the new developed areas (light color) vs. the old built-up areas (dark color). It can be seen that a homogenous expansion takes place moving way for the city center. But in the case that a new development takes place inside the old built-up area an overlapping problem surge. Perhaps one of the most popular examples of a static map is the Napoleon Russian campaign created by Charles Joseph Minard, (see Fig. 2.10.). The combination of multiple variables and symbols to depict the changes (troops loses, temperature, location, date, battles, and direction) make it one of the kind static map.

Advantages:

- 1) Easy for portability, navigation and contemplation (G. Langran 1992).
- 2) Easy to understand the changes if proper graphic variables are used.
- 3) Additional information can be include I order to clarify the visual changes.

Disadvantages:

- 1) Occlusion depicting changes in time and space can occur.
- 2) No means of manipulation.

Series of Static maps: A single map in the series represents a snapshot in time. Together, the maps depict a process of change. Change is perceived by the succession of individual maps depicting the situation in successive snapshots.



Fig. 2-11. Sequence of maps depicting the urban growth: Maastricht city, The Netherlands (Principles of Geographic Information Systems book).

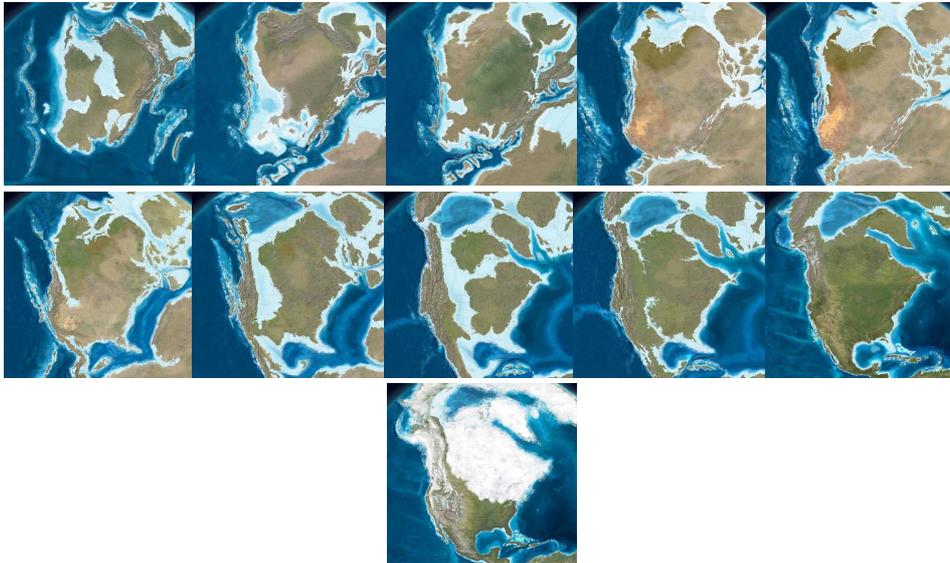


Fig. 2-12. Map series representing the “paleogeography” of North America over the last 550 million years of geologic history (Ron Blakey).

Fig. 2.11, shows a moderate time series sequence maps. In order to understand the map series the user has to follow a determine order. Change in temporal domain can be represented as shape, color, or any suitable visual variable. Conversely, the saturation of a series of maps displays make it's difficult to the human eye to follow (see Fig. 2.12). The user can present confusion at the time to identify spatio-temporal changes.

Advantages:

- 1) Good to avoid overlapping between images.
- 2) The images represent an exact time period allowing comparisons between them.

Disadvantages:

- 1) Depicting changes are limited to the number of images.

Animated maps: Change is perceived to happen in a single image by displaying several snapshots after each other just like a video cut with successive frames. The difference with the series of maps is the variation is deduced not from a spatial sequence but from real change in the image itself.

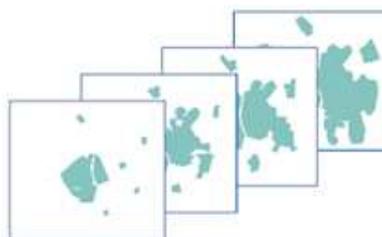


Fig. 2-13. Simulation of an animated map (Principles of Geographic Information Systems book).

Cartographic animation has been implementing since 1960 with the non-digital cartoon maps. During 1980s technological developments gave rise to a second phase of cartographic animation, with first computer-produced imagery (Kraak and Ormeling, 2003). Now days the large amount of mapping software (Arc.GIS, Erdas, etc.) allows the user to create easy going animated maps with a vast number of tools (blinking, panning, zooming, etc.) and functionalities (forward, backwards, fast, slow, stop, pause, etc.).

Advantages:

- 1) Good representing multifaceted process.
- 2) Stimulate user’s visual feeling process (patterns, trends and relationships).
- 3) Good performance in representing a real case simulated process.

Disadvantages:

- 1) Require more computation power and map preparation.

Below, a schema involving the perception and transcription of the geometric, thematic and temporal components of spatio-temporal maps by Koussoulakou and Kraak.

			Component		
			geometric	thematic	temporal
Map display	static	single	plane	visual variable	visual variable
		series	plane	visual variable	spatial deduction
	animated	plane	visual variable	memory deduction	

Table 2-1. A table graph describing the relation of maps displays, and its components (Koussoulakou and Kraak, 1992).

The combination of static, series and animated maps give the user the opportunities to explore and gain more insight into the spatial process under scrutiny. But their limitations pose a great challenge for more complex spatio-temporal problems. Additional methods are needed in order to overcome such difficulties. In the next section a more elaborated technique is discussed.

2.1.2.3. The Space-Time Cube

As mentioned before Hägerstrand's model has served as a framework in which (TG) built its roots. Taking advantage of the most important geographic elements of the model (STC) (see Fig. 2.3) and computer software, TG scientists can now recreate in a digital environment of the cube reference system in a more dynamic and complex way in order to investigate the spatio-temporal phenomena under scrutiny.

In the late of 1960 in "What about people in regional science?" Torsten Hägerstrand introduces a 3D method the "space time aquarium" to represent the people in space and time. In his model space and time are seen as inseparable. It could be said that a relative view of time is adopted as a framework of the spatio-temporal process under study. The basic environment of the STC to represent spatio-temporal data is based on the position of the temporal dimension (e.g., years, seasons, periods, or any kind of time resolution, (see Fig. 2.1) into cube heights (Z-axis) and the base of the cube (X, Y-axis) depicting the two dimensional geographic coordinates. STC's main objective is to visualize the individual movement across the landscape. For that, series of line segments belonging to each individual (or object) depicting the trajectory movement is graphically drawn in to the cube in order to visualize different locations at different times. This can be seen in the traditional case of the Napoleon's campaign in Russia discussed by Kraak 2003 (see Fig. 2.14). In here points represent places where battles were fought, steepness of the lines represent troops' movements, path thickness corresponds to the number of troops, and time of the events defines the cube height.

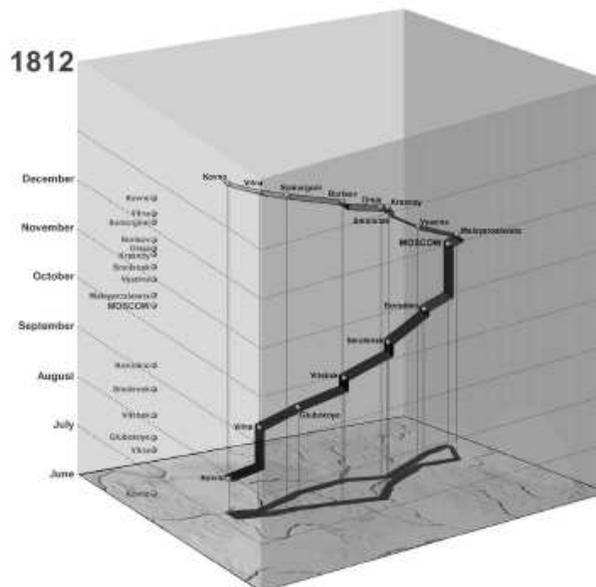


Fig. 2-14. Space-Time Cube: Napoleon's Russian campaign (Kraak, 2003).

One of the main properties of the STC is the possibility to answer the three basic core questions (what, where, and when) of spatio-temporal phenomena (see Fig. 2.3 for details). For example, the question "what?" is related to objects or events (e.g., people, earthquakes, etc.) and usually are represented by any kind of graphic variable that satisfies the case, (see Fig. 2.15). The "where?" question is related to locations occupied by the objects, usually aided by using a base map as a reference system for geographic coordinates and the cube's vertical reference lines, (see Fig. 2.16). Regarding the question "when?" the cube height is usually preferred to depict the time domain.

Stations and colors in conjunction with the cube horizontal reference lines in most of the cases is the “best choice” to explore time issues, (see Fig. 2.17).

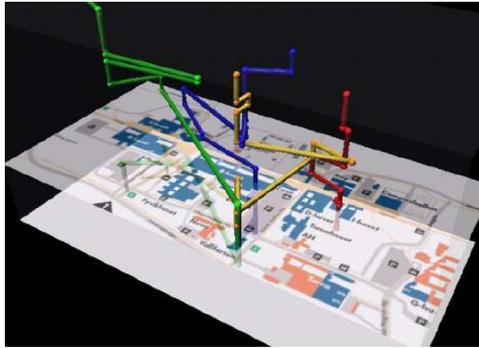


Fig. 2-15. Human walkthrough visualization in Valla Campus: Linköping University. Different colors represent different persons (Kristensson, et al. 2007).

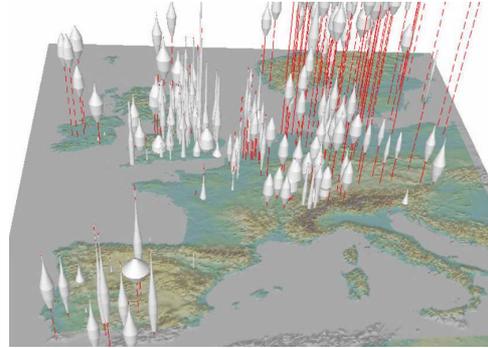


Fig. 2-16. The Black Death epidemic case: a base map of Europe serving as geographic reference for the disease locations (Kraak and Madzudzo, 2005).

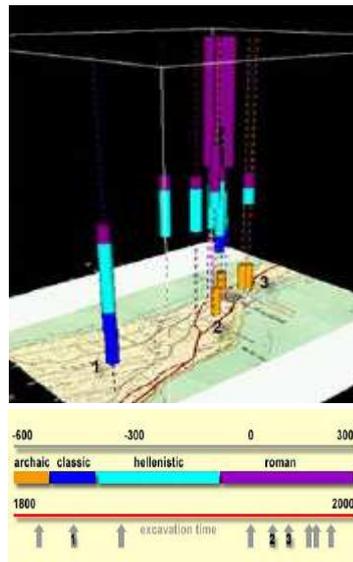


Fig. 2-17. Different time lines representing archaeological periods (Kraak, 2005).

As mention before the STC environment is base on a representation model for tracing human movements around the landscape. In that regard additional functionalities’ are included like:

- **Space-Time-Paths (STP):** lines depicting trajectory of an event, also called “lifelines”.
- **Space-Time-Prism (STPrism):** represent locations that can reach by a particular person in a particular time interval.
- **Potential Path Space (PPS):** the extent radius of the STPrism.
- **Potential Path Area (PPA):** the foot prints of the PPS into the geographic plane.

In order to fulfil a more realistic representation, the STC environment allows the constrains of such paths based on the original model by Hägerstrand, 1970. “One can distinguish between capability constraints (for instance mode of transport and need for sleep), coupling constraints (for instance being at work or at the sports club), and authority constraints (for instance accessibility of buildings or parks in space and time)” (Kraak, 2003). Others functionalities apart from the traditional ones can be found in all the previous surveyed literature and STC prototype software developed by ITC University. These are: cube manipulation (rotation, zooming, panning), highlight elements, multiple linked views, query functions, activation and deactivation of layers, footprint movement parallel to the z-axis, and the extrusion of surface base maps. As these functionalities posses the power to answer many questions related spatio-temporal phenomena, new problems arise when trying to deal with other kind of temporal data that must be visualize and explore in a different way. For that reason, new and old function must be taking into a consideration as well as the posed questions in order to (re)design appropriated tools for a suitable data exploration.

Advantages:

- 1) Time, location, and object domain are intertwining for the display.
- 2) Good representation of individual moving objects.
- 3) Extremely flexible to combine with others linked views to explore the data.
- 4) Excellent for visualization purposes.

Disadvantages:

- 1) Limited in the number of graphic variables for represent objects.
- 2) Conflict representing 3D heights and the cube height for the time component.
- 3) Spatial analysis limitation.
- 4) Almost impossible to visualize multivariable attributes without the aid of linked views.

Till now, I have disuse how changes in temporal and spatial domain can be perceived; what notions (absolute or relative) of time, researchers tend to adopt, and the surge of Time Geography to satisfy the need for exploring spatio-temporal phenomena. I also mention how representational view and frameworks, and Spatio-Temporal Maps are used to analyses, explore and disseminate such geographic phenomena. Is also stated, how new technological developments gives space to a more sophisticated software like the STC. On one hand, that all of the above (in some extent) compounds the main core knowledge needed for spatio-temporal problems related. On the other hand, too many limitations still exist when such phenomena involve multiple attributes (e.g. multivariate data). The next's sections will touch some of the techniques and applications that satisfy those needs.

2.1.2.4. Multivariate data Visualizations

Now days, high volumes of data that consist of many attributes are accessible to the public. The users' usually is interested in retrieving meaningful information hiding in the data and in a more compressive visualization methods. On the other hand, the complexity of such data sets demands more computational power, interaction, and flexibility. The new development in technology provides the support need to deal with such rich data sets. In here, information visualization and data mining process help the user to extract, explore and understand the data in hand. This section introduces

different visualization techniques for multivariate time dependent data focus on tools and functions capabilities.

2.1.2.4.1. Traditional Graphs

Simple graphs are usually preferred when individual tries to gain a quick glance about the data on hand. 1D, 2D and 3D graphs conform to the standards of most frequently used graphs in publications, surveys, and media. “For instance Points Graphs for point data, Line Graphs for continuous data, Bar Graphs for communicative data, and Circle Graph for cyclic data (Harris 1996)”, (Muller and Schumann, 2003), (see Fig. 2.18). Although they serve well in some cases, their limitation to represent multidimensional, time series and the combination of both data types make them unsuitable for a more complicated spatio-temporal phenomenon.

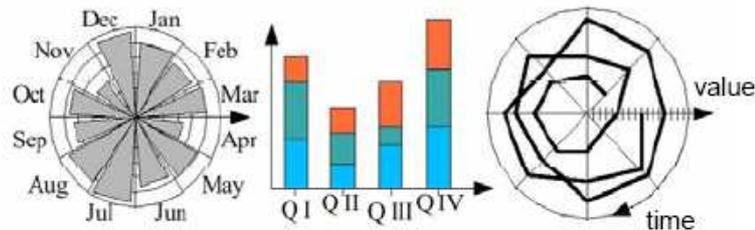


Fig. 2-18. Graphs for time-dependent data: left: Sector Graph, middle: Stacked Bar Chart, right: Circle Graph (after Harris, 1996), taken from (Muller and Schumann, 2003).

Advantages:

- 1) Easy to manipulate and generate.
- 2) Supported by an out numbers of software.
- 3) Easy going to understand.
- 4) Good for dealing with quantitative and qualitative data.

Disadvantages:

- 1) Limitation in the amount of data presented.
- 2) Visual overlapping using Stacked Bar chart.
- 3) Usually suitable for portraying 2 or 3 variables.

2.1.2.4.2. Parallel Coordinate Plot (PCP)

Another kind of graph representation that serves well for the depiction of multi-dimensional data into a single display is the Parallel Coordinate Plot (PCP). The PCP environment represents the geographic objects as a series of continuous lines intersecting a number of vertical lines. Each vertical line represents the respective variables belonging to the geographic objects. The maximum and minimum values of the variables are positioned at the top and bottom of the vertical lines and the horizontal lines cross the vertical units relative to the maximum and minimum values (Ormeling and Kraak, 2003) (see Fig. 2.19). “The result is a (possibly unique) multivariate signature for each observation, and a visual representation of relationships among many variables” (Edsall, 2003).

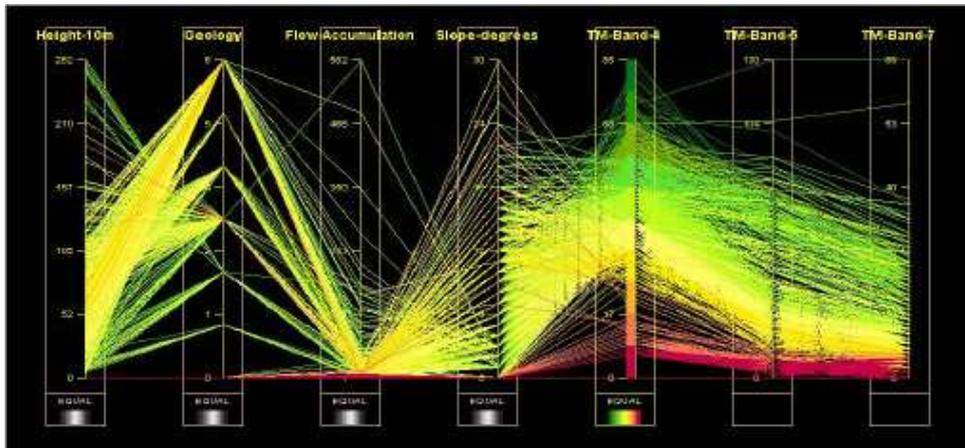


Fig. 2-19. Seven dimension PCP. Seven attributes (4 geographic and 3 spectral) are used to explore geographic phenomena, (Gahegan et al, 2001).

Since the introduction of the PCP by Inselberg in the 80's much effort has been put to enhance the capabilities of the non-interactive PCP into a one much more dynamical.

The user can highlight, select, filter, and coloring the variables in use. A discussion and implementation of the most used interactive tools can be found in Edsall, (1999, 2003), Kleim and Kriegel, (1996), and Andrienko and Andrienko (2004). Below some of the most interesting interactive capabilities are explained:

- **Brushing:** some kind of selection option in were objects are selected in one display and simultaneous highlight into others.
- **Strumming:** a brushing variant. Objects are “mousing over” highlighted.
- **Filtering or focusing:** reduce the plotting lines into selected ones to avoid saturation and clustering problems. The user can center the attention into particulars subset lines.
- **Variable assignation:** the user can state what variables the PCP will present.
- **Classification and coloring:** the objects are colored according one or more attributes to facilitate the exploration.
- **Dragging and repositioning:** axis variables can be move freely with the help of the mouse.

Talking about the temporal aspect, PCP can represent time depend data in few ways. On one hand, similar time terms can be positioned along the vertical axis and subsequently others variables related to it. The problem is that in order to perceive changes in time the total amount of related variables must be repeated for each time term. This cause a quick overload (vertical axis) and confusion to follow lines (PCP horizontal extent exaggeration) and detect patterns. On the other hand, time can be positioned in one of the vertical axis (preferably the first) and treaded as a (ordinal-level) classified variable in where additional variables can be explore according their appropriate time interval. The main problems is, that is impossible to detect changes in time for a single objects unless the same object is repeated for a given time interval. If that so the case, the extend display of the PCP will be saturated with multiples lines. However some dynamics tools have been develop to deal with temporal aspects. For example: temporal brushing and focusing. In the first case, time lines are rendering for an appropriate graphical representation. The second case focuses on presenting a detailed but static view for the selected time series (Monmonier, 1990) taken from (Muller and Schuman, 2003).

Advantages:

- 1) Can be linked to maps and others graphics displays.
- 2) Support both spatial and temporal domains for spatio-temporal phenomena.
- 3) Good for representing multivariate data.
- 4) Patterns can be detected if a limited number of attributes in conjunction with a moderate data set are inputted.
- 5) Support any kind of data type e.g. nominal, ordinal, interval and ratio.

Disadvantages:

- 1) The input of a huge geographic database overloads the PCP environment. Causing occlusion and difficulties to extract the information.
- 2) The order of the axis depicting the variables in use influences the amount and quality of the insights.
- 3) The space dimension is difficult to depict if is not reduce into one dimension (Edsall, 2003).
- 4) Temporal domain is difficult to represent. Usually requires the duplication of the entire variables according their specific time term.

2.1.2.4.3. Chernoff Faces

In 1973 Herman Chernoff design a new method of visualization the Chernoff faces, named after him. In this technique, a simplified image of a human face is used to display the data variables in use. Different facial features, such as the length of the nose, the curvature of the mouth, the size of the eye, the shape of the head, etc. are assigned to the data variables in order to perceive patterns. The reasons behind the implementation of the faces for data analysis are the universal communication signals, and the easy understandable type of object that they form.

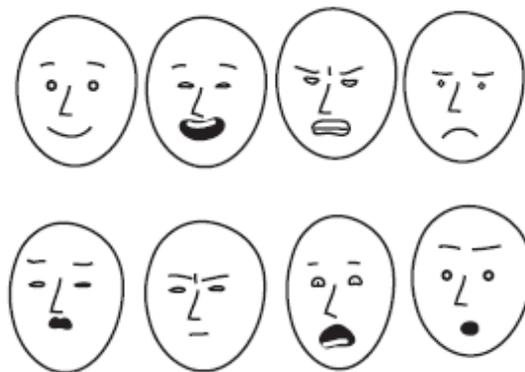


Fig. 2-20. Eight universal expressions: happiness, elation, anger, sadness, disgust, determination, fear, surprise (Ekman and Friesen, 1975; Ekman, 2003).



Fig. 2-21. Chernoff faces. Different data variables are mapped to the sizes and shapes of different facial features (Ware, C., 2004).

Advantages:

- 1) Good to represent patterns (using shape, size, eyebrows, and mouth).
- 2) The communication message is intuitive (according to feeling status of the faces).

Disadvantages:

- 1) Faces saturation of the around the mapping area can cause a first wrong impression to the user.
- 2) In order to deal with the temporal component faces features must be mixed with animation or series of maps display technique.

2.1.2.4.4. Lexis Pencil

The Lexis Pencil technique represents events histories as a pencil-like multi-facet objects in three dimensional spaces (Francis and Pritchard, 1999). The idea of the 3D Lexis Pencil comes as an attempt to improve the visual effects and capabilities of the concept developed by W. Lexis (1875), The Lexis Diagram (see Fig. 2.22). The diagram itself was considered an effective tool for demographic research studies, because of the ability to deal with a simple and legible time graph representation. But as much more complex event history data is introduced, the practicalities of using a lexis pencil in a dynamic environment become problematic.

In the paper “Visualization of Events Histories” by Francis and Fuller (1996), a review of proposed ideas to extend the Lexis Diagram are discussed. Is not the intention here to discuss each idea in detail, rather, the chapter will move on to introduce the actual 3D Lexis Pencil visual and dynamic environment.

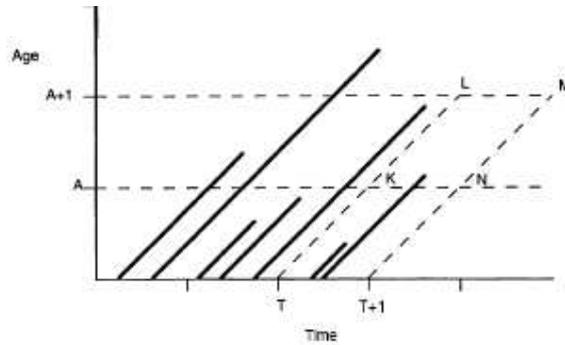


Fig. 2-22. A classical Lexis Diagram. The x-axis depicting calendar time and y-axis representing age or time occupied by the events (solid sloping line) (Pressat, 1961).

The purpose of the 3D Lexis Pencil does not differ from the original model, which was the study of events in times as mention before. The pencil is presented as a 3D solid object, positioned into a geographic plane. The location domain is characterized by the tip of the pencil pin-pointing the exact location using the x and y-axis of the reference plane. The time domain is represented using the height of the pencil depicting the total time of occurrence of the individual events. Regarding the attribute domain, “the pencil maps different time-dependant variables onto a different faces of a pencil, starting with the top of the pencil” (Muller and Schumann, 2003) (see Fig. 2.23).

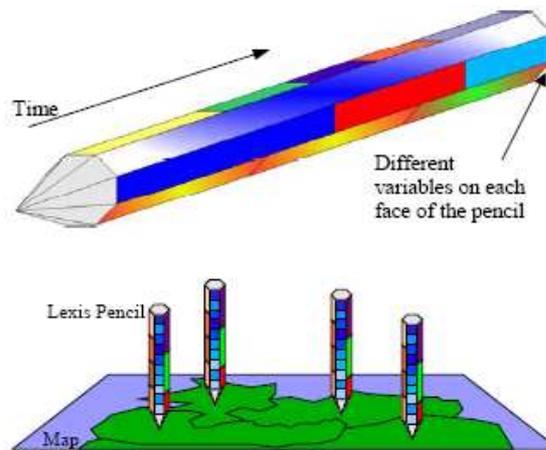


Fig. 2-23. The 3D Lexis Pencil. Each pencil stand for an object or events, the pencil height describes time and colors variables (Muller and Schumann, 2003).

Now, in order to understand the changes in time the user has to assign each desire variables on the face of the pencil and applied the graphic variables (Bertin, 1983) that best suit the case. “Each part of the rectilinear construction is thus proportional to the amount of time the individual spends in that state before moving to a different state. With many time-dependent variables, we can produce a planar multi-faceted image, consisting of multiple rectilinear components side by side, or by multiple concentric rings” (Francis and Pritchard, 1998). To summarize, the user can track changes in attribute domain following the color changes across the length of the pencil, changes in time domain comparing the differences in pencil heights and the lengths of the color variables boxes, and changes in location domain manipulating the reference plane for events retrieval. Some of the dynamic tools

that possess the 3D Lexis Pencil are: rotation, zooming and panning, flight trough, view point fixed, focus, interactive enlargement, clipping, cropping, color redefinition and lighting, and spinning. Regarding the exploration capabilities are: identify, variable assignation, selection and bruising, and linked views. Below, a description of the un-reviewed tools concepts.

- **Clipping:** works as a form of a clipping plane, in where the events are displayed above or below the intersected plane. (similar as STC in ITC uDig prototype.)
- **Spinning:** the ability to rotate the pencil around their own axis.
- **Cropping:** a variant of clipping. A user defined 3D box in where only events within the box are displayed.
- **Flight trough:** interactive navigation around the scene.
- **View point fixed:** the display is anchored into a steady position in (x, y, and z) coordinate, using as a center of projection the selected objects. The display rotated around the events without losing sight.
- **Interactive enlargement:** enlarge and resize the pencil height freely in order to explore the faces variables. Usually become handy when multi-faceted faces of the pencils are saturated.

Advantages:

- 1) Continuous and discrete variables can be combined.
- 2) Support the spatial, time and attribute domains.
- 3) Good representation for time-dependent data.

Disadvantages:

- 1) The use of large numbers of variables can overload the multi-colored faces of the pencil, and can cause confusion to detect patterns.
- 2) Occlusion problems may come out, if many events are presented.

2.1.2.4.5. Circular Axes-Base Visualizations

Another visualization technique for handling multivariate time series data has been proposed by (Tominski, et al., 2003). Similar as the PCP and some of the Graphic Drawings techniques, the TimeWheel and MultiComb diagrams present a new interactive environment to examine time-dependant data. Both techniques use time domain as independent variable, and the attribute domain as dependent variables. As they belong to the category of axes representation retaining the real spatial locations of the events was not the main purpose. That does imply a loss of an important dimension if they are not represented as separate attributes (x, y coordinates). In this approach axes represent the minimum and maximum values of the assigned variable and are displayed in bar form.

In the case of the TimeWheel visualization the time axis is positioned at the center of the display. Additional variables are circular arranged around the time variable (see Fig. 2.24). Colored lines segments portray connectivity between attributes values (dependent variables) and the corresponding time (independent variable) values. In order to avoid occlusions, achieve straightness of the lines and variable focusing the following options are available:

- **Color fading:** lines segments are automatically faint colored in order to redirect the attention of the variables in use.

- **Interactive Rotation:** the user can positioned the variables perpendicular to the time variable in order to reduce the angle of the connections for an easy examination.
- **Length and Position Adjustment:** lines connections that belong to the un-used variables are diminished in length and located at the background of the two perpendicular variables (according the time variable).

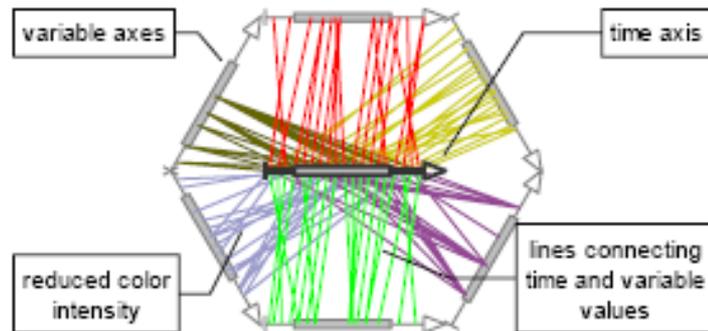


Fig. 2-24. The 2D TimeWheel compounded by six variables (circle set) and time variable (central placement) (Tominsk, et al, 2004).

On the other hand the MultiComb uses a similar approach but using time plots to represent the relation between the time variable and attributes variables. The idea is to positioned the plots of the different the dependent variables circular around the independent variable same as the TimeWheel case. The signal can be placed parallel to the faces of the variable in use or extended outward for the center display (Tominsk, 2003), (see Fig. 2.25). One of the main advantages of the MultiComb is that posses almost the same functionalities of the TimeWheel plus depict a spike glyph in the center of the display to show additional information.

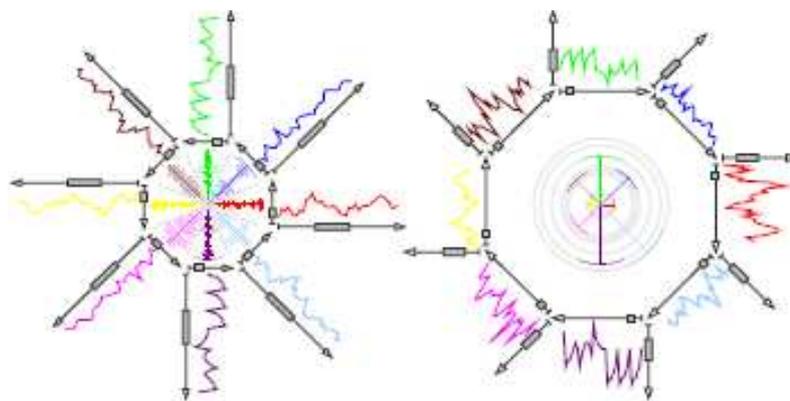


Fig. 2-25. Two variants of the 2D MultiComb (Tominsk, et al, 2003).

Extensions of the two dimensional to three dimensional graphic visualization has been proposed, (see Fig. 2.26). But perhaps a more compressive real case implementation is needed in order to address the effectiveness of such. The main advantage in using a 3D version of both cases ranges on the ability to navigate freely the graph; the 3D axes provide a more legible visualization environment to avoid lines intersections, and the dynamic cocktail standards tools (e.g., zooming, panning, focusing, brushing, etc.) for visual exploration that any 3D environment should support.

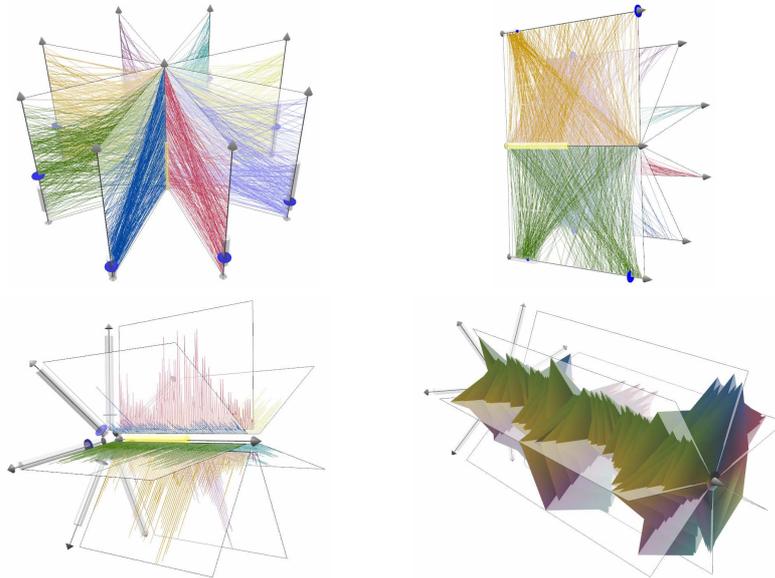


Fig. 2-26. Variants of the 3D Axes-Base representations. Upper left: 3D Time Wheel, Upper Right: TimeWheel “folder style”, Down Left: 3D MultiComb, and Down Right: 3D Kivat tube, (Tominsk, et al, 2003).

Advantages:

- 1) Support ordinal, nominal, discrete, and continuous data.
- 2) Allow interactive exploration.
- 3) Design specifically for time-dependent and multivariate data.

Disadvantages:

- 1) The level of interpretation rest in the amount of data and variables in use (overload easily).
- 2) Dynamics are strictly needed.
- 3) Limited in the amount of time-dependent variables.

2.1.3. Methods, Techniques, and Concepts of Graphic Representations

2.1.3.1. Introduction

Often people are interested in visualizing information considering a relationship between objects. For example, a person must be interested in what train passes certain stations? Or in the case of a student, what relationship does his/her university possess with other universities for possible internships? In this case, we can easily display the information for the user in the form of graph. "A graph contains nodes and edges where the nodes represent the items or concepts of interest, and the edges connect two nodes together according to some associational scheme" (P. Rogers, 2005) . Looking into the graph the user can easily explore the information, see cluster and discover patterns hiding in the data set. Automatic process now allows the user to visualize complex graphs faster and in a number of ways. In the following sections a discussion of the most interesting graphic drawings techniques are presented.

2.1.3.2. Readability parameters

In order for a graph to present an effective communication message certain rules must be followed. Similar to cartographic mapping techniques conditions and criteria are taken to improve the visual look of the representation. In graphic drawing technique this is referred as readability or aesthetics issues, that is "the capability of conveying the meaning of the diagram quickly and clearly". The Readability of the graph mostly depends on the graphic drawing method adopted and the data set in use. But, usually (as all are connections between nodes and edges) certain standards are adopted. For example, edge straightness, minimizing crossing edges, geometrical symmetry of the nodes, nodes size and shape, length of the edges, symbol and colour consistency, and edge separation. These standards as mentioned before makes the graphic drawing representation more suitable for interpretations, but it is worthy to say that additional rules variants' are applied according to users' preferences and the desired message. The next sections introduce some of the most widely used graphic drawings techniques.

2.1.3.3. Hierarchical Drawing Methods

"A hierarchical drawing of an acyclic digraph is an upward polyline drawing where the vertices and bends are constrained to lie on a set of equally spaced horizontal lines, called layers. In some applications the assignment of vertices to layers is given, e.g., by the semantics of the graph. Such graphs are called layered digraphs, or hierarchies"

(Di Battista, G. D et al., 1994)

Hierarchical drawings usually are presented as graphs compounded by nodes (points) connected by edges (lines) that follow some kind of ordering or ranking for the nodes in a fixed coordinate system. The nodes are positioned in layers. Root nodes are usually in the top layer and consequently low level nodes are further down. When a node does not resemble any additional connectivity is called leaf nodes and are positioned at the bottom of the graph. Dummy nodes frequently are used to aid the drawing process. They serve as ghost nodes to help to reduce the number of edge crossings and edge bend points. An efficient hierarchical graph should be able to communicate the message

intended or provide the necessarily information to explore the data in hand. The efficiency of the graph has to fulfill the following requirements according to Sugiyama et al., 1981.

- **Element A:** “Hierarchical” layout of vertices.
- **Element B:** “Less-crossings” of lines
- **Element C:** “Straightness” of lines
 - **Element C1:** “Straightness” of one-span edges.
 - **Element C2:** “Straightness” of long span edges.
- **Element D:** “Close” layout of vertices connected to each other.
- **Element E:** “Balance” layout of lines coming into or going from vertex.

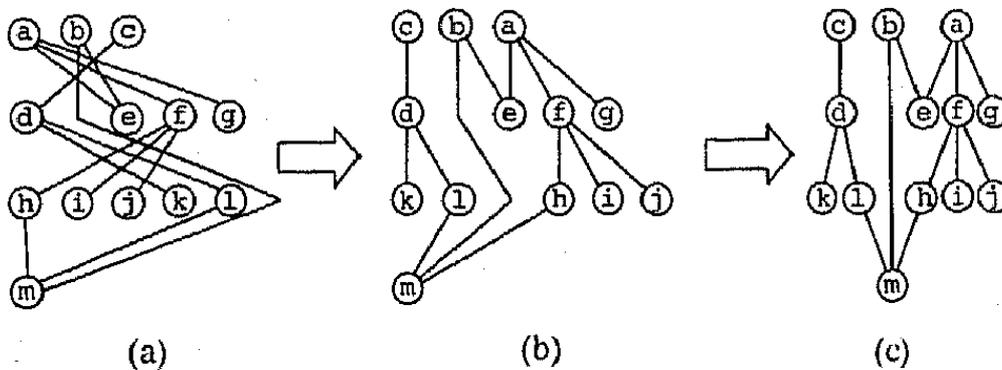


Fig. 2-27. Example of improvements of a hierarchical drawing. (Sugiyama et al., 1981).

- (a) Nodes are placed in layers (*element a*).
- (b) Nodes are rearranges to reduce edges crossing (*element a*) → (*element b*).
- (c) Horizontal position of vertices (*element a*) → (*element b*) → → (*element c + element d + element e*).

Advantages:

- 1) The hierarchy structure suite well for complex diagrams.
- 2) Easy to understand if the basic readability standards are followed.

Disadvantages:

- 1) Does not work well with dense graphs.
- 2) Other algorithms are needed to improve the aesthetics of the graph.
- 3) Problems arise with a limited display space.
- 4) Node movement is needed.

2.1.3.4. Planar Layout Techniques

A planar graph is a graph that can be sketch with no edge crossing. Conversely a nonplanar graph is the one that posse edges crossing. Usually planar graphs are preferably desirable because their aesthetic looks make it easy to understand. A common strategy to convert nonplanar to planar graph is to planarize the graph and then applied a planar drawing algorithm. The term planarization refer to the

transformation (*nonplanar* \rightarrow *planar*) and additional operations (*edge deletion*, *NP-hard and splitting*) to simplify the graph. After planarization, a planar drawing algorithm can be applied, to enhance the aesthetic look of the graph. An out number of graph drawings algorithms exist to convert nonplanar to planar graphs. Some examples are: spring embedded, force directed, spring embedded (SA).

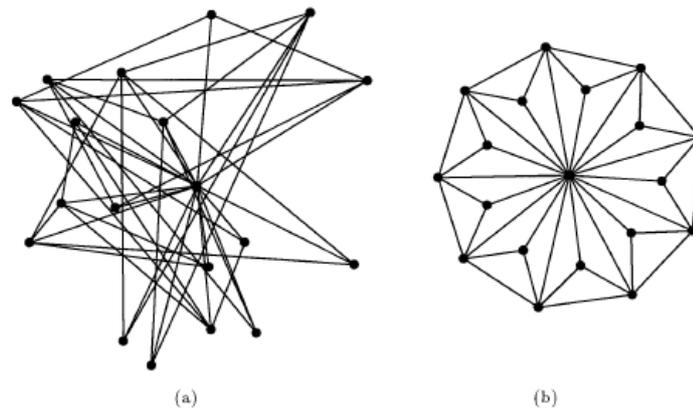


Fig. 2-28. Example graph implementing the SA algorithm base on user define aesthetics (R. Davison and D. Harel, 1993). (a) Input nonplanar graph. (b) Output planar graph map.

Advantages:

- 1) Suitable to make the first modifications (*nonplanar* \rightarrow *planar*) of the drawing graph.
- 2) Quick performance to achieve a nonplanar graph (not always a good looking graph is obtained).

Disadvantages:

- 1) Other algorithms are needed to enhance the graph.
- 2) Node movement is often necessary.
- 3) Optimal geometric symmetry is needed.

2.1.3.5. Force Directed Layout Techniques

Force directed methods as mention before, are often used techniques for graph drawing mapping. They foci on graph aesthetics make them attractive and adaptable to any kind of data for graph representation. Their easy implementation often gives the users satisfactory results. The nodes are positioned in a two dimensional graph in where the minimum edges crossing and length are achieved. The algorithm assigns forces to the nodes and edges. The nodes acts as charged particles that repel all others and the edges are treaded as springs that attract connected nodes (Eades, 1984). The algorithm works recursively until the minimum state of energy is achieved according to specifics readability constrains. Some examples are:

(a) *The barycenter approach (Tutte, 1993)*: “the nodes are drawn at the average distance of their neighbors (witch is known as their barycenter). This method is repeated for a set number of iterations, or until nodes stop moving, having reached a minimal energy state. A subset of nodes must be fixed.

To indicate the outer borders of the drawing space, or the nodes will simply collapse to a single point” (P. Rogers, 2005).

(b) *Spring embedded (Eades, 1984)*: is a heuristic based on a physical model. The straight-line standard is adopted. The drawing process is to simulate a mechanical system, where vertices are replaced by rings, and edges are replaced by springs. The springs attract the rings if they are too far apart, and repel them if they are too close (Di Battista, G. D et al., 1994).

(c) *Spring embedded (SA) (Davison and Harel, 1993)*: an energy function is defined in terms of the desired aesthetics: for instance, the number of edge crossings plus a measure of the closeness of vertices. A layout of minimal energy (and thus maximal beauty according to the energy function) is obtained by Simulated Annealing (SA) (Di Battista, G. D et al., 1994).

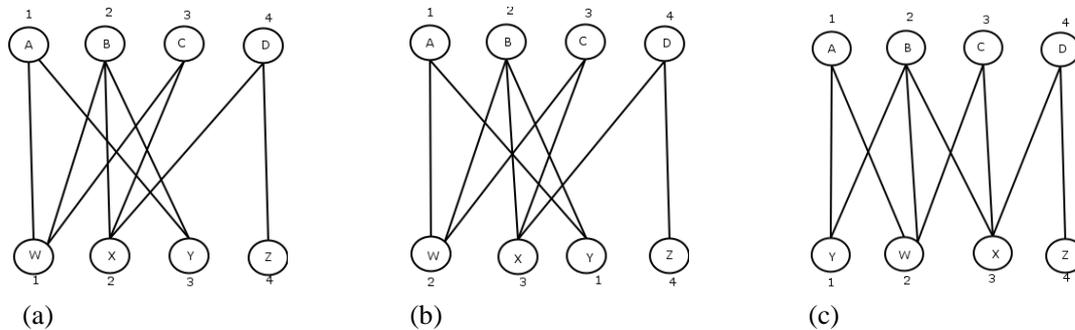


Fig. 2-29. Example graph implementing barycenter approach (W. Ahmada and A. Khokhar, 2007).

(a) Original Graph G. Crossing Number($C(G) = 9$)

(b) Updated Positions of Lower Layer Nodes using Barycenter Heuristic.

(c) Graph G' after one iteration of Barycenter Heuristic. $C(G') = 2$

Advantages:

- 1) Geometrical symmetry is achievable.
- 2) Good and clear graph drawing output.
- 3) Allows aesthetics rules flexibility according users preferences.

Disadvantages:

- 1) Computational cost of computing a force between every pair of nodes.
- 2) Inability to scale to handle large graphs.
- 3) Node movement is needed.

2.1.3.6. Optimizing Aesthetic Criteria

Optimizing Aesthetic criteria are methods that focus on the improvement of the aesthetic qualities of the graphs drawings. The main target is to allow the user to manipulate some visual criteria's to enhance the graphic drawing map. The algorithm for this process works almost automatically, once the user specifies a series of conditions the algorithm weighs them and organizes the graph in order to achieve the desired output. Is important to mention that as some constrains are considered others are affected. That implies that much iteration is required in order to reach a “good looking” graph layout. Multi-criteria optimizers like the SA (see above (c)) and the genetic algorithm are typical examples.

For the algorithm to work effectively the used criteria must be quantifiable. For example, “ number of edge crossings, statistics, relation to even node distribution; the area or aspect ratio of rectangle containing the graph; the angular resolution of edges emanating form the nodes; the number of edges blend; evenness of edge length; methods for measuring the symmetry in the graph; and encouraging directed edges in particular direction” (P. Rogers, 2005).

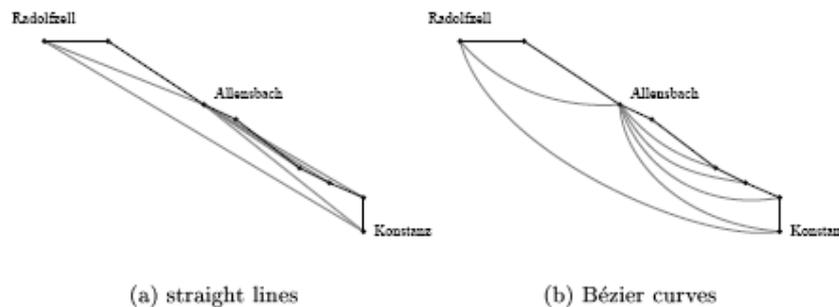


Fig. 2-30. Example graph implementing criteria base method (Ulrik and Dotothea, 1998).

- (a) Main train stations preserver as a straight lines.
- (b) Secondary train stations enhanced by Bezier curves preventing edge overlap.

Advantages:

- 1) Allows explicit improvements.
- 2) Permit multiple criteria optimizing techniques according to users’ preferences.
- 3) Good for fixed geographic locations.
- 4) Good for specific requirements (knowledge about the data).

Disadvantages:

- 1) Much iteration is required.
- 2) Takes time to reach reasonable drawing.

2.1.3.7. Dynamic Graph Drawing

Dynamic graphs drawings are usually graphs that present some changes in the data structure. A user may be interested in visualizing some kind data that is in constant change. For example, mapping some part of the internet network, a metadata for a library or business related associates for a corporation. In here the target is on preserving the user mental map of the graph. A metal map is the current view of an individual to retain the map in his internal memory. According to R. Loyd (2000), reference points, spatial simplification and cognitive distance, help the individual to capture the initial image of the map. These main characteristics can be translated to the main standards (focus nodes, hierarchy layers and evenness of the edges) of the graphic drawing maps in order to maintain the user mental graph. In Fig. 2.31 (left) a selected focus node A and his parent node B. Followed by Fig. 31 (right) in which a graphic transition animation has been applied to smoothly redraw the graph in order to minimize the disruption of the mental graph map.

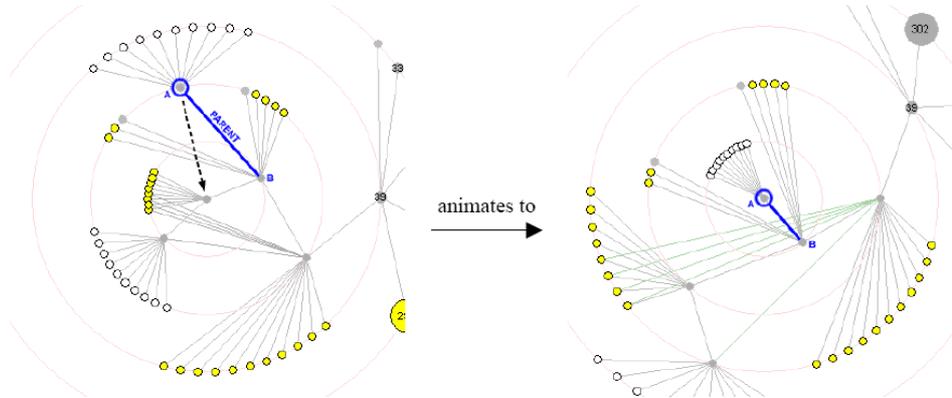


Fig. 2-31. Example of a Dynamic Hierarchical Radial Layout. Node A (left image) is selected to become the new focus. Node A's current parent is node B; the new layout (right image) is oriented so that the edge AB maintains the same direction (Yee et al., 2001).

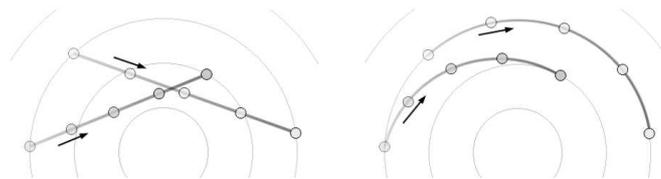


Fig. 2-32. Interpolation in rectangular coordinates (left) can yield a confusing animation. Interpolation in polar coordinates (right) works better for radial layouts (Yee et al., 2001).

Advantages:

- 1) Good to retain the mental map of the user.
- 2) Animations allow an easygoing track changes and avoids confusion.
- 3) Excellent at maintaining the context in order to explore the graph.

Disadvantages:

- 1) Limited for a smaller graphs.
- 2) Node movement always presents.
- 3) Not suitable for retaining geographic locations.

2.1.3.8. Drawing Graphs in 3D

As we see, different kind of information structures can be represented as graph maps. However as the technology and computer systems keep growing more functionality and computing power are available. As a response, more complex data structures has surge catching the attention developers interest to design more sophisticated graphic drawing methods. For example the addition of an extra dimension allow the designer to create complex three dimensional graph drawings in order to investigate the 3D visual capabilities to gain more insight about the data. In the case of the larger scale blocks of information, we can either learn by navigation the spatial layout of the map. As mentioned before, computers give us the opportunity to go beyond real-world navigation: In virtual 3D spaces we have the option of creating artificial navigation techniques (zoom, pan, rotation, translation, rapid zooming, non-spatial navigation, elision, hyperlinks, and 3D widgets. An important issue in drawing

3D graphs layout is the intuitive recognition of the user. In order to aid the navigation capabilities in such notion edges, colors, textures, elementary motions and depth should help to extract the basic information structure of the graph map. Is true that 2D and 3D graph guard too much in common, perhaps one is the consequence of the other. The unique characteristic of the 3D graph layout is the ability to deal with occlusion.

“Occlusion is the single most important depth cue, over-riding others and because of occlusion we can see far less information in the z-direction. Design of the interfaces implies taking into consideration viewpoint and 2D visual field, transparency and textures. To deal with this problem the implementation of navigations widgets must be considered” (C. Ware, 2000).

Usually these methods are variants of a two dimensional graph drawings: hierarchical layout, force directed graph, ext.

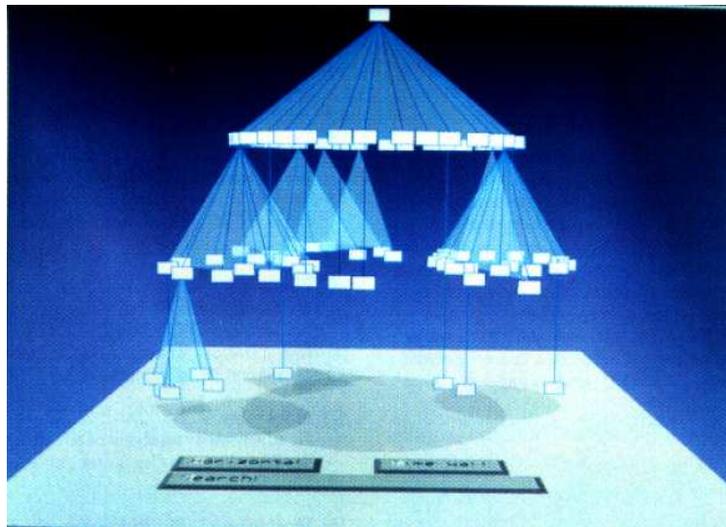


Fig. 2-33. Example of 3D Cone Tree graph (G. Robertson et al., 1991).

An interesting approach (Robertson et al., 1991) is the three dimensional cone tree graphs. A standard hierarchical structure was taking as the base model for the graph representations. The hierarchy is presented in 3D to maximize effective the use of available screen space and enabling the visualization of the whole structure. The force directed algorithm is used to maintain the evenness of the edges length and node separation. In the picture above, branches are arranged as series of circles projecting child nodes. Casting shadow at the plane of the map allows the user to identify clustered chunk of information. The implementation of the elision functionality allows the user to hide relative's nodes into the proper parent node to simplify the graphic drawing map. Others functionality like rotation, translation, highlight, and rapid zooming were implemented to aid the navigation and cognitive recognition of the user mental map. A similar approach is the graphical representation of a 3D hyperbolic space to visualize a structure section of the World Wide Web (Munzner and Burchard, 1995). Nodes represent documents on the web pages and lines are links of others related web pages.

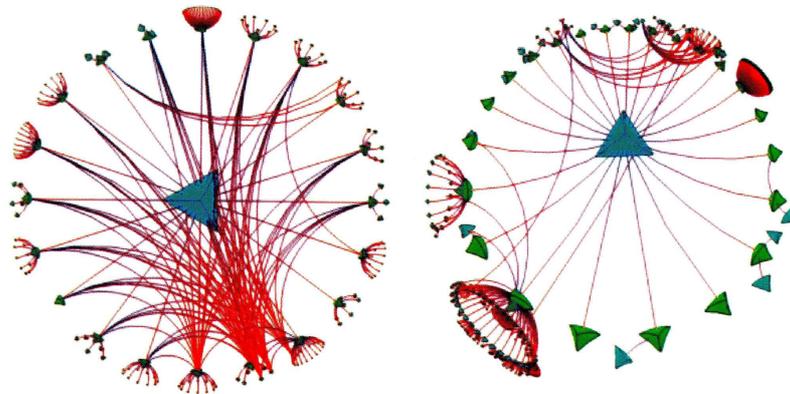


Fig. 2-34. Example of 3D Hyperbolic Space graph (Munzner and Burchard, 1995).

It can be said that the 3D representations of a graph require much more complex user's interfaces to allow interaction and accessibility of the information. And is not necessary convenient to do so if the data structure allows a simple representation in 2D fulfilling the basic graphic standards.

Advantages:

- 1) Maximization of the screen space.
- 2) Facilitate the visualization of the whole structure.
- 3) The ability to use deepness to display more information.
- 4) Shadows and light effects help the users to focus on data subsets.
- 5) Rich user interactivity.

Disadvantages:

- 1) Limited size of the graph (Computational power is require for a 3D display).
- 2) Others disadvantages depend on the standard model taken (e.g. planar, hierarchy, hyperbolic, etc.).

2.2. Summary

Base on what it has been discussed in this chapter; we can clearly say that many geovisualization and visualization environments possess particular similarities and differences. For example, the kind of data to be used, the way to present and disseminated the information, tool and functions and sometimes the environment itself. It also evident, that the advantages of some environments (geovisualization or visualization) can be satisfies by the limitations of another. In that regard, designer can take advantage of the technological developments and the portability benefits that some applications provide in order to combine and create a more robust application for data exploration. Many methods and techniques for exploring spatiotemporal and multivariate data are easily found now days. However these environments possess the necessarily tools and functions to manipulate and explore their respective data characteristics (e.g. time and multi-attributes), they still lack the ability to visualize more complex data (e.g. multivariate spatio-temporal data).

A model involving the STC environment and the Graphic Drawing techniques will be proposed in order to address some of the limitations that posses the STC to deal with multivariate spatio-temporal

data. A potential combination between these two environments (STC “geovisualization” and Graphic Drawing “visualization”) lays in the sense that both deals with lines and node to connect events. The STC uses lines and nodes to displays singulars events (compounded by multiples points) moving in time. On the other hand Graphics Drawings’ uses multiples nodes (points) to state the connections (lines) according to a semantic relationship (similarities in attributes). The next chapter will discuss in detail such combination.

3. Model design

3.1. Introduction

Based on the literature review stated above we can say that in order to design and select appropriate tools for exploring and visualizing spatio-temporal phenomena (see Fig. 3.1) the following issues must be considered:

- 1) Based on the literature review stated above we can say that in order to design and select appropriate tools for exploring and visualizing spatio-temporal phenomena (see Fig. 3.1) the following issues must be considered;
- 2) A comprehensive understanding behavior of the spatio-temporal phenomena will provide the necessary clues to state how should be represented;
- 3) The adoption of the representational view (vectors, raster, or both) in conjunction with the data sets will influence the available options for the visualization environment and suitable tools;
- 4) Embracing a time notion (absolute, relative, or both) and time resolution (granularities), will serve up a guide line to answer possible temporal questions;
- 5) The desired “search levels”, and tasks define the appropriate selection and combinations regarding the exploratory tools;
- 6) Given all of the above issues and the questions to be answered, special tools can be designed to deal with specific users needs if the data set supports this;
- 7) The levels of extraction (data mining) provide the user to pull out meaningful information or detect patterns behind the data.

The following sections discuss the most suitable visual environments’ and tools for exploring archaeological sites as well as others related data types.

3.2. Space-Time Cube

3.2.1. Justification of Space-Time Cube

On the main advantages of the STC is the dynamic environment to represent events in within a temporal dimension, and hence the basic components of spatio-temporal data (what, where, and when) in order to answer questions. In Fig. 3.2, objects or events are related to what, where is related to locations on the base maps (X, Y coordinates) aimed by the (X, Y –axis) cube reference lines, and the when (t-time) is related to the Z- axis cube reference lines and one of the faces of the cube.

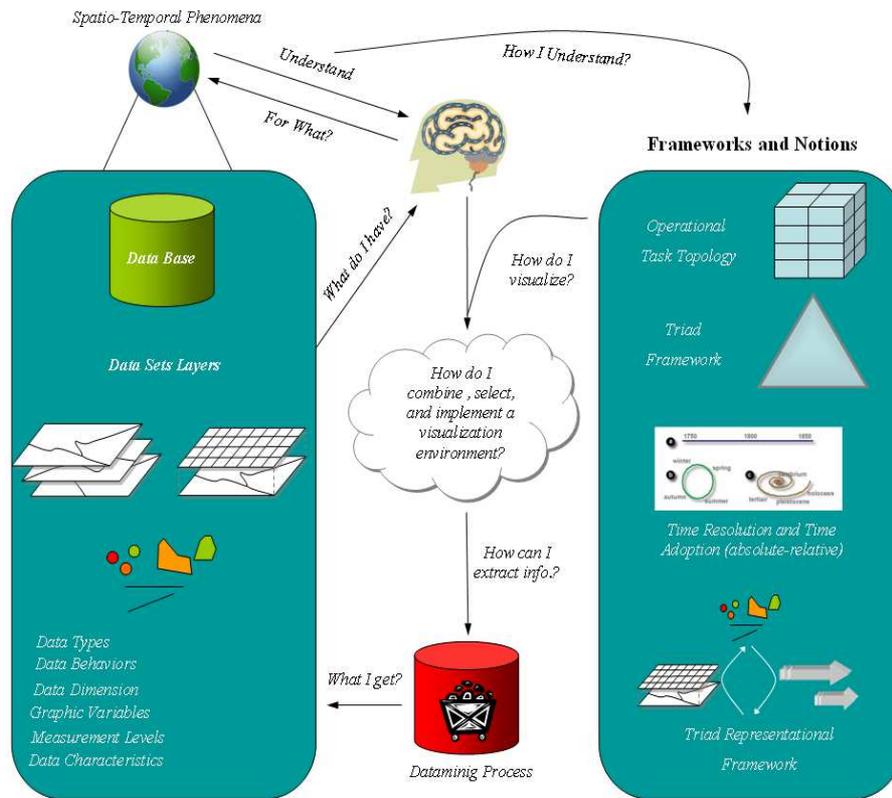


Fig. 3-1. Steps for a TGS to design a geovisualization environment and understand spatio-temporal phenomena (Source: Author).

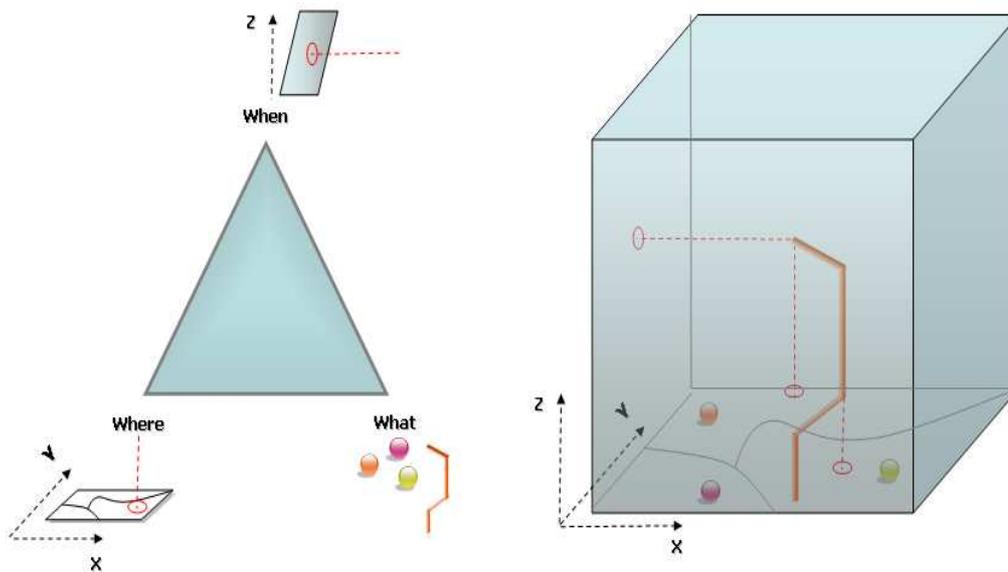


Fig. 3-2. The Relationship between the Triad Framework (Peuquet, 1994) and the STC environment (Source: Author).

- *when + where* → *what*

Regarding the question (*when + where* → *what*), the main interest is to describe sets of objects or particular objects in a place in time. In here the what question is strictly related to the where (locations). It can be said that a particular object exists in a particular place or a set of objects are preset in a set of locations. The case variate when we are talking about moving objects, a particular object is a function of space and time as its moves across the landscape. In here the whole idea of movement involves different locations (sets of locations) in a given time terms (sets of time). For a particular object a simple query function can retrieve the desired information. For example, (X= 234.54, Y= 324.76, T= 1970 yrs.) gives O = object. In reality this kind of simple query does not seem convenient to explore particular objects. The probability to query an object in such way is influenced more by the precision of the coordinate values than the desired time, unless the user knows the exact position, which is unlikely to be the case for most queries. In Fig. 3.3, others options to deal with objects are presented.

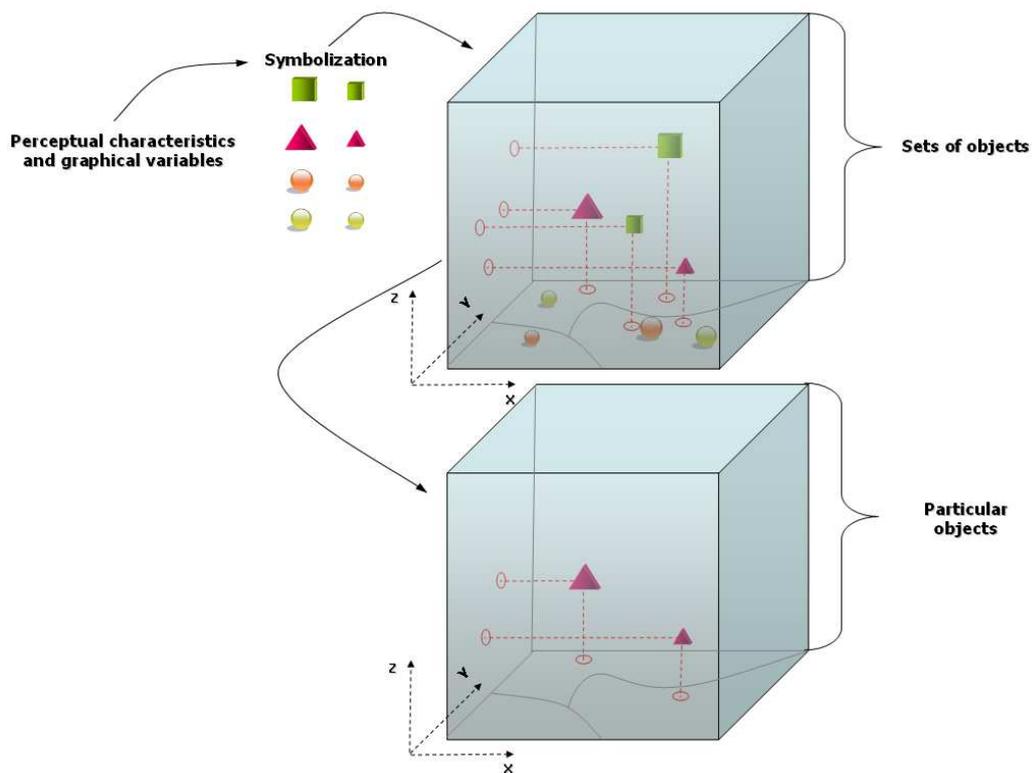


Fig. 3.3. The Relationship between STC, (Source: Author) and (*when + where* → *what*) Triad Framework, (Peuquet, 1994).

The advantage of the STC ranges for the ability to combine the perceptual characteristics of the data and the graphic variables according the desired attributes. In that way the user can identify patterns of the objects, and query them according to a particular class or value.

- *when + what* → *where*

Concerning to the question (*when + what* → *where*), the attention lies on describing the location or sets of locations of objects in or over a given time. As explained above, the location (where) question

has a strong relationship with the objects (what) question. A location loses its (exploratory) meaning if there is no objects to investigate. But on the other hand an “empty space” (no interest objects) can be useful if the user plan to compare the geographic characteristics it with characteristics of other locations that do hold objects. The representation of the objects locations in the STC is aided by the combination of the 3D cube and the 2D map. The objects on the 2D maps are presented as well in the 3D map allowing the user to retrieve the X, Y coordinates (2D map) with a simple mouse click and state the relationship of space and time on the (3D map)cube. As the user click on any of the views a highlighting “blink” will connect the respective object on both views. Additionally, the user can manipulate the area to be depicted on the cube to target specific locations. In Fig. 3.4, a relation of space locations is shown.

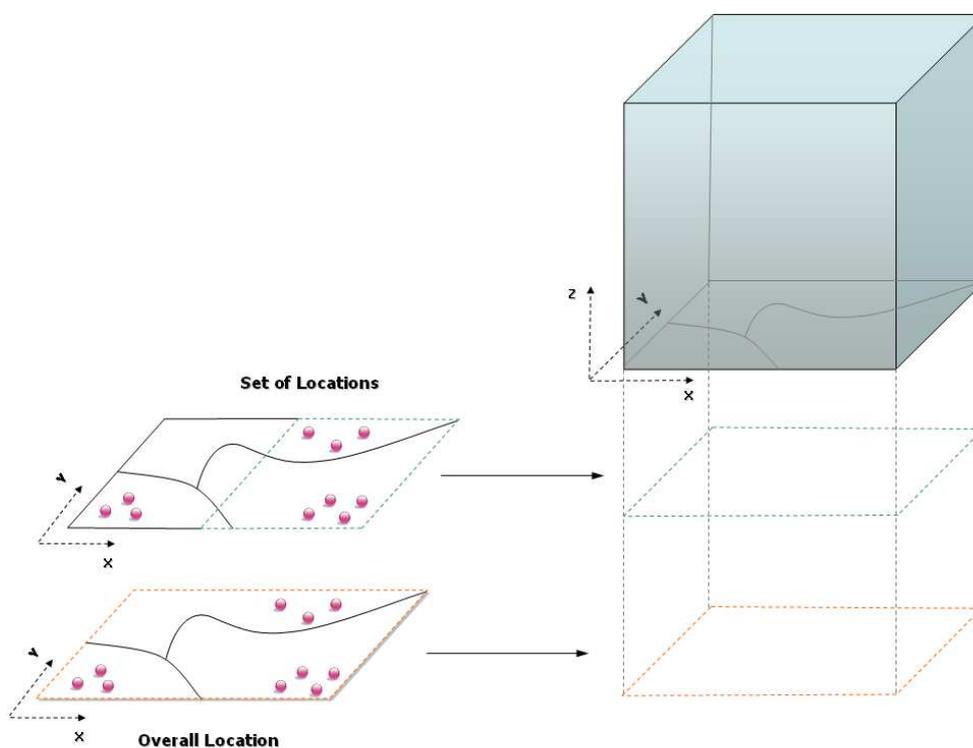


Fig. 3.4. The Relationship between STC and (*when* + *what* → *where*) Triad Framework (Source: Author).

- *where* + *what* → *when*

In relation of the question (*where* + *what* → *when*), consideration on describing the times or time terms for object/s in given location/s are addressed. In here, the time of the objects are represented using the height of the cube (z – axis). The user can link space and time for any particular object following the X, Y reference lines (X, Y, coordinates) and the Z – reference lines (T = time component) of the cube. Additionally the user can filter objects belonging to any particular time term or dress the time with the use of the mouse. Regarding the filtering, the time resolution adopted will definitely influence the amount of objects belonging to any specific time term. Fig. 3.5, shows an example of such case in where different objects are affected by different time’s terms.

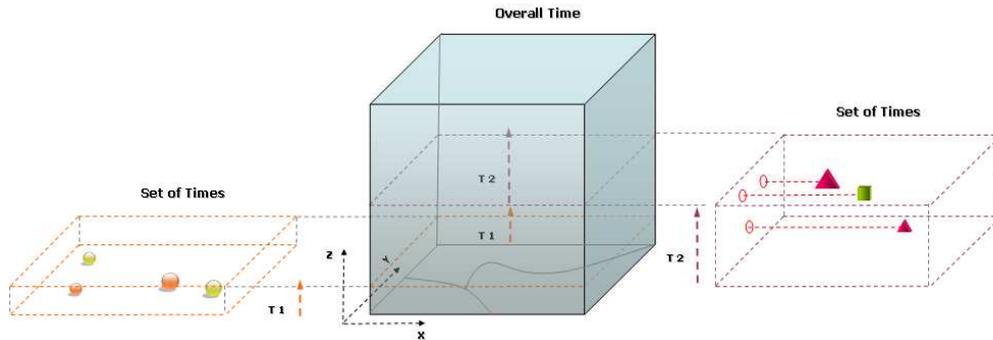


Fig. 3.5. The Relationship between STC and (*where + what* → *when*) Triad Framework (Source: Author).

In the next section a discussion of the possibilities of the STC to manage archaeological sites are presented.

3.2.2. Space-Time Cube Visualization for Archaeological Sites.

In the previous section the relationship of the Triad Framework and the STC was discussed. Moving a little bit further on exploring spatio-temporal phenomena we can say that the combination of all components (what, where and when) in the STC, can help to pose different kind of questions regarding the desired granularity. For example, a user could be interested in a set of objects that belong into particular locations in a specific time term or describing the overall time related to particular objects in the whole area, (see Fig. 3.5). The possibilities to combine the desired search levels to explore spatio-temporal phenomena makes it more than suitable for explore archaeological events.

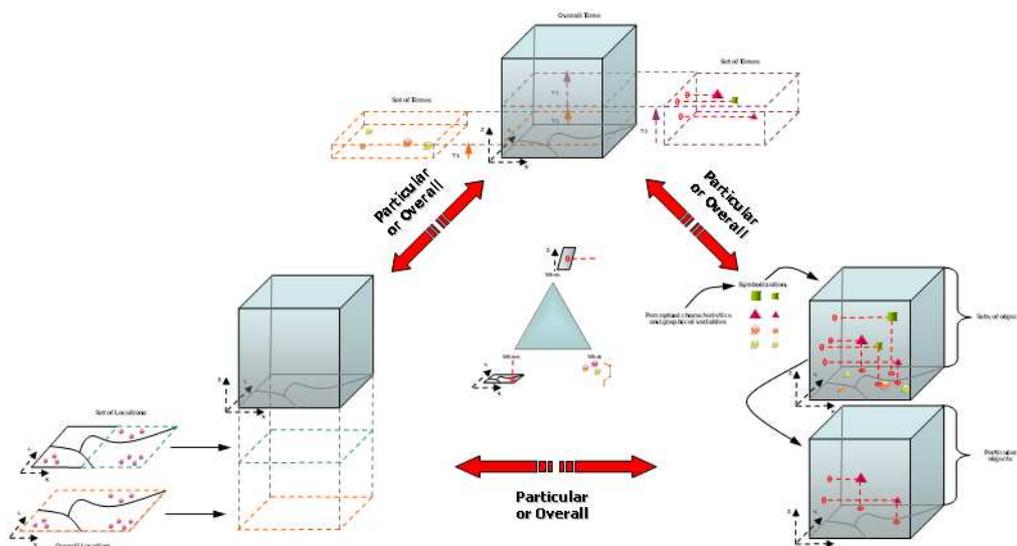


Fig. 3.6. The Relationship between STC and search levels (Source: Author).

Even more interesting than the capabilities of combining the search levels is the facility of the STC to deal with the three spatio-temporal dimensions (objects, location, and time), that make it attractive to explore archaeological events. For example, archaeological events can be visualized in the cube according their respective time term. In that sense the use of stations is the best choice to depict the appearance, disappearance, and duration of the events. Visualizing the events in that way the user can make use of the (X, Y, Z) references lines of the cube as a follow up to retrieve their temporal and spatial characteristics. Additionally he can make use of a particular time term to explore possible interaction in time, (see Fig. 3.7).

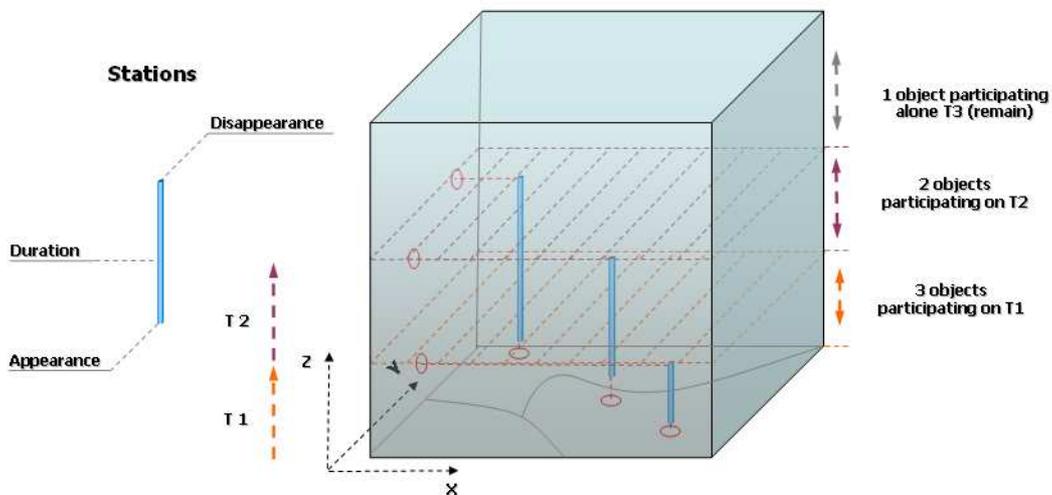


Fig. 3.7. The STC depicting (where and when) for archaeological sites (Source: Author).

Regarding the what question the user can use the facility of the STC to depict events by classification schema. One of the main interests of an archaeologist is to visualize the events according their respective class cultures. Coloring and labeling can provide meaningful information about what kinds of events are interacting or spread in a certain location in time, (see Fig. 3.8). Furthermore, the combination with stations can complete in a certain degree the visualization environment needed to deal with archaeological events.

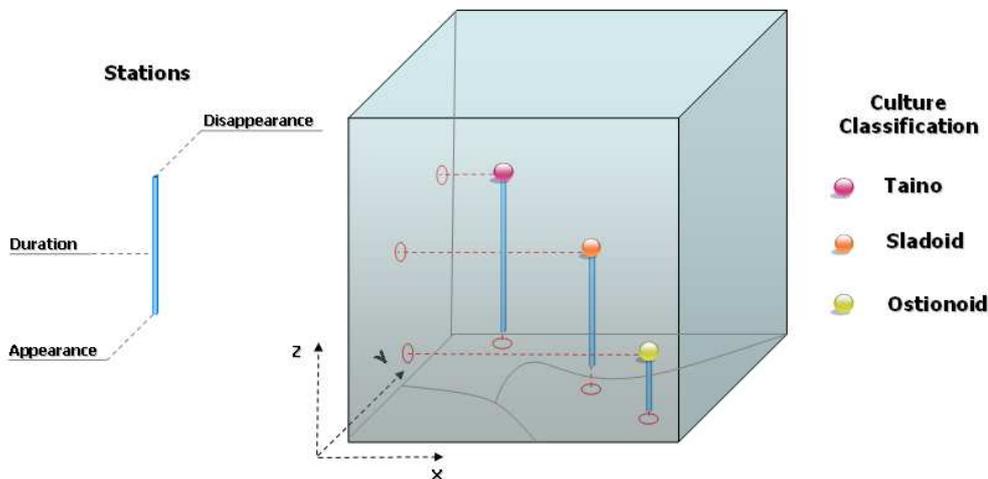


Fig. 3.8. The STC depicting (what, where and when) for archaeological sites (Source: Author).

More emphasis should be given to the characteristics of the archaeological events. What is discussed above seems to provide the essential visualization environment for dealing with simple characteristics (location, time and a singular attribute). But the fact is that when we try to explore or visualize complex data (e.g. multivariate data), more options and functionalities are needed. In that regard PCP, histograms, and scatters plots pose a standard solution for an effective geovisualization environment (see Fig. 3.9). For example, a user can be interested in describing the behavior characteristics of an object or a set of objects; investigating the frequency for a particular attributes; or verify if two variables present a correlation. The combination of linked views into the cube can help to answer the questions related to a particular or set of objects. Despite that a reliable solution for exploring multivariate data can be found in using additional linked views, it is necessarily to think how the STC environment can be enhanced to manage the multivariate components of the data in order to provide a more robust and complete application for exploring spatio-temporal phenomena.

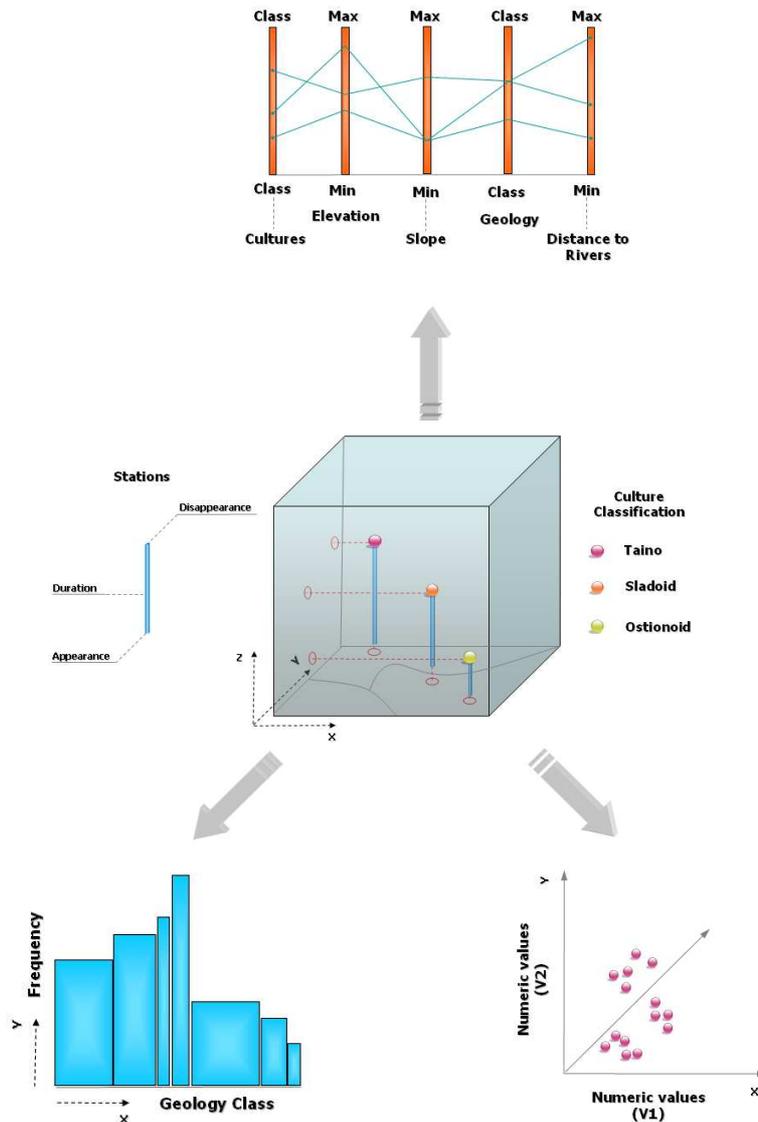


Fig. 3.9. The STC and linked views (top: PCP, middle: STC, bottom left: Histogram, and bottom right: Scatter Plot) for exploratory analysis (Source: Author).

3.3. Semantic Network Graph

3.3.1. Justification of Semantic Graph

As mention before in section 2.1.3 graphic drawings techniques allow the user to visualize relationship between objects according attributes of interest. The visualization takes the form of nodes and edges in where the objects are connected base on similar characteristic or some kind of association schema. The main advantage of the graphic drawing techniques is that strictly concentrates on objects and theirs attributes (e.g. multivariate data). For example, in Fig. 3.10 a basic form of a graphic drawing technique is presented. Nodes and edges are defined (nonplanar graph) by the similarities of attributes for each object. In there, node P1 resemblance the same attribute value (a) in the same field (W) for the nodes P3 and P9 (attributes table). Subsequently, a connectivity table in the form of a matrix state what nodes (objects) are connected to each other. The result is a graphic drawing map, in where the user can explore and visualize the events relationships.

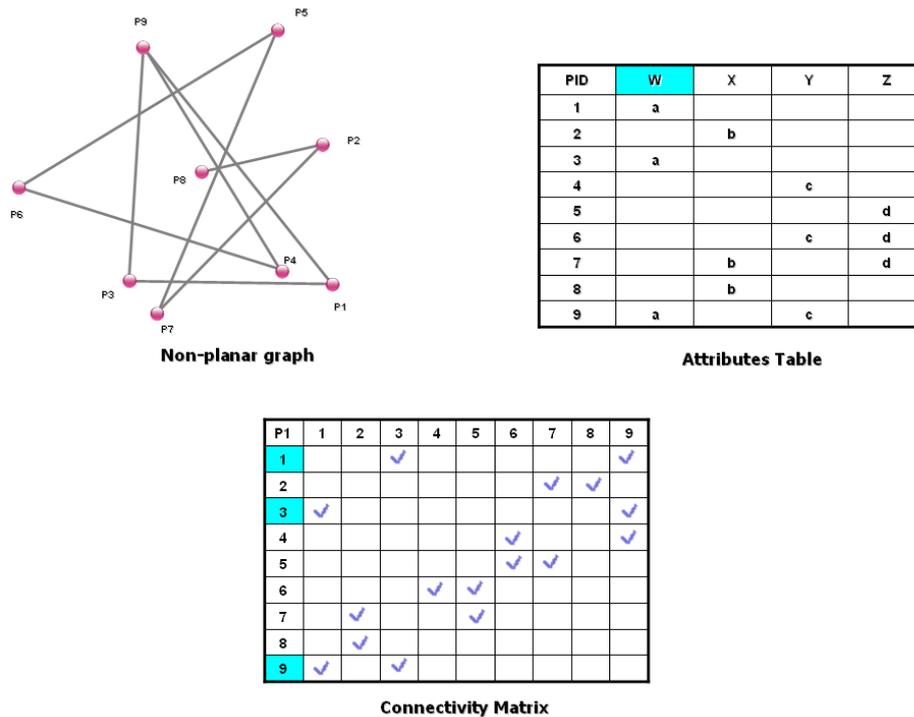


Fig. 3.10. A basic form a non-planar graph: top left: non-planar graph, top right: table attributes, and bottom center: connectivity matrix (Source: Author).

- *when + where* → *what*

Regarding to the question (*when + where* → *what*), graphic drawing techniques focus on much more complex data sets. Compared to the STC environment in where a limited set of attributes can be depicted with out the aid of additional linked views, graphic drawings methods offers the capability to handle the multivariate components of the data and display it in a more meaningful and simple from. Similar to the STC, graphic drawings techniques can use the perceptual characteristics and the graphic variables to depict the objects. In the same way a simple query can retrieved objects or sets of objects possessing similar or particular attributes, but much more that that is the ability to visualize and

display the connections between nodes according any desired attribute value or values. In that sense graphic drawings techniques provide the necessarily visual effect to quickly explore and retrieve the information.

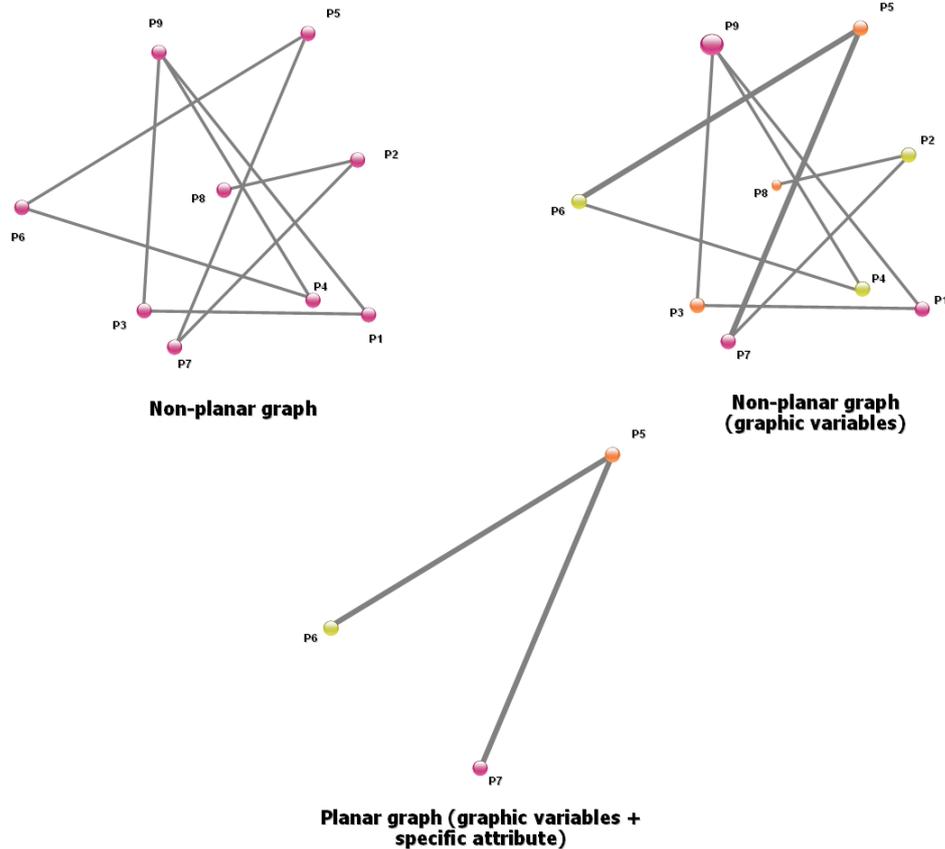


Fig. 3.11. The Relationship between Graphic Drawings techniques and (when + where → what) Triad Framework (Source: Author).

- *when + what → where*

Concerning to the question (*when + what → where*), graphic drawings techniques present different alternatives to deal with the location components. For example, preserving the exact location of the objects, moving the nodes into a relative position to each others, or totally arrange the nodes according particular attributes of interest or based on hierarchy. Because the purpose of this is study is to manage archaeological sites, retaining the exact position of the objects is desired. In that regard graphic drawing techniques act almost in the same way as the STC environment to depict an objects location. The user can simply retrieve the objects by querying the X, Y coordinates or going directly to the map and clicking the objects. Some particulars problems occur when the locations are fixed and too many objects are presented. Connectivity saturation, lines intersections and occlusions are the most common. In order to overcome such difficulties graphic drawings techniques can use the ability to depict the relationships according some specific attributes to filter the network, (see Fig. 3.11 bottom). Additionally the user can select a subset of objects and query them by attributes or by simply dragging a box around the objects of interest for exploring a particular set of locations, (see Fig. 3.12).

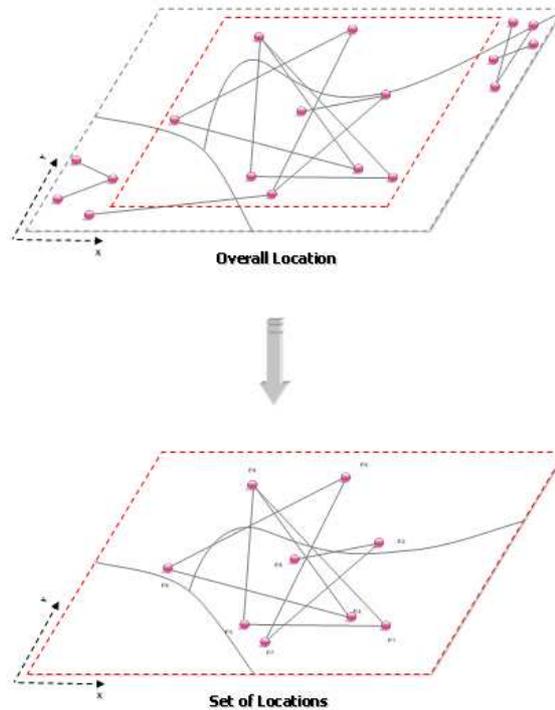


Fig. 3.12. The Relationship between Graphic Drawings techniques and (when + what → where) Triad Framework (Source: Author).

- *where + what → when*

In relation of the question (*where + what → when*), graphic drawings techniques can handle time in the following ways. Time can be considered as an additional attribute or the user can use the advantage of animations. Using the time component as an additional attribute the user can make use of the graphic variables to depict objects in order to visualize trend and patterns between locations or set of locations for an object or sets of objects. In that sense the user can manage the question (*where + what → when*) in the same manner as the question (*when + where → what*). But much attention should be give to the kind of time resolution adopted, thus this will influenced the amount of objects belonging into a particular time class. In the case of animations two options are available. First the user can make use of time as an attribute (as explained before) or use the specific time term belonging to the objects to address the appearance and disappearance and attributes (network changes) to build the animation sequences. In Fig. 3.13, the possibilities to for the graphic drawings technique to handle time are presented.

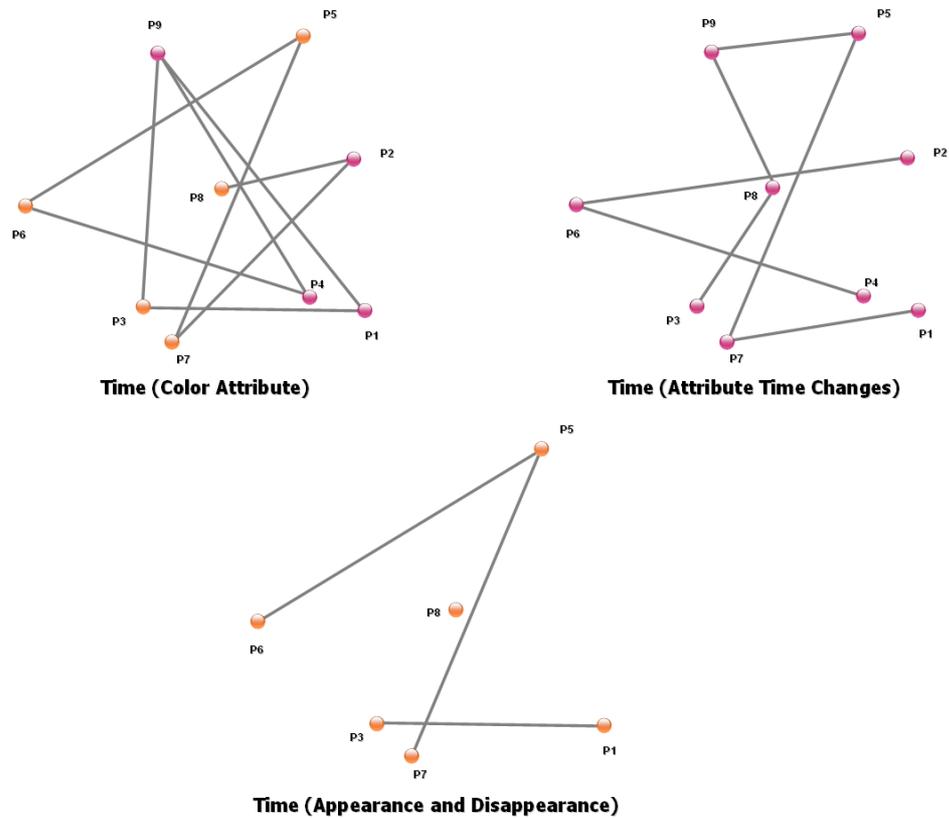


Fig. 3.13. The Relationship between Graphic Drawings techniques and (*where + what* → *when*) Triad Framework (Source: Author).

3.3.2. Graphic Drawings Visualization for Archaeological Sites.

Another main interest for the archaeologist is to establish associations between archaeological sites. Graphic drawings techniques can provide the necessarily medium to generate a more formal analysis in the form of a network to relate objects and investigate them. For example, a graph network can be generated using geographical, cultural attributes or both. Colors, size, and thickness can be applied to the nodes and edges using any desired attribute value. Additionally the network can change to explore any specific attribute of interest and filter any particular of sets of nodes to avoid occlusions. A table (see **Error! Reference source not found.**) stating the similarities between objects defines the drawing graph map in Fig. 3.14.

Table 3-1. Attribute table depicting similarities between objects. Dashed lines: as differences in attributes values (Source: Author).

PID	Geology	Soil Type	Slope	Dist. Rivers	Cluster Dist.	Elevation	Time
1	KJa	---	---	---	---	---	---
2	---	Rs	---	---	---	500 mts.	800 yrs.
3	KJa	---	---	---	---	---	---
4	---	---	40 - 60	---	0 - 2 miles	---	---
5	---	---	---	1.5 miles	0 - 2 miles	500 mts.	---
6	---	---	40 - 60	1.5 miles	0 - 2 miles	---	---
7	---	Rs	---	1.5 miles	---	500 mts.	800 yrs.
8	---	Rs	---	---	---	---	---
9	KJa	---	40 - 60	---	---	---	---

Attributes Table

The resemblance between P1, P3, and P9 according similar geology value KJa (Jayuya Tuff formation) can be seen in Fig. 3.14: top left. Once the network is generated considering all attribute similarities, the user can exploit the capabilities of the graph drawing map to explore the data. In Fig. 3.14, a series of potential steps to explore archaeological sites is presented. As mentioned before, an archaeologist’s main interest ranges in exploring the time period of particular or sets of events in order to establish the relationship between them. With the use of the graph map archaeologists can make use of the graphic variables (see Fig. 3.14: top left) to visualize the time culture of the events plus explore the network to see what kind of characteristics describe the objects. For example, in Fig. 3.14: top right, a filter using the geology field is applied followed by a coloring of the network (bottom left) base on another attribute. Additionally the user can focus on any particular node (see Fig. 3.14: bottom center) to explore the relationship with others objects and keep filtering the network to any attribute of inters inside the focusing node (see Fig. 3.14: bottom left). The combination of this dynamic interactions and the use of the graphical variables give the user the opportunity to depict at least 5 attributes into the graph map. The following possibilities are: Nodes with colors, size of the nodes, filter the network, coloring the network, and thicker lines. Using this combinations archaeologist can explore the multivariate characteristics of the archaeological site in a more comprehensive manner.

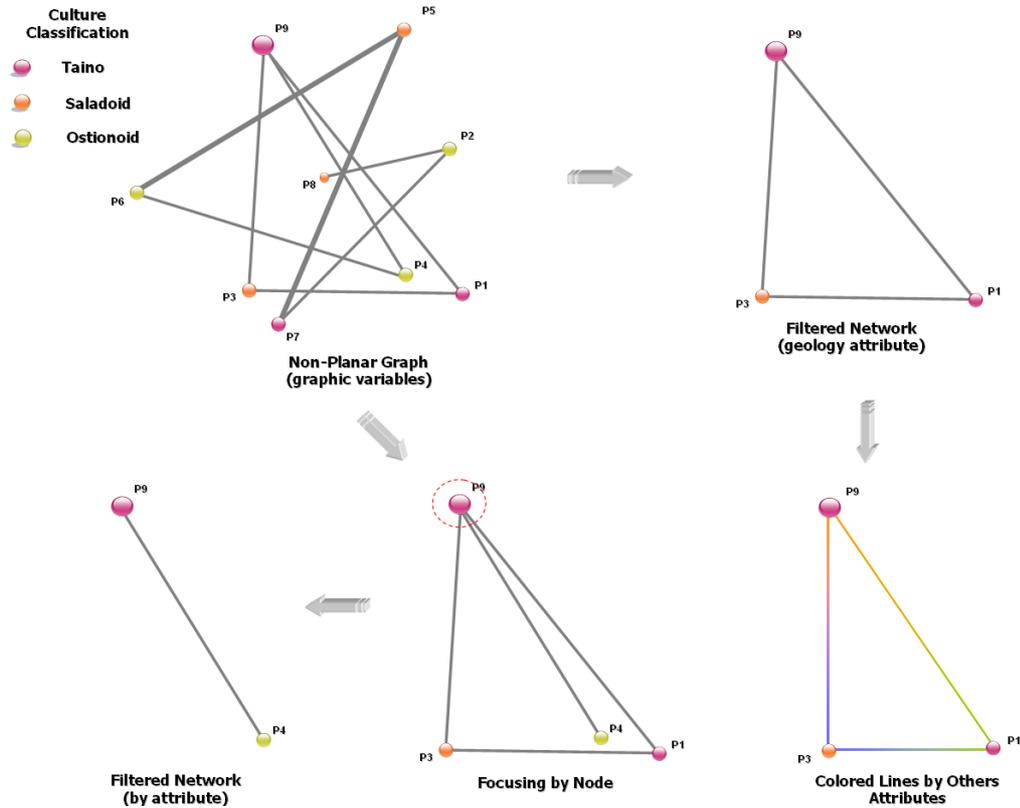


Fig. 3.14. A Graphic Drawing depicting multivariate characteristics (what) for archaeological sites. (Source: Author).

3.4. Semantic Multivariate Time Cube

3.4.1. Justification of Semantic Multivariate Time Cube

As mentioned in section 2.1.2.3 the STC environment focuses on depicting trajectories of moving objects. Making the use of the time domain and points locations the STC generate STP to depict the trajectory of the objects around the landscape. On the other hand, in section 2.1.3 the principles of Graphic Drawing were discussed. A graphic drawing map uses lines (edges) to establish relationship between points (nodes). One of the main advantages of a graph map for geographic mapping is the ability to retrieve point's locations (objects) as input nodes. In here lines can be seen not only as connections between objects, but more as the relationships on the multivariate characteristics that they possess.

The STC as well as the Graphic Drawing techniques use points and lines to display the objects. On one hand, the STC focuses on the spatio-temporal characteristics and on the other, the graphic drawings on the multivariate characteristics. As both practically work on the same data format, the STC environment can take the advantage of the Graphic Drawing techniques to fill the gaps to depict the multivariate characteristics of the objects (especially for non-moving objects). A Semantic Multivariate Space-Time Cube is proposed in order to extend the capabilities of the STC environment to manage spatio-temporal multivariate data, (see Fig. 3.15).

In the Semantic Multivariate Space-Time Cube, or better called Semantic Multivariate-Time Cube (SMTC), the STP depicts the relationships between objects. Any line that connects the moving object in the STC is used to connect non-moving objects in the SMTC according to a semantic relationship, see Fig. 3.10 for the connectivity method. One of the main differences in the SMTC is that for every line display (2D and 3D view); a color classification can be applied to depict any particular attribute of interest. On the contrary to the STC that only permits one color per line. In the 3D view (SMTC) this is achievable assigning a particular line Identifier for each related node to fragment the STP into different lines. On the 2D view (SMTC) the footprint of the original STC is changed from poly lines into lines. The result is a 2D non-planar graph view.

- *when + where* → *what*

Regarding to the question (*when + where* → *what*), the SMTC as mention before take the advantages of the two visualization environment (STC and Graphic Drawings) to depict the objects. The objects can be still retrieved by a simple query form or by a simple mouse selection. On one hand the user has the ability to use the perceptual characteristics of the data to assign the proper graphic variables and on the other, use the line connections to explore what kind of similar attributes posses the objects. In Fig. 3.16, alternatives of possible choices to explore the objects are presented. The user can: visualize the drawing map inside the cube, depict particular node connections, retrieve the similarities using the footprint, and explore the multivariate characteristic using the linked views and drawing map coloring.

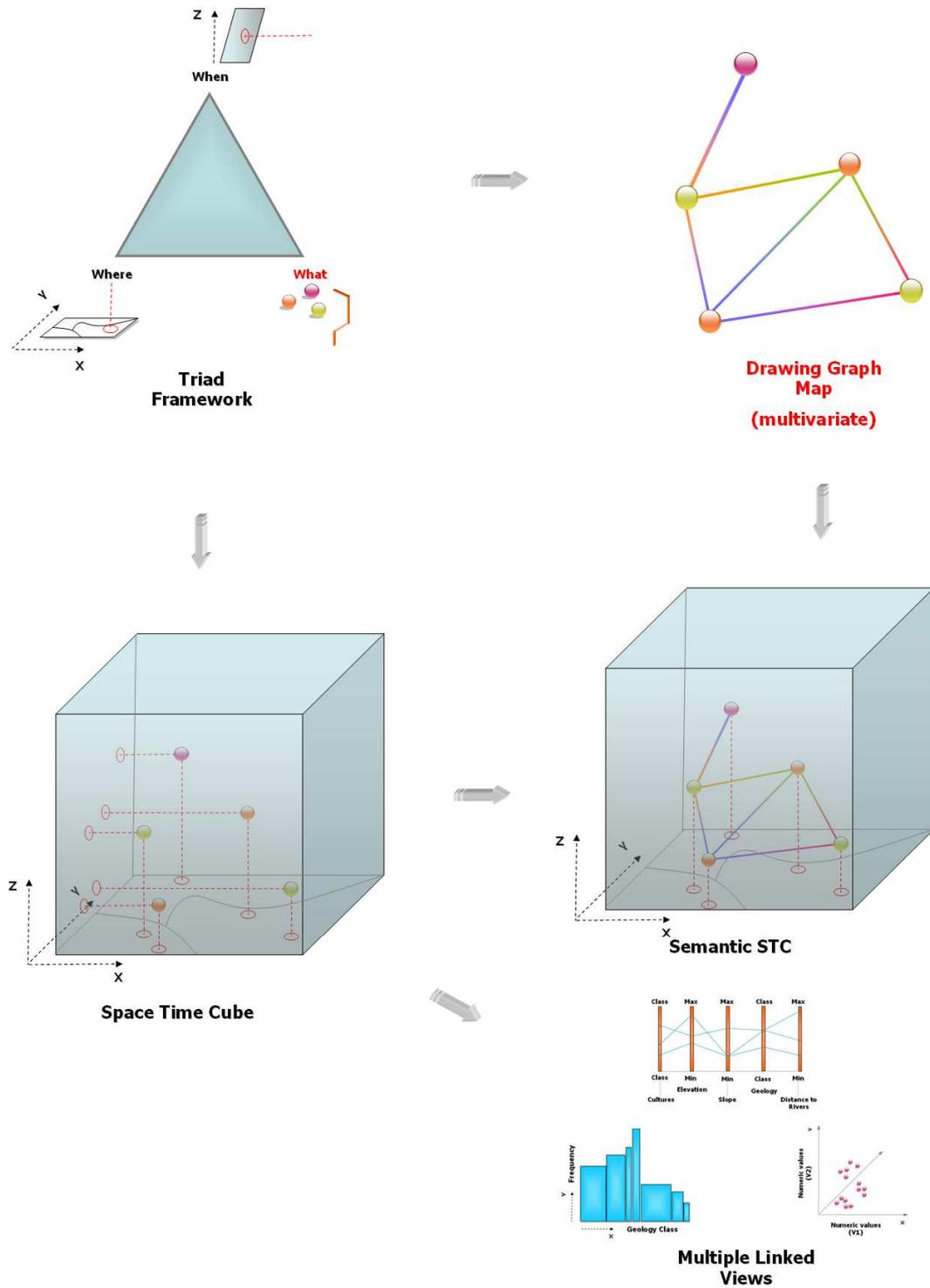


Fig. 3.15. The Relationship between the Triad Framework (Peuquet, 1994), STC, Graphic Drawings, and SMTC (Source: Author).

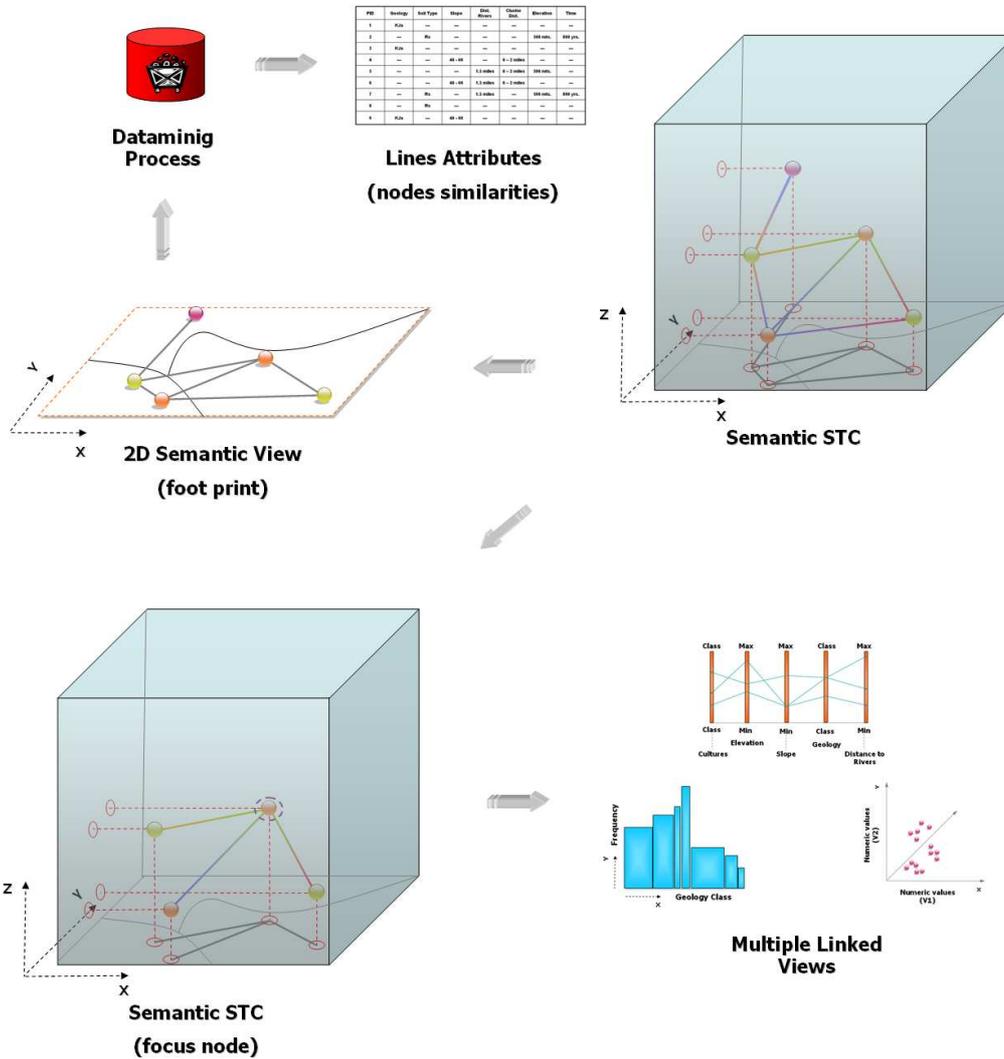


Fig. 3.16. The Relationship between SMTC and (*when + where* → *what*) Triad Framework (Source: Author).

- *when + what* → *where*

Concerning to the question (*when + what* → *where*), the SMTC environment reacts exactly the same way as the STC. Locations are presented in the 2D map view linked on the 3D cube view. The user can retrieve the coordinates of a particular object clicking in the 2D map view and visualize the respective object in the 3D cube view, drag a box around the objects. In the same manner as in the STC, the location of the cube can be fixed to depict particular cluster areas. Regarding to the connections, areas can be occluded if too many line are displayed. This presents a problem if the user is interested in retrieving any particular line or node that is positioned near other objects. Here the user can filter the network using any particular attribute to avoid occlusions and then retrieve the specific locations of the object of interest. As mention before a geovisualization environment always reveals a close relationship between the question what and where. For that reason a proper combination of what and where functions allow the user to maximize the performance in order to answer the related questions.

- *where + what* → *when*

In relation of the question (*where + what* → *when*), the SMTC environment treat the time component as another attribute for the objects to be depicted. As the SMTC is more oriented to manage non-moving objects time is strictly treated as a time term classification in order to allow a semantic relationship in time. That means that objects that belong to particular time class will be related by lines connections. In that sense the SMTC can make use of the height of the cube (z – axis) to position the objects and show the relationships in time as well as other attributes in a form of a graph map. As mention in section 3.3.1 the user then can answer the question (*where + what* → *when*) in the same mode as (*when + where* → *what*), (see Fig. 3.3).

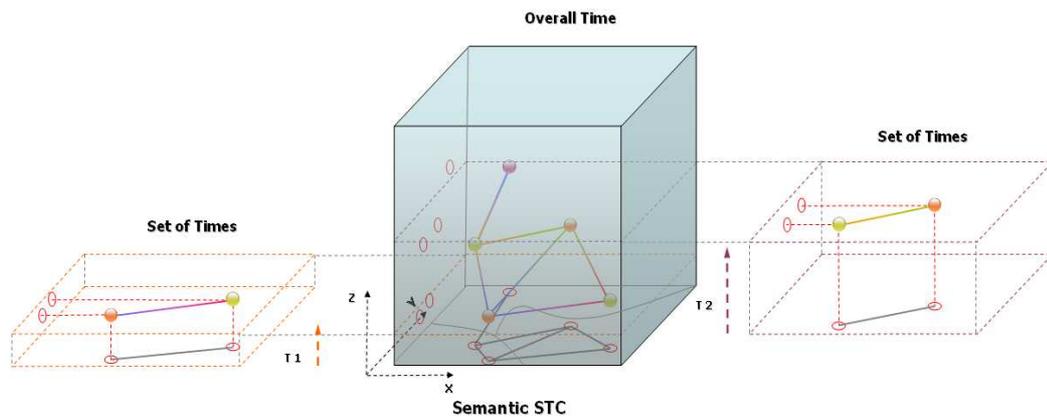


Fig. 3.17. The Relationship between SMTC and (*where + what* → *when*) Triad Framework (Source: Author).

3.4.2. Semantic Multivariate-Time Cube Visualization for Archaeological Sites

In section 3.2.1, the possibility of the STC with the aid of linked view was discussed in order to explore spatio-temporal events. It was also mention that the STC is limited to manage more complex data (e.g. multivariate data) without the inclusions of such linked views. In section 3.3.1, this limitation could easy be address exploiting the capabilities of the Graphic Drawings techniques as a form of a semantic graph map. The result is a SMTC that could easy satisfy the need to display, explore, and retrieve meaningful information for any particular spatio-temporal multivariate data set. It was also mentioned what possible solutions the STC and the Graphic Drawings can offer in order to deal with a particular data set (archaeological data). Using the proposed model (SMTC), for archaeological events the user can not only visualize the context, but also discover, explore, and answer more complex questions regarding time, location and multiple attributes within one geovisualization environment. In Fig. 3.18, the main core of the proposed geovisualization environment (SMTC) for exploring archaeological events is presented. The user can simply visualize the events with the use of the station to identify the associations in time. Making use of graphical variables to distinguish events as well as using the graphic network map to see and state who is related to whom, and determine why this relationship takes place (attributes). The ability to change the connections as many possible ways (if the data sets allows) to visualize any particular attribute

relationship and prevent incursion into other relationships using additional graphics variables, makes the application flexible for exploring the multivariate components of the events. The rotation, zooming and panning as well as other native tools and options inherited, provide the complexity needed for a geovisualization environment to deal with such difficult data set.

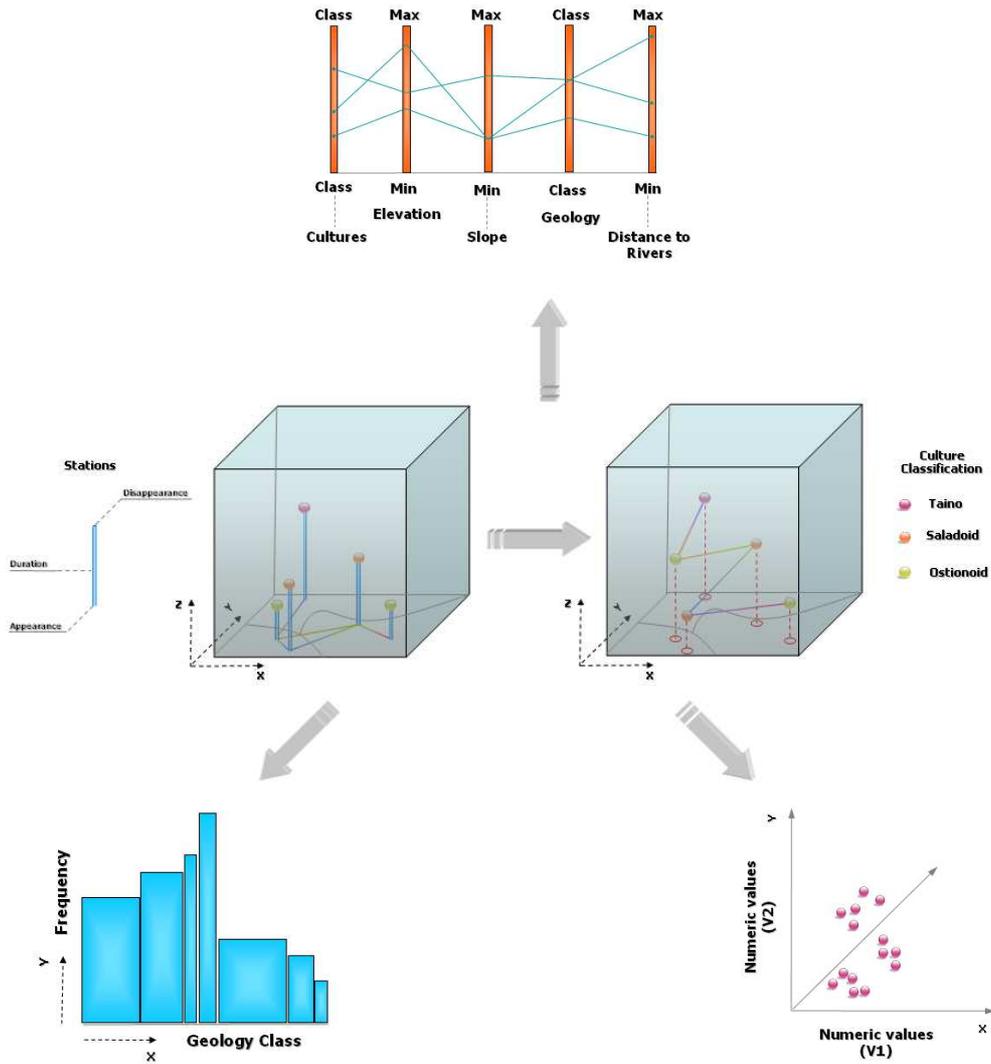


Fig. 3.18. The SMTC and linked views for an exploratory analysis (archaeological events) (Source: Author).

3.5. Multivariate Temporal Framework

Table 3-2. A Multivariate Temporal Framework (Source: Author).

<i>Component</i>	<i>Visualization Environment</i>	<i>Visual Output</i>
What	Net Work + graphic variables	Classification Colors + Labels
Where	STC (Base Maps + Coordinate Grids)	Base Maps + Coordinate Grids
When	STC (time visualizations)	STC (Stations, point in time (z-axis))
What-Where	STC (time visualizations) + Base Maps	(Classification Colors + Labels) + (Base Maps + Coordinate Grids)
What-When	Net Work + STC (time visualizations)	(Classification Colors + Labels) + STC (Stations, point in time (z-axis))
Where-When	STC (Base Maps) + STC (time visualizations)	(Base Maps + Coordinate Grids) + (STC (Stations, point in time (z-axis)))

3.6. Combined Tools and Options Evaluation

Error! Reference source not found. below represents the tools and options that have been evaluated and implemented for the proposed SMTC model. Priorities functions are marked with orange colors. Due to time limitations, not all priorities functionalities were implemented. The remaining sections provide the series of steps design taken and a description of the implemented functionalities.

The criteria for the reviewed functionalities are based on the literature review (Chapter Two) focuses on the search levels and operational task topology (Andrienko, 2003), the nature of the environments to be combined as well as similar visualizations and govisualization environments, and the actual status of the STC developed by The International Institute For Geo-Information Science and Earth Observation (ITC).

3.7. Prototype Stages

In order to clarify how the prototype development was made the subsequent section are discussed. The prototype setup was mainly compound by three stages:

- 1) First Stage - consists mainly on the design of an external application (Java), which allows the users to convert the input information (e.g. data set attributes) into a matrix table specially formatted for the STC software. The output takes forms as a semantic text file in were all possible relationships between the objects are declared;
- 2) Second Stage - the design of a programming code that retrieves only the similarities of two connected events and added to the STC footprint attributes;
- 3) Third Stage - restoration, addition, and debugging of options that were absent in the previews version of the STC program.

Table 3-3. Tools and Options Evaluation (Source: Author).

<i>Tools Name</i>	<i>Technique</i>	<i>Standard (S) Specialize (E)</i>	<i>Suitable (yes - no)</i>	<i>Implemented-(I) On progress-(P)</i>	<i>Search Level Elemental (E) General (G)</i>
Select	All	S	yes	I – P (update)	E
Clipping	Lexis Pencil	E	yes	P	G
Spinning	Lexis Pencil	E	no	N/A	
Cropping	Lexis Pencil	E	yes	P=(focusing)	E
Flight trough	All 3D	S	yes	N/A	
View point fixed	Lexis Pencil	S	yes	P	E
Interactive enlargement	Lexis Pencil	E	no	N/A	N/A
Color fading	Lexis Pencil	S	yes	P	E
Length and Position Adjustment	Lexis Pencil	E (multi V)	no	N/A	N/A
Brushing	PCP	S	yes	P	E - G
Strumming	PCP	E	yes	P	E
Filtering or focusing	PCP	S	yes	I (partial) – P (update)	E
Variable assignation	PCP	S (multi V)	no	N/A	N/A
Classification and coloring	PCP	E (multi V)	yes	P	G
Dragging and repositioning	PCP	E (multi V)	no	N/A	N/A
Rotation	STC	S (3D)	yes	I	G
Zooming	STC	S	yes	I	G
Panning	STC	S	yes	I	G
Highlight	STC	S	yes	P	E
Linked Views	STC	E (multi V, 3D)	yes	P	E-G
Query Functions	STC	S	yes	I	E-G
Footprint movement	STC	E (3D)	yes	I	G
Activation of layers	STC	S	yes	I	G
Extruded surfaces	STC	E (3D)	yes	I	G
Focus node movement	Graphic Drawings	S (multi V)	no	N/A	N/A
Rapid Zooming	Graphic Drawings	E	yes	P	E
3D widgets	Graphic Drawings	E (3D)	yes	I	G
Elision	Graphic Drawings	S (multi)	yes	P = focusing	E
Reference Grids	STC	S	yes	P	G

3.7.1. Stage 1 (Java Code Matrix Implementation)

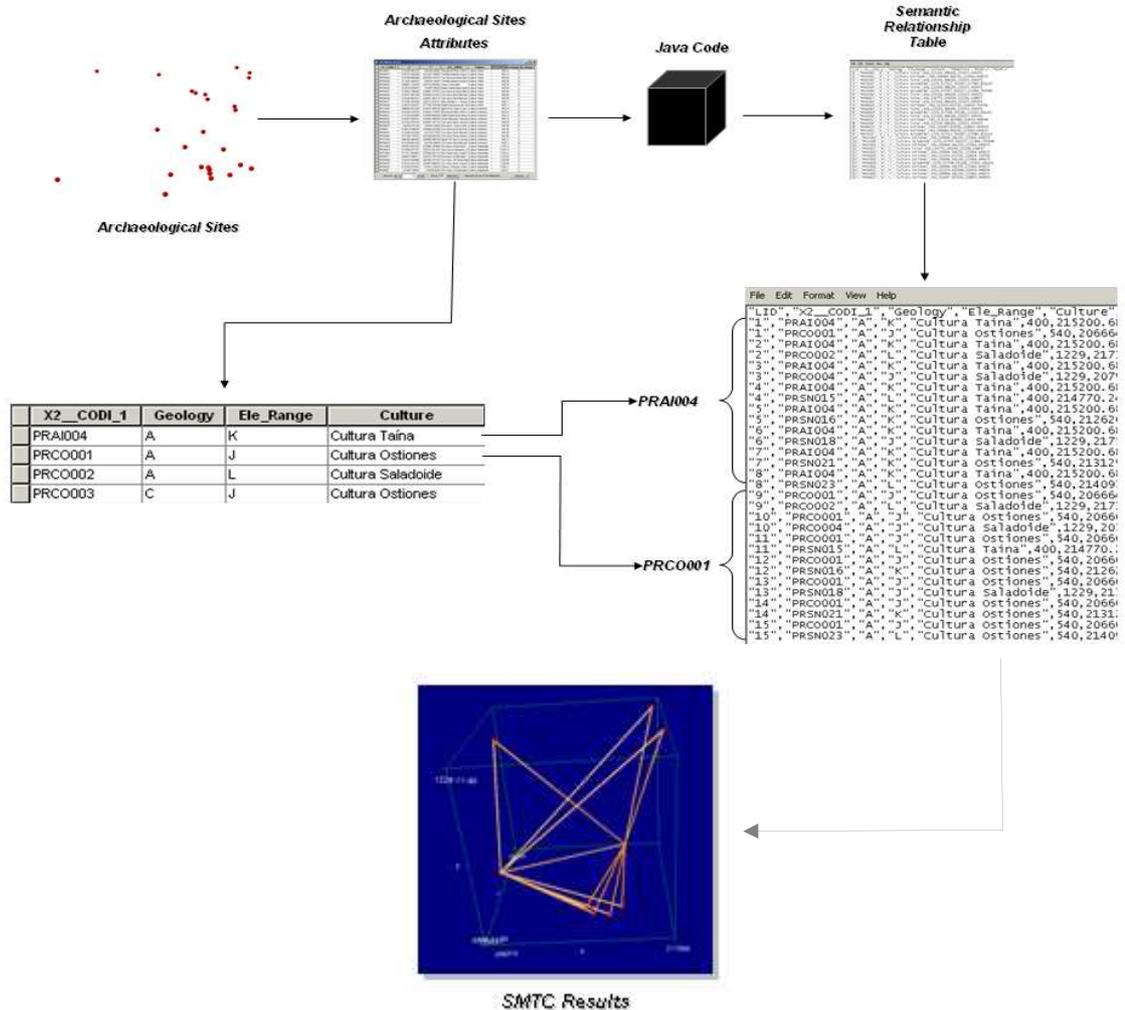


Fig. 3.19. Matrix code implementation (Source: Author).

The developed application begins by retrieving similar values field by field of the input table. Once the code finds any similar values involving two rows in the same field, a relationship is created using a line identification field (LID) for a new text file (STC format). In order to avoid duplicated relationships at the time to compare the remaining fields. All the attributes from the first comparison are retrieved. Hence, all possible relationship on the remaining fields are automatically inherited, and the code starts comparing the all the values in the second field to corroborate if a LID value in the new table was already declared. If that not the case, a new relationship is declared by adding a new LID value. The process continues running until all field values are compared. Fig. 3.19, shows an implementation test using fixed attributes values from a sub-set archaeological data.

3.7.2. Stage 2 (Java Code Footprint Data Mining Implementation)

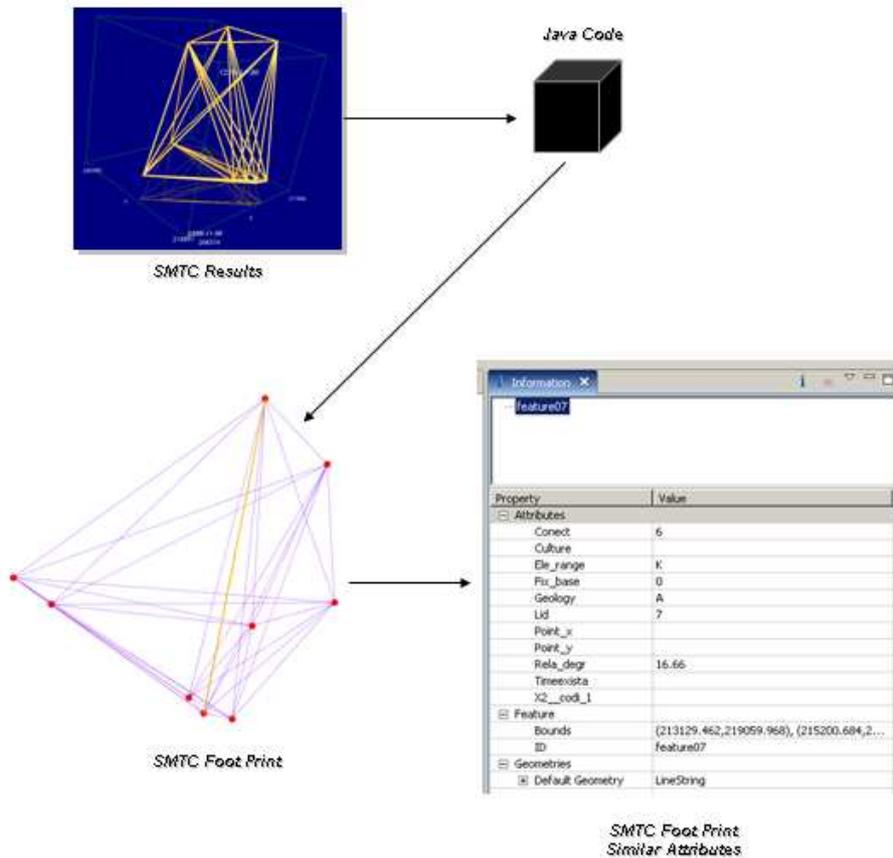


Fig. 3.20. Footprint code implementation (Source: Author).

In order to adapt the footprint of the original STC software, which in the first place was designed to depict the trajectories of moving objects, more java code was added to improve the exploration capabilities of the SMTC. As the main purpose of the SMTC is to depict and retrieve information of related objects, the new semantic footprints will allow the access of the similar characteristics between related objects. Any similarity founded between two objects on the 3D graph will be inherited by the line footprint. Hence, the footprint becomes a layer with meaningful information that can be queried/explored. Furthermore, as the footprint is a mirror image of the 3D graph and the original version of the STC has the ability to export the footprint as a layer file for external applications (e.g. Erdas, or any GIS software), different relationships can be extracted according the users preferences of the 3D graph limited only by as many attributes field and combinations that the user can think. This gives the user the flexibility and ability needed to handle the complexity of multivariate data, (see Fig. 3.20).

3.7.3. Stage 3 (Additional Options)

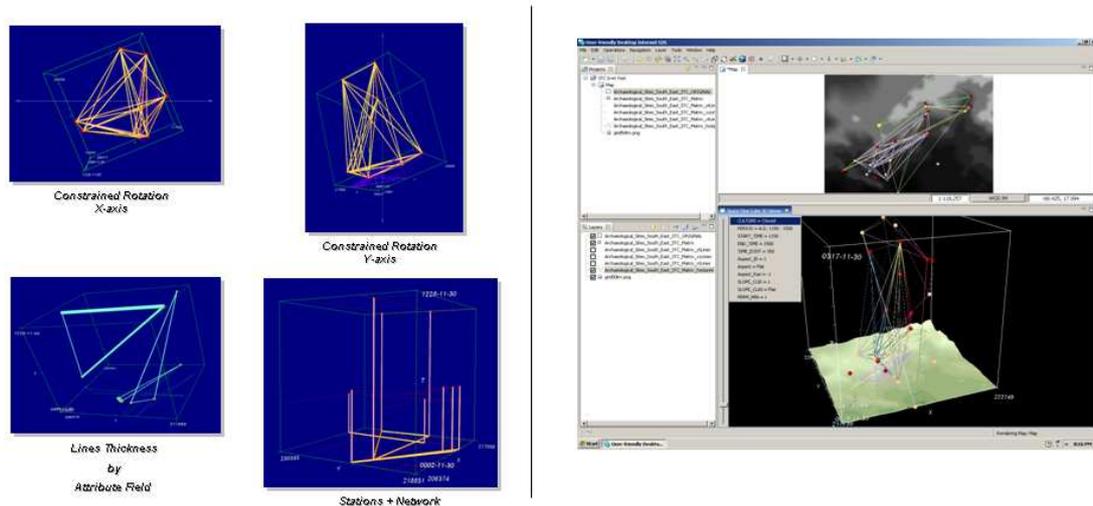


Fig. 3.21. Left: Additional options, Right: Final Version (SMTC) (Source: Author).

In order to maximize the performance of the previous STC version and make it more suitable for the SMTC prototype the following option were included or restored:

1. *Constrained Rotations*: fixed rotations in X, Y, and Z-axis were implemented in order to provide a full control movement of the cube. This facilitates the user to maintain any perspective in X, Y or Z direction without affecting the others.
2. *STP Lines Thickness*: the user will have the option to assign any numerical attribute to the thickness option field in the STP to depict the information.
3. *Stations*: the stations option was not available on the previous version of the STC. To facilitate the exploration of the archaeological sites, the station option was repaired. The user not only can use the station, furthermore combine it with the semantic graph map.
4. *Coloring + Gradient*: coloring classification and gradients effects where implemented. As the main goal of the SMTC is to display multivariate data, coloring and gradients effects play a main role for geovisualization purposes.
5. *Similarity Info. Box + STP Highlight*: with the use of the mouse the user can quickly retrieve similarities between two objects or sets of objects. A highlight effect in both views (2D and 3D) helps the user to identify the related objects.

3.8. Summary

This chapter has discussed how the STC and the Graphic Drawing techniques can answer the question regarding the what, where, and when of spatiotemporal events. It has also discussed how archaeological events can be represented by these two (geovisualization and visualization) solutions. A proposed model for the combination of the STC and Graphic Drawings was presented and explained.

As illustrated in the chapter, a range of tools and functions have been implemented for the visual exploration of space-time data. Simulated tests have been run in order to corroborate the

functionalities of such tools and functions. Still, a real case implementation and evaluation is needed in order to demonstrate the effectiveness of the proposed model (SMTC). The following chapter will introduce the archaeological case implementation using the SMTC prototype, and Chapter Five will discuss its evaluation.

4. Prototype: case study archaeology

4.1. Introduction

In Chapter 3, the potential to combine Graphic Drawings techniques with the STC was addressed. Considering the technology and options available, a prototype was designed called the Semantic Multivariate Cube (SMTC) in order to explore multivariate spatio-temporal characteristics of the data. This Chapter will discuss the implementation and exploration of archaeological sites on the SMTC as well the STC.

4.2. Geography of Puerto Rico (Area of Study)

The study area for this research is Puerto Rico (PR) Island, where the archaeological sites are spread around the landscape. Puerto Rico is part of an archipelago located between the Caribbean Sea and the Atlantic Ocean (east: Dominican Republic, west: Virgin Islands, north: Atlantic Ocean, south: Caribbean Sea), and is the smallest island located at the east part of the Greater Antilles. The location of PR Island is around the $18^{\circ}15'$ N, $66^{\circ}30'$ O and the extent is about 170 km by 60 km. The surrounding small islands that compose PR are: Vieques, Culebra, Mona, Desecheo, and Caja de Muertos. Its location is inside the tropical margin of the globe and it is characterized for being a volcanic island.

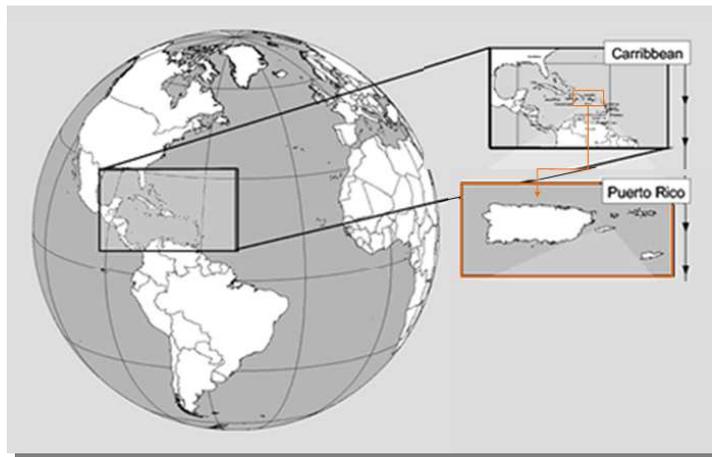


Fig. 4.1. Puerto Rico Location Map, (Landrau, 2002).

PR is compound by approximately 75% of irregular landscape, divided by three physiographic regions (central mountainous, northern karst, and alluvial valleys) and (54) perennials rivers flowing down hills from the central mountainous regions to the valleys into the sea. The central mountainous region is formed by sedimentary volcanic rocks as well as calcite, magma, and serpentines rocks crossing from east to west splitting the island in three parts (north, central, and south). The most elevated peaks

are: Cerro Punta (4,389 fts.), Rosa (4,156 fts.), and Guillarte (3,952 fts.) and the lesser ones are: El Yunque (3,494 fts.), Penuelas (3,414 fts.), and Torrecillas (3,093 fts.).

The karst regions of PR are situated at the northern part of the island with an extent from the north-west to north-east center. Characterized by their superficial and subterranean formations the karst of PR serves as an extremely important component supporting water to the hydrologic system and supporting a large number of native species. Caves, aquifers, and subterranean rivers are the most important subterranean features. On the surface level, dolines, “terraplen”, “mogotes”, pits, and lapies are found.

The alluvial valley extension covers the north, south, west, and east shore part of PR. Smaller islands, “cayos”, mangroves, dunes, and beach shore lines are the basic components of the alluvial valleys.

4.3. Case Study: Archaeology Case PR

Archaeological sites are locations or groups of locations in where evidence of past activity is found. Is the discipline of archaeology that focuses on the investigations of such cultural remains (historic or pre-historic) in order to understand, corroborate the existence of, and state the influences upon society. In Puerto Rico Island, archaeological sites are found spreading off all around the landscape. The archaeological sites not only play an important role in the society (establishing people’s identity), but also is a key factor for a number of applications like urban planning, and environmental impact assessment.

It is argued here that over the years, the manipulation of such rich data sets have missed the true potential that archaeological sites possess in terms of providing meaningful information that can be used for decision support applications. Based on this need, the STC and Graphic Drawing environments (brought together in the SMTC), will serve as a vessel in the geovisualisation of archaeological data, by considering their temporal, location, and attribute domain in order to retrieve and understand interrelationship between them and with the landscape .

4.4. Data for the Case

As the study aims to visualize the relationships of geographical variables in time (see section 1) the following layers were provided by the Natural Resource Department of Puerto Rico in conjunction with different entities (USGS, SSURGO, NRCS) that supplies most of geospatial information for the country. Geology, soil types, rivers, elevation contours, and digital elevation models (DEM) comprise most of the layers retrieved by those entities. Archaeological sites layers were provided by the “Council for the Protection of Terrestrial Archaeology of Puerto Rico”. Additional information will be generated for example, sub-sets of DEM, slope, aspect, distance to rivers, and a clustering analysis will be carried on to simplify the selected areas to be studied. All layers use the State Plane (NAD 83) Puerto Rico Virgin Island FIPS 2500 with a coordinate system in meters.

The following section will present a cluster analysis in order to target the most important areas for the case study in this thesis.

4.5. Cluster Analysis

4.5.1. Introduction

In the Archaeology field most of the information generated for the archaeological events are disseminated frequently as report forms (stating the amount of artifacts) and writing surveys (describing different issues like: sociopolitical system, moving patterns, feeding behaviors, etc.). With the introduction of GIS into the archaeological field it is now possible to retrieve the geographic locations of such events. Nowadays, archaeologist use point's data as short hand for positioning archaeological events into geographic space (X-Y coordinates). Not only the main position (regional scale) of the archaeological site is usually retrieved, but in some cases different locations are taken within a particular archaeological survey extent in where test pits (artifact locations) are accurately located. As mentioned before this research study will be concentrate using the archaeological sites at regional scale. Cluster Analysis methods as well as Data Visualization techniques will be applied in order to reduce the spreading complexity of the archaeological sites and justify the test subjects to input into the SMTC.

Cluster Analysis is defined as the process to “finding groups in the data, such as data points within each group (segment) being similar to one another by some metric” (U. Fayyad, et al, 2002). Their ability to discern and extract groups within the data sets can help to easily visualize, understand, and explore particular elements that embrace some kind of relationship (e.g., locations, attributes, etc.). A wide variety of cluster analysis methods exist, and are available within most GIS software packages. The intention here is not to provide a list of all the existing cluster analysis method but the most popular ones. For example, Nearest Neighborhood, Quadrant Methods, Dispersion Around the Mean Center, K-Function, Standard Distance, etc.

One of the major disadvantages for all these clustering methods is that they are strictly influenced by the amount of input points and the extent that comprise them. The spatial scale of the analysis must be considered as they perform better on describing possible clusters than extracting them for the whole data set. For example, performing a cluster analysis for country region can easily identify groups of objects that in a city region will look sparse and vice versa. On the other hand the result to execute a cluster analysis for an entire country can lead to a false impression of the behaviour of the objects if a description of particular clustering events is desired. In that regard, a prior Data Visualization of the whole data set (points) may provide the necessarily clues to identify clustered objects in a relative larger scale (subject to the researcher criteria) when the data sets is only composed of sets of locations and lacks other relevant information. The main purpose of Data Visualization is to detect patterns, structures, trend, relationships and anomalies in the data. In similar way geovisualization environments use a variety of map displays to stimulate the visual thinking of the geospatial data at hand. For example, choropleth maps, surfaces, cartograms, proportional dot symbols and isoline maps, compose the most traditional map views for a quick exploration of the geospatial data. As mentioned before Data Visualization techniques as well as the Clustering Analysis methods will be combined for the proposed model. The reason for this is that cluster analysis can be an effective method for determining areas presenting high concentrations of archaeological sites. But only by

using visual recognition, cluster analysis can provide much better results adding prior knowledge regarding the data set under study. Below the selected mapping techniques and cluster analysis tools are described in order to justify the choices.

4.5.2. Data Visualization Techniques

- 1) *Choropleth Map*: to first identify important areas, and have the first impression of the spatial patterning of the archaeological sites, and to quickly detect those areas (municipalities) in where the majority of the archaeological points are located.
- 2) *Quadrant Methods Maps*: will serve as good indicators to detect cluster areas considering the whole extent of the island in a predefined grid cell extent. Running various grid cell extents will provide the necessary visual output to clean up overlooked areas of the choropleth map.
- 3) *Proportional Dot Symbol Map*: based on the optimal quadrant method a dot symbol for each quadrant depicting the amount of the archaeological sites will provide a clear picture of the highly populated grid surrounding by less populated ones.
- 4) *Thiessen Polygons Maps*: the area of the thiessen polygons will be calculated in order to detect smaller areas, based on the principle of the thiessen polygons, smaller areas will show clustered points. The following considerations will be addressed in order to define the extent in where the clustered points are. Smaller polygons will be selected according to a classification based on area and the neighborhood of bigger polygons. A discrimination for the bigger polygons in order to include them as part of the cluster will be made by visualizing the position of the correspondence archaeological site to the boundary of the most near smaller polygons. A false impression using the thiessen polygons is that bigger areas can be generated by the isolation of the boundary points of the selected clusters with the remaining ones.
- 5) *Dot Symbol Map (Elevation Mean)*: in order to corroborate in some extent the selected areas, the archaeological points will be displayed against the voronoi map. Following the principle that things that are close tend to be similar, it is expected that a homogenization in the elevation values for the selected clusters will be seen.

4.5.3. Optimization of the Selected Clusters

4.5.3.1. Questions to be Answered

- 1) Do the selected possible clusters reflect a spatial pattern?
- 2) How confident is the distribution of the selected areas?
- 3) What is the best distance that describes the clusters?

4.5.3.2. Cluster Optimization Techniques

- 1) *Average Nearest Neighbor*: a simple Nearest Neighbor approach can tell us the information needed to corroborate the distribution of the selected clusters.
- 2) *Dispersion Around the Mean Center*: the dispersion around the mean center will be used to optimize those clusters that the standard deviation technique fails to achieve a 68% of the total points. Those archaeological sites that are completely within both (mean radius and standard ellipse) will be given more weight. And those that are in between are considered as possible outliers given more emphasis to include those events that are positioned perpendicular to the ellipsoid. As archaeological sites do not always present a

- clear structures is reasonable to think that no automatic clustering process can fully describe the events.
- 3) *Directional Distribution (Standard Deviation Technique)*: the directional distribution will be used to establish the spatial pattern for the archaeological sites and tell us something about the randomness.
 - 4) *Multi-Distance Spatial Cluster Analysis (Ripleys K-Function)*: the K-Function will be used to detect the possible distances that best describe the clusters. The use of the mean distance of the output of the K-Function will be use to establish the best relation in distance. Is expected that the mean of the K- function inter connect the majority of the nodes resemblance the shape of the cluster as well as possible outliers. The mean distance will be use as threshold to related the points on the SMTC as well as 1 and 2 miles distance in order to explore others geographical variables.

$$L(d) = \sqrt{\frac{A \sum_{i=1}^N \sum_{j=1, j \neq i}^N k(i, j)}{\pi N(N-1)}}$$

K-Function formula

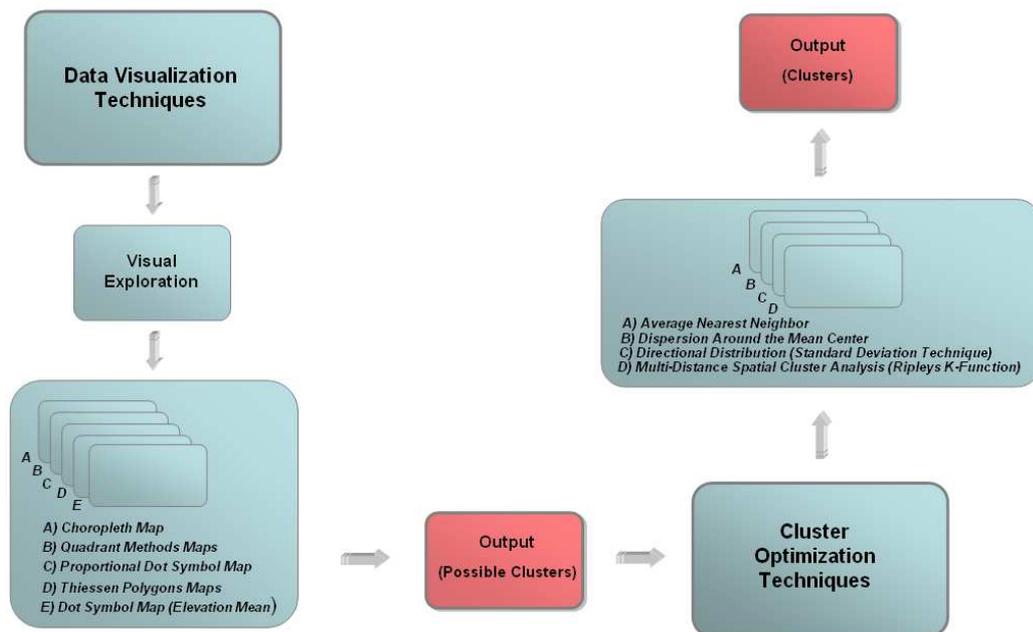


Fig. 4.2. Model for detecting clustered archaeological sites.

4.5.4. Visual Exploration

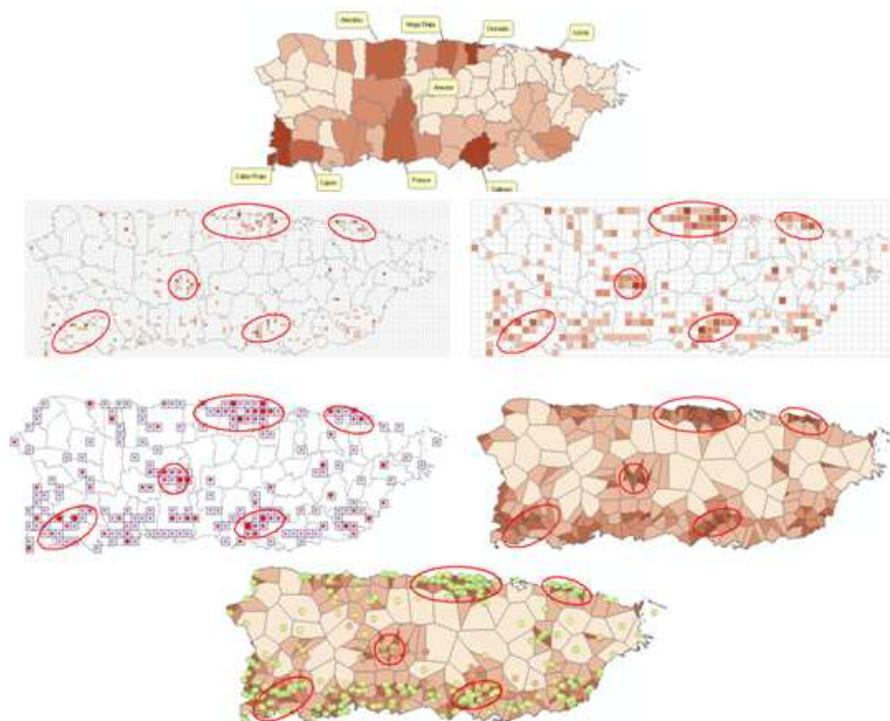


Fig. 4.3. Puerto Rico Island Maps (Visual Exploration). Top: Choropleth Map, middle top: Quadrant Maps (left: 1 mile grid, right: 2 miles grid), middle bottom: (left: Dot Map Symbol Map, right: Voronoi Map), and bottom: Voronoi vs. Arch-sites Elevation measurements.

The following describe the results for the Visual exploration techniques.

- 1) The choropleth map identifies the municipalities of Arecibo, Vega Baja, Dorado, and Loiza on the northern part of the Puerto Rico Island possessing a significant number of archaeological sites. On the southern part the municipalities of Cabo Rojo, Lajas, Ponce, and Salinas, and on the centre, Jayuya.
- 2) The Quadrants Maps clearly diminish the area extent of the possible clusters location. Comparing both quadrants maps (1 and 2 miles), the 2 miles grid seem to offer better results for visualization purposes.
- 3) The Dot Symbol Map clearly show those grid cells that posses a high frequency surrounded by those with low frequency. This provide a usefully information about possible sub-groups and distribution of the clusters.
- 4) Regarding to the Vorionni Map, the calculation of the areas for each polygon corroborates in some extent the identifying clusters. Smaller areas are depicted using a classification color in order to visualize smaller areas clustered together vs. the archaeological sites elevation.
- 5) The results of the Visual exploration are five possible clusters located at north-east, north center, center, south west, and south east part of the Island.

4.5.5. Cluster Optimization

4.5.5.1. North East Cluster

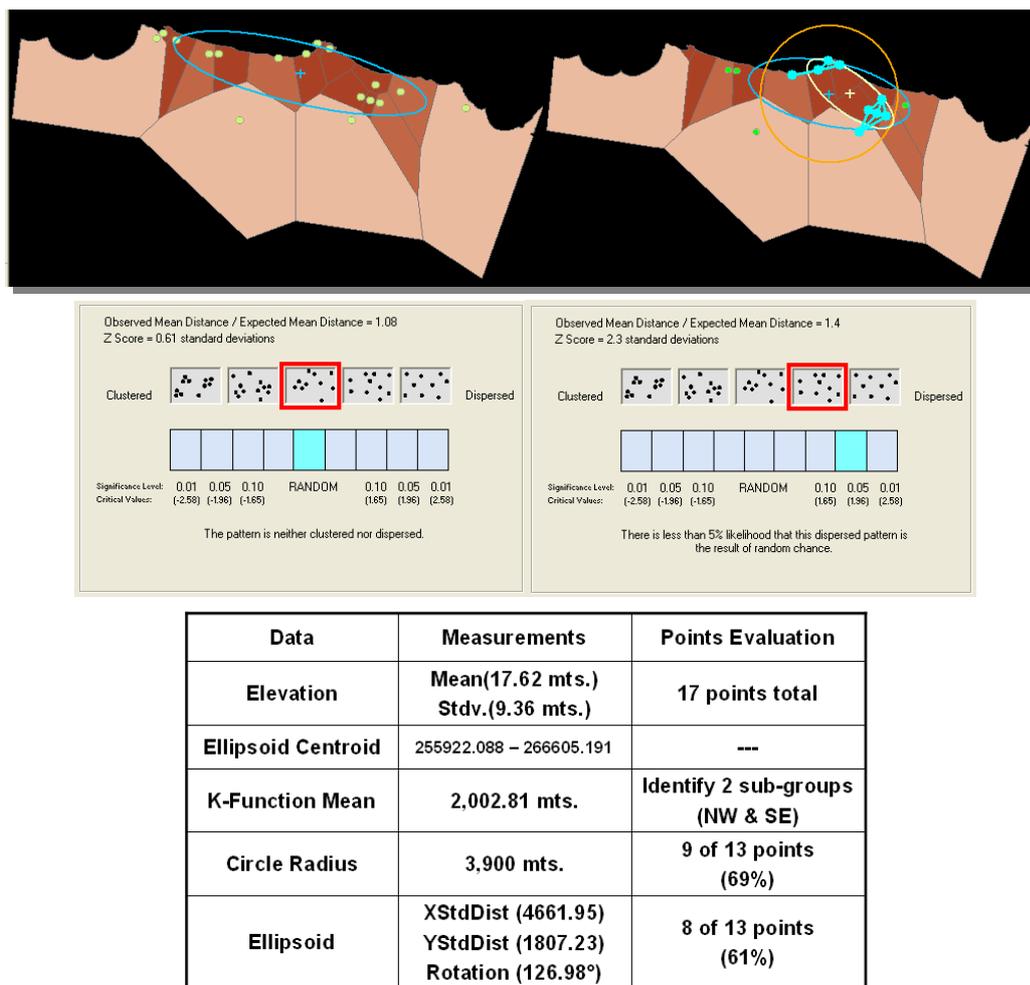


Fig. 4.4. Cluster Analysis Results. Top: Cluster exploration, Middle: Nearest Neighborhood results, and Bottom: Additional information.

North East Cluster Results: the cluster seem to guard a strongly relationship with the northern coast formation with a rotation of 126°. Using the maximum points (17) for the Average Nearest Neighborhood (detected by Visual Exploration Techniques), the results indicate that the cluster present a random behaviour with a Z-score value of .61 stdv. using a confidence level of 95% for the critical values -1.96 and +1.96. On the other hand the first try of the Directional Distribution (ellipsoid) state, that only 58% of the point's falls inside the ellipse. Hence, the data set does not present a normal distribution around the mean. After the attempt for optimization, removing possible outliers using the Dispersion Around the Mean Center and visualizing the points distances of the ellipsoid centroid, the Average Nearest Neighborhood shifted towards a more disperse cluster with a Z-score value of 2.3 stdv. exceeding the confidence level. The second Directional Distribution slightly

increases the result with 61%, corroborating the results of the first Directional Distribution. The K-Function clearly identifies to subgroups, the first on the north-west and the second on the south-east part of the ellipsoid. In conclusion, the optimization fails to improve the Visual Exploration results for the northern cluster.

4.5.5.2. North Center Cluster

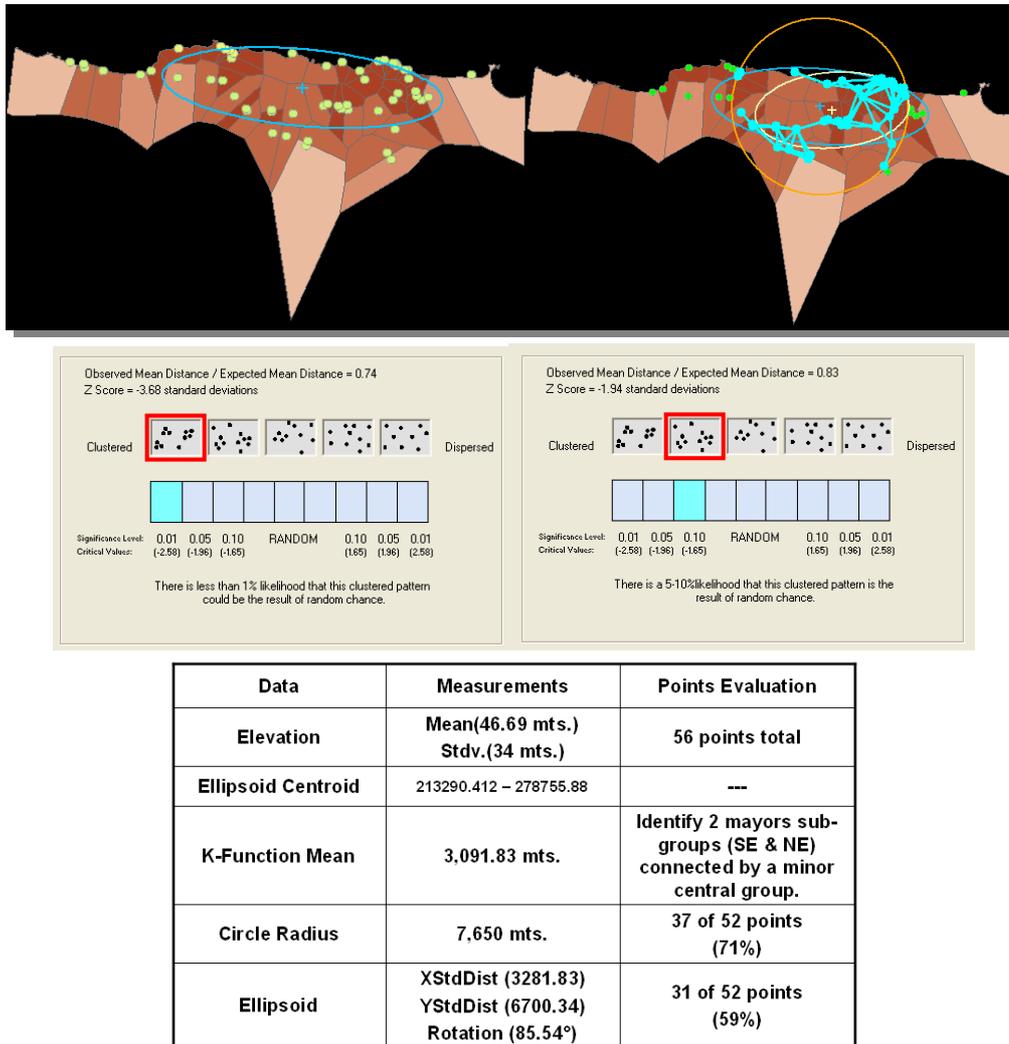


Fig. 4.5. Cluster Analysis Results. Top: Cluster exploration, Middle: Nearest Neighborhood results, and Bottom: Additional information.

North Center Cluster Results: Using a maximum of 56 points. The Average Nearest Neighborhood (detected by Visual Exploration Techniques) indicate that the study area present a clustered behavior with a Z-score value of -3.68 stdv. totally exceeding the confidence level of 95% for the critical values -1.96 and +1.96.

The first Directional Distribution (ellipsoid) state, that only 58% of the point's falls inside the ellipse. Hence, the data set does not present a normal distribution around the mean. An orientation of 95° (ellipse rotation) clearly relates the data set with the north center shore with a particular extent towards inside the island. The second Directional Distribution using 31 points (removing possible outliers) shows an increase of 1%, corroborating the results of the first Directional Distribution.

After the attempt for optimization, the second Average Nearest Neighborhood shift towards a more random cluster with a likelihood of 5-10% with a Z-score value of -1.94 stdv. inside the critical values. A third Directional Distribution is used only to state the direction for the cluster indicating rotation of 85° . The K-Function clearly indentifies two major sub groups (north-east and south-west) and one minor sub-group (center) of the ellipsoid. In conclusion, the optimization seems to offer better results than the Visual Exploration results.

4.5.5.3. Center Cluster

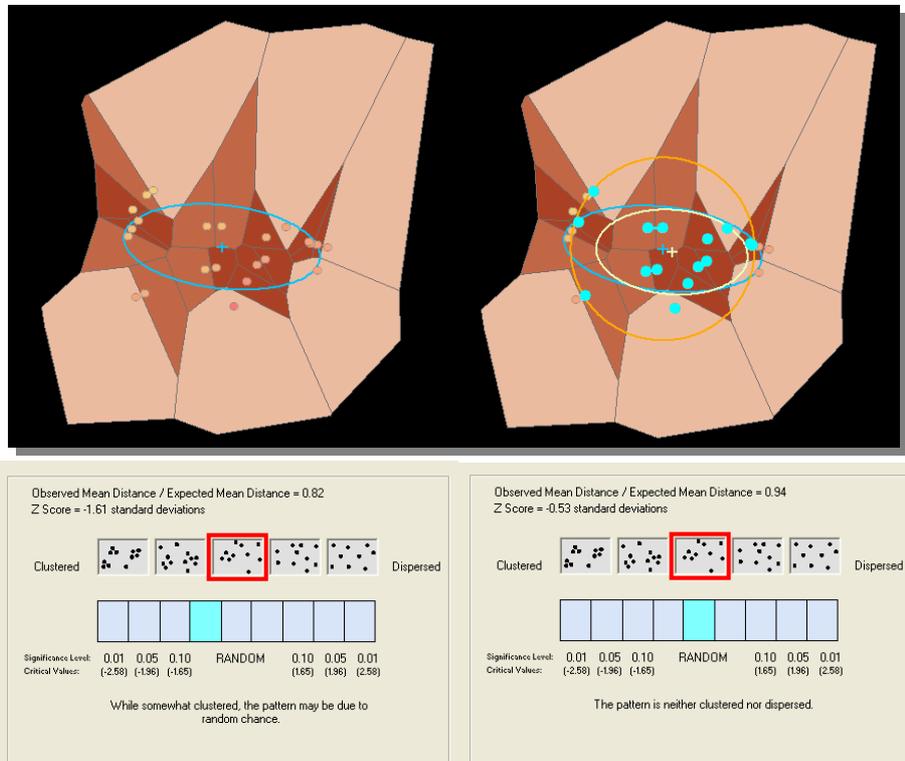


Fig. 4.6. Cluster Analysis Results. Top: Cluster exploration, Middle: Nearest Neighborhood results, and Bottom: Additional information.

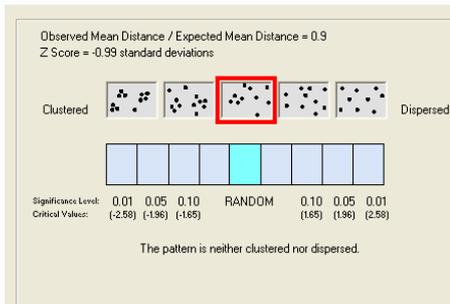
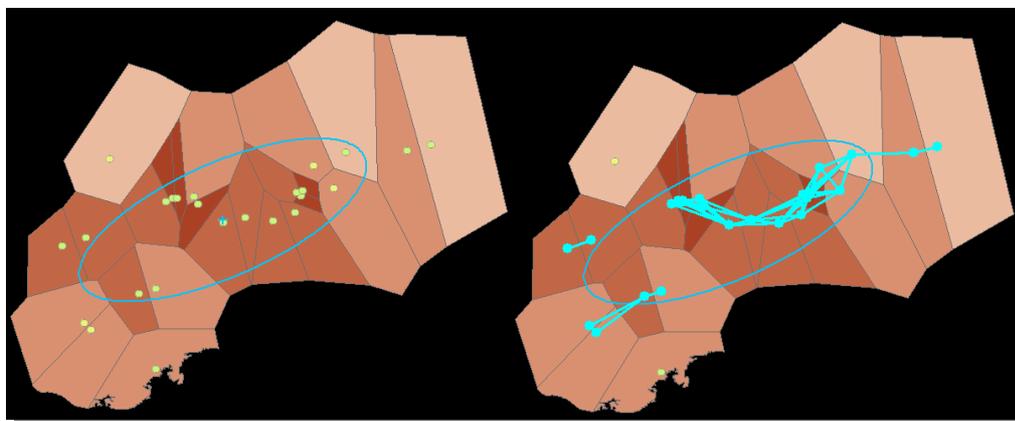
Center Cluster Results: Using a maximum of 23 points. The Average Nearest Neighborhood (detected by Visual Exploration Techniques) indicate that the study area present a random behavior with a Z-score value of -1.61 stdv. inside the critical values.

The first Directional Distribution (ellipsoid) state, that only 60% of the point's falls inside the ellipse. Hence, the data set does not present a normal distribution around the mean. An orientation of 96° (ellipse rotation) relates the data set with a possible movement towards the west and east direction.

The second Directional Distribution using 19 points (removing possible outliers) shows a decrease of 8%, with a rotation of 98° corroborating the results of the first Directional Distribution.

After the attempt for optimization, the second Average Nearest Neighborhood shift towards a more random cluster, with a Z-score value of -0.54 stdv. inside the critical values making a more solid statement about the distribution. A third Directional Distribution is used only to state the direction for the cluster indicating rotation of 94°. The K-Function fails to identify any particular sub-group. In conclusion, the optimization seems to offers better results than the Visual Exploration results.

4.5.5.4. South West Cluster



Data	Measurements	Points Evaluation
Elevation	Mean(130.50 mts.) Stdv.(5.22 mts.)	25 points total
Ellipsoid Centroid	133939.677 – 222644.732	---
K-Function Mean	3,594.73 mts.	Identify 1 mayor sub-group with a particular S-shape (River Related)
Ellipsoid	XStdDist (3006.31) YStdDist (8197.76) Rotation (65.84°)	17 of 25 points (68%)

Fig. 4.7. Cluster Analysis Results. Top: Cluster exploration, Middle: Nearest Neighborhood results, and Bottom: Additional information.

South West Cluster Results: Using a maximum of 25 points. The Average Nearest Neighborhood (detected by Visual Exploration Techniques) indicate that the study area present a random behavior with a Z-score value of -0.99 stdv. inside the critical values.

The first Directional Distribution (ellipsoid) state, that 68% of the points fall inside the ellipse. Hence, the data set seems to present a normal distribution around the mean. An orientation of 65° (ellipse rotation) relates the data set with a possible movement towards the south-west and north-east direction. The second Directional Distribution using 19 points (removing possible outliers) shows an increase of 4%, with a rotation of 75° corroborating the results of the first Directional Distribution.

Looking at the result, there is no need for an optimization attempt. The K-Function identifies one subgroup with a particular S formation resemblance to a perennial river crossing the archaeological sites (not presented in the display). In conclusion, a Visual Exploration technique was able to clearly describe the cluster.

4.5.5.5. South East Cluster

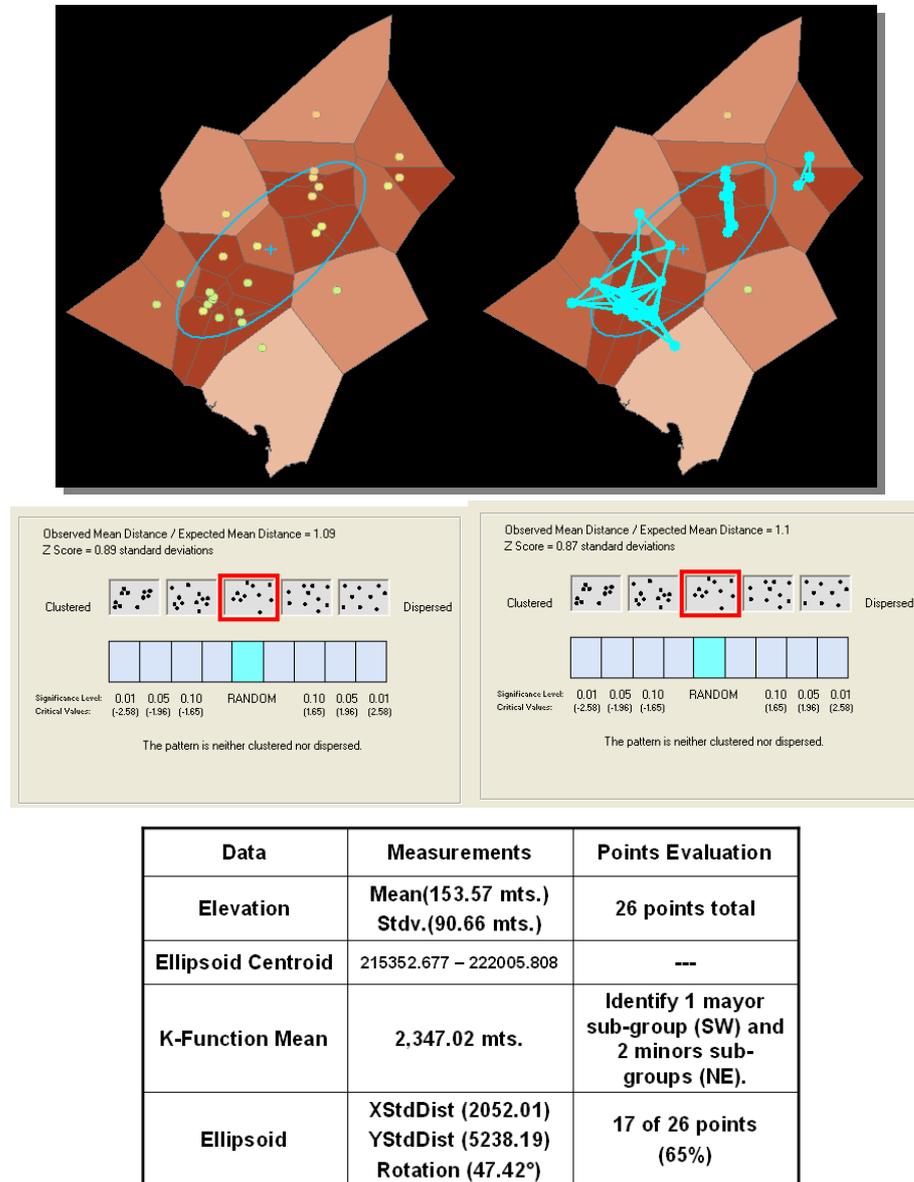


Fig. 4.8. Cluster Analysis Results. Top: Cluster exploration, Middle: Nearest Neighborhood results, and Bottom: Additional information.

South East Cluster Results: Using a maximum of 26 points. The Average Nearest Neighborhood (detected by Visual Exploration Techniques) indicates that the study area present a random behavior with a Z-score value of -0.89 stdv. inside the critical values.

The first Directional Distribution (ellipsoid) state, that 65% of the point's falls inside the ellipse. Hence, the data set cloud to present a normal distribution around the mean. An orientation of 47° (ellipse rotation) relates the data set with a possible movement towards the south-west and north-east direction. The second Directional Distribution using 21 points (removing possible outliers) shows an

increase of 7%, with a rotation of 48° stating that the data set does not present a normal distribution around the mean.

After the attempt for optimization, the second Average Nearest Neighborhood still indicates that the study area presents a random behavior with a Z-score value of -0.87 stdv. inside the critical values. A third Directional Distribution is used only to state the direction for the cluster indicating rotation of 42°. As no major improvements were achieved using the optimization model, the original dataset will be used for the K- Function. The K-Function clearly indentifies one major sub group (south-west) and two minors sub-group (north-east) of the ellipsoid. In conclusion, the optimization and the Visual Exploration seem to offers similar results.

4.5.6. Cluster Analysis Conclusions

After a revision of the results of the cluster analysis two clusters will be chosen for the input of the SMC. The south east cluster will be chosen because using visual exploration techniques and the optimization model produce similar results with a reliable confidence level. On the other hand the K-Function clearly associates sub-groups and outliers that are obvious by visual recognition.

Regarding to the second cluster (center cluster), both results are inside the critical levels. Even that the K- Function does not provide a clear results; the center cluster will be given priority as it comprises one of the most important archaeological sites in Puerto Rico island (Caguana Ceremonial Center). Hence it is worthy of further exploration.

4.6. GIS Analysis Model (attribute transfer)

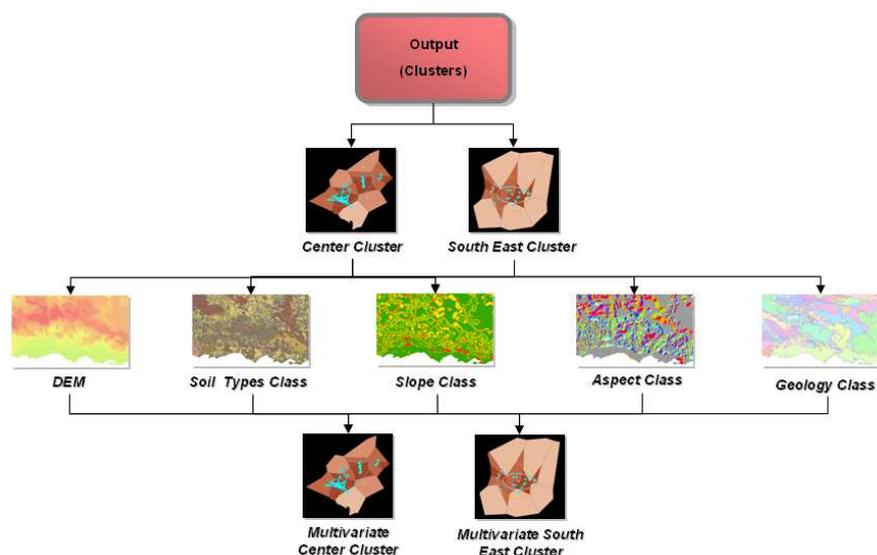


Fig. 4.9. GIS Analysis Model. Top: Cluster analysis results, Middle: classification layers, Bottom: Multivariate Clusters.

In order to capture the geographic variable for the chosen archaeological clusters, a simple GIS analysis was performed. DEM, soil, geology, slope and aspect comprises the impute variables for the SMTC environment. The soil types, slope and aspect classification was made following the “Guidelines For Soil Description (4th ed.)”. The geology classification was made using additional information (e.g. geologic period and class) on the metadata that was not available on the attributes layer. The DEM was generated by using the contours lines (10 mts.) for a 20 mts. sqr. grid cell. All the analysis was performed using Arc. GIS 9.2 and Erdas 9.1.

4.7. Independent and dependent Variables

Table 4-1. Data set in use. Orange color is given to the independent variable (archaeological sites).

<i>Field Name</i>	<i>Dimension</i>	<i>Behavior</i>	<i>Shape</i>	<i>Triad Framework Cognition</i>
Archaeological Sites	Temporal, Geometric	Object	Point - (0 simplex)	Time Base
Elevation	Geometric	Field (continuous)	Line - (1 simplex)	Location Base
Slope	Geometric	Field (continuous)	Location Base
Geology	Geometric	Object-Field-discrete)	Polygon - (2 simplex)	Location Base
Soli Type	Geometric	Object-Field-discrete)	Polygon - (2 simplex)	Location Base

Table 4-2. Attributes information (archaeological sites clusters).

Field Name	Data type	Measurement Level	Shape Inheritance	Data Characteristic	Graphic Variable Berlin's points and lines	Triad Framework Cognition
X-coordinate	Number	-----	-----	-----	-----	Object-Base
X-coordinate	Number	-----	-----	-----	-----	-----
Elevation (max, min, men, std.)	Number	Ratio	Point and Lines	Quantitative	Size	-----
Slope Class	Number	Ratio	Point and Lines	Quantitative	Size	-----
Aspect Class	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Soil Class	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Soli Type	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Soil Plasticity	Number	Interval	Point and Lines	Quantitative	Value	-----
Soil Permeability	Number	Interval	Point and Lines	Quantitative	Value	-----
Agrarian capacity	Number	Interval	Point and Lines	Quantitative	Value	-----
Culture Period	Date/Time	Ordinal	Point and Lines	Quantitative	grain/texture	-----
Stat Time	Date/Time	Ordinal	Point and Lines	Quantitative	Value	-----
End Time	Date/Time	Ordinal	Point and Lines	Quantitative	Value	-----
Culture	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Time of Existence	Number	Ratio	Point and Lines	Quantitative	Size	-----
Geology Class	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Geology Period	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----
Geology Type	Text	Nominal	Point and Lines	Qualitative	Colour, orientation and shape	-----

4.8. Semantic Multivariate Cube Implementation

4.8.1. Spatio-Temporal Questions

After the realization of the SMTC environment and retrieving the information needed for the case study, the three components for the multivariate spatio-temporal data will be addressed. The following sections will target the visualization as well the exploration of archaeological events using the developed prototype.

4.8.2. Visualizations

- *Exploring archaeological sites*

One of the first steps to explore any kind of information at hand is to establish the frequency of certain objects. In order to understand the archaeological sites distribution and culture diversity, it is necessary to address who and how many are present in the study area. A linked view of a simple frequency bar chart can provide the necessary information to answer these kind of questions. However, in this research that was not achieved. On the other hand, the SMTC still has the ability to suffice this need in others ways. As mentioned in Chapter 1, one of the main objectives of this research is to relate and extract meaningful information about archaeological events. Is in here that the new SMTC environment uses his ability to group and relate to answer this kind of questions. Fig. 4.10

shows a demonstration of such a case. The SMTC as well the STC uses the Space-time Path (STP) to group the events by culture. A Coloring by culture (STP) was applied to differentiate the events. The height of the cube is used to separate the groups according their Time of Existence. Clearly the user can see (top left image) the differences of grouping effects between the STC and SMTC environment. The advantage of the SMTC is that massive lines resemblance a high frequency node networks, providing the user the ability to recognize in a matter of second the information needed. Moreover, the user can corroborate the visual display by clicking the groups to gather the amount of participating events (see bottom image). The following results are given: Chicoid (5), Ostionoid (4), Elenoid (11), Elenoid/Chicoid (2), Santa Elena (2), Elenoid/Ostionoid (1), and Esperanza (1).

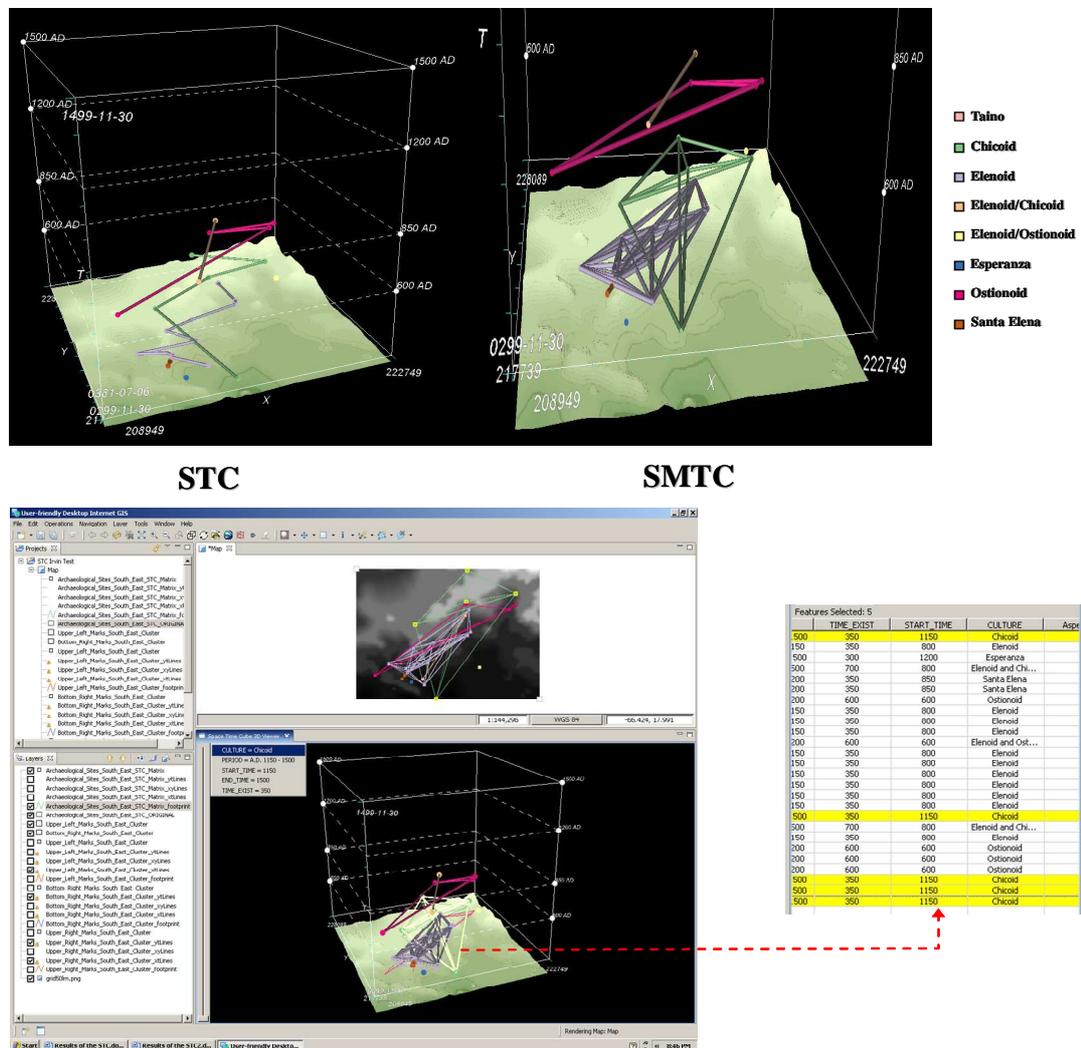


Fig. 4.10. Grouping effect. Top Left: STC, Top Right: SMTC and Bottom: Information retrieval options.

- *when + what* → *where*

In Chapter 3, possible combinations (box selection, filtering, etc.) to address the questions regarding to the where question were presented. However, the majority of these tools were unable to be implemented due to time constraints. Section 3.3.1 discussed how the graphic drawings techniques can adapt and change according any particular attribute in use. In this regard, the exploration of locations or sets of locations can easily be tackled using the location characteristic (municipalities, towns or neighborhoods) nominal values. Fig. 4.11, shows a demonstration of exploring archaeological sites at municipality level. Once the user sets the SMTC environment to group the events in this manner, noticeable patterns appear. Instantaneously the user can indentify Jayuja municipality as the one that presents a higher frequency (14 events) of archaeological events. The saturation of lines relating the group's events corroborates this.

It is important to mention that the part of study area is comprised mostly by Jayuya municipality. However, the technique is still effective to explore subsets of events. Similarly, the user can apply the same method (grouping) to move further the exploration, in this case at neighborhood level. In Fig. 4.12, an exploration of archaeological events is demonstrated. Notice how the network changes compared with the municipality exploration according users preferences. Not only can the user detect the neighborhoods like: Veguitas (5 events), Paso Palmas (6 events) and Coabey (4 events), as the ones with high frequency; but also similar characteristics that cannot be detected grouping by municipality level (see, information similarity box in both Fig.s). In the case of the municipality level exploration, the information box retrieves 6 characteristics (1 locational and 5 temporal) that describe the whole group. Conversely, in the exploration at neighborhood level, 10 characteristics (2 locational, 5 temporal and 3 geographical) are detected.

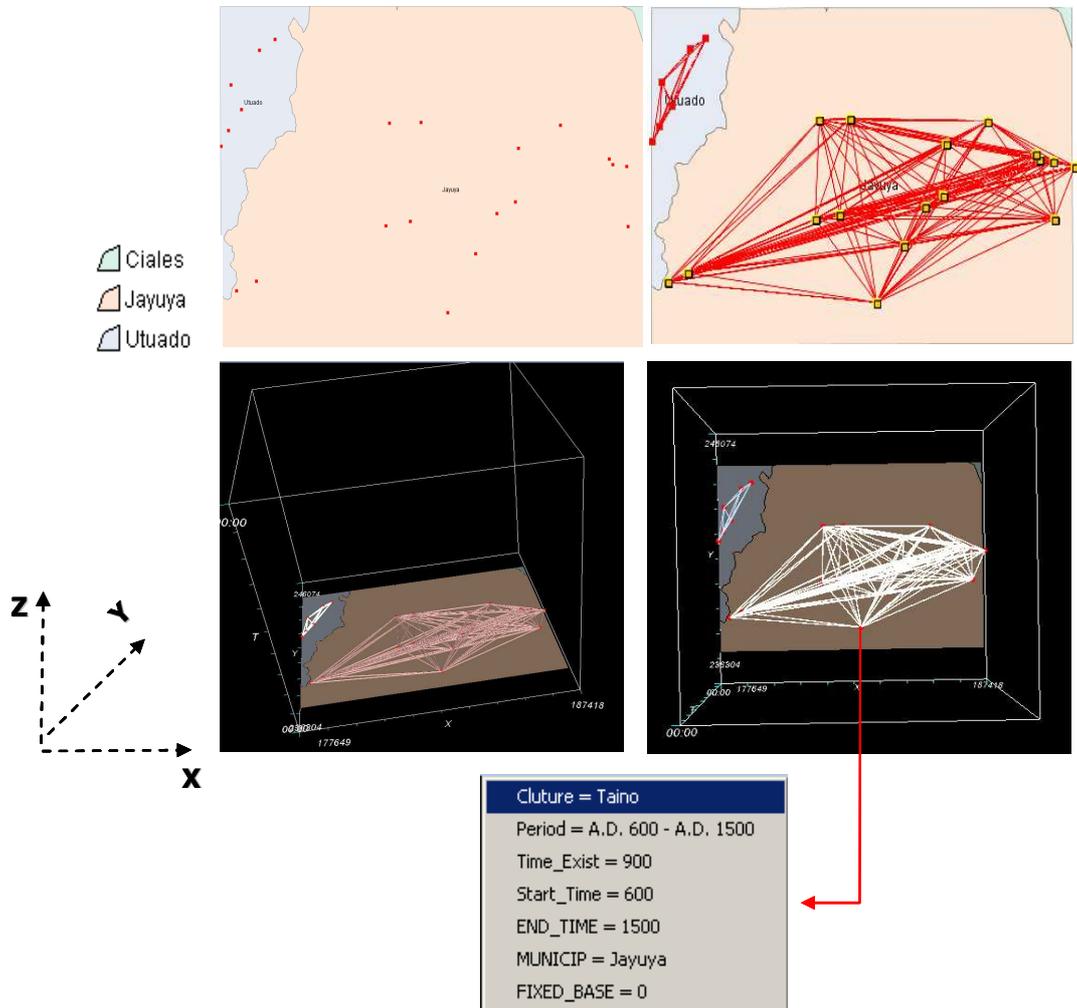


Fig. 4.11. Exploring archaeological sites frequency at municipality level. (Top Left: Municipality divisions 2D Map, Top Right: archaeological site semantic network by municipalities, Middle Left: SMTC, Middle Right: SMTC interactive selection, Bottom: semantic information box.)

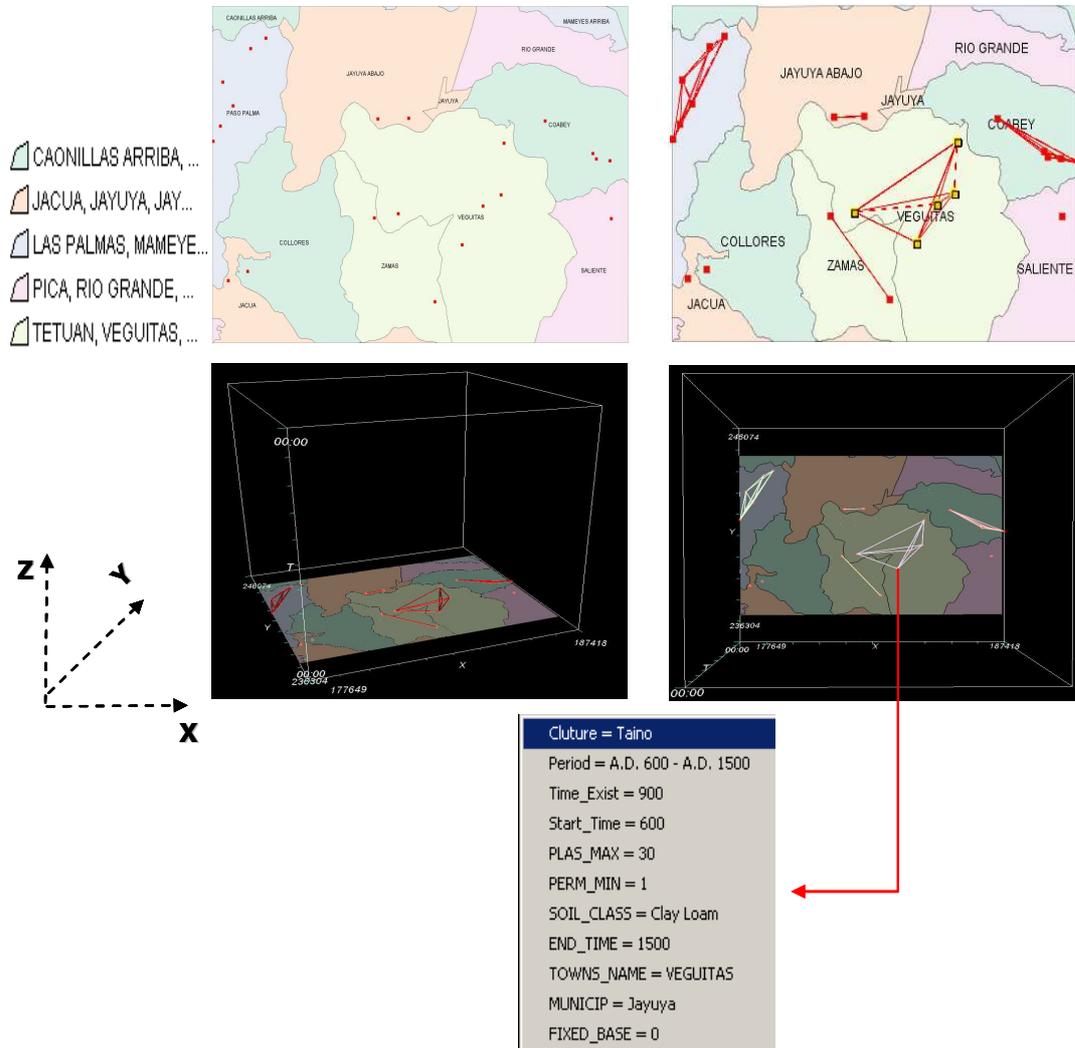


Fig. 4.12. Exploring archaeological sites frequency at neighbourhood level.). Top Left: neighbourhoods divisions 2D Map, Top Right: archaeological site semantic network by neighbourhood, Middle Left: SMTC, Middle Right: SMTC interactive selection, Bottom: semantic information box.

- *where + what* → *when*

Regarding to the question when, the SMTC uses the main capabilities of the STC (stations, time options, rotations, etc.) to depict the events in time. In here the user can assign any temporal, geographical and attribute characteristic to differentiate the objects (events) in time. As the SMTC was built taking the main frame of the STC the user can simply use both environments to gain insight about the time domain of the events. Additionally, a linked PCP view will offer the users the ability to quickly detect patterns on the data sets. This in some extent facilitates the users to identify noticeable attributes patterns further to explore in the SMTC/STC. A demonstration of how archaeological sites can be explored combining the SMTC, STC and Linked View (PCP) is presented below.

Fig. 4.13(a), shows the STC linked with PCP for the North Center Cluster (see section 4.5) exploration. The stations on the cube clearly depict all archaeological events in the region having similar appearance (Start Time), disappearances (End time) and durations (Time of Existence). The coloring of the station (see, legend key) states the Taino culture as the ones who dominated the area. Additionally the PCP (right side) corroborates these assumptions. In the PCP, the Objects ID (vertical axis) are used to separate the lines for identification purposes. It is clearly visible how all the events (lines) are cluster in the same Cultures Class, Start Time, End Time and Time of Existence. Hence, they are a good example of an extensive indigenous settlement compounded by Tainos.

In the same manner the South East Cluster was explore (see Fig. 4.13(b). A diversity of cultures (colors) ranging in different times (stations: start, end and length) can be appreciated on the cube display. The main advantages of the PCP as mentioned before are the recognition of patterns and the detection of characteristic behaviors. The PCP (right side) shows these dynamics quite clearly, and two noticeable patterns are detected. The first one at the Start Time with massive lines clustered at 850 AD. And the second on the End Time at 1200 AD. Following the results of the PCP we can assume that a significant culture interaction could lie between 850 AD. -1200 AD. A change on perspective of the cube and the use of additional gridlines (cube corners) for the specific time terms corroborate the results (see Fig. 4.13(c)). Even that clear statements have been give by the STC and PCP exploration, they still lack the ability to relate particular objects or sets of objects in time. Considering the grate amount of events is almost impossible for the STC to depict clearly those events with similar Start and End Time.

In order to further explore the dynamic behavior in time of the South East Cluster the SMTC environment will be implemented. Fig. 4.14(a), shows a demonstration of how the SMTC groups events by time. STP on the SMTC is used to link the events by Start (right image) and End Time (left image) and point colors to represent the class cultures of the events. Following the shape of the graph and the colors of the STP to differentiate the grouping times terms, interactions in time can be perceived (see, Fig. 4.14(b)). We can clearly see (bottom to top) that the first group of events (orange shape) on the Start Time shifts the order in the End Time by positioning on top of the second group events (yellow shape). Moreover, the inclusion of additional events on the green shape (End Time) state the longer duration of some archaeological event that participates in different Start Time. That corroborates in some extent events possessing a long time term interaction or simply that a mix of culture have been takes place. Not necessary the SMTC have to use two cubes to explore the

information in hand. Moreover, the user can simply combine the two versions into one homogeneous cube. Of course, is important to mention that this technique is not suitable for unskilled users, due to the complexity involved. Fig. 4.14(c), shows the results of such combination. A short time interaction (50 years) is detected between 1150 AD. – 1200 AD. This is a remarkable advantage that surpasses the STC and PCP capabilities to explore the information.

In order to tackle those events participating in different time (Start Time and End Time) another version of the SMTC is generated. In Fig. 4.15, events are grouped by Start Time and the STP is colored by End Time. Two events (Elenoid/Chicoid culture) starting in the second group order (grouping by Start Time) are shifted to the third group order (grouping by End Time). Others two events (Santa Elena culture) belonging to the same group in the Start Time are shifted to second group order (grouping by End Time). This effect can be seen by the change of the STP colors (purple lines) in the top right image. As well an isolated event (Esperanza culture) in the Start Time is added to the third group order by End Time.

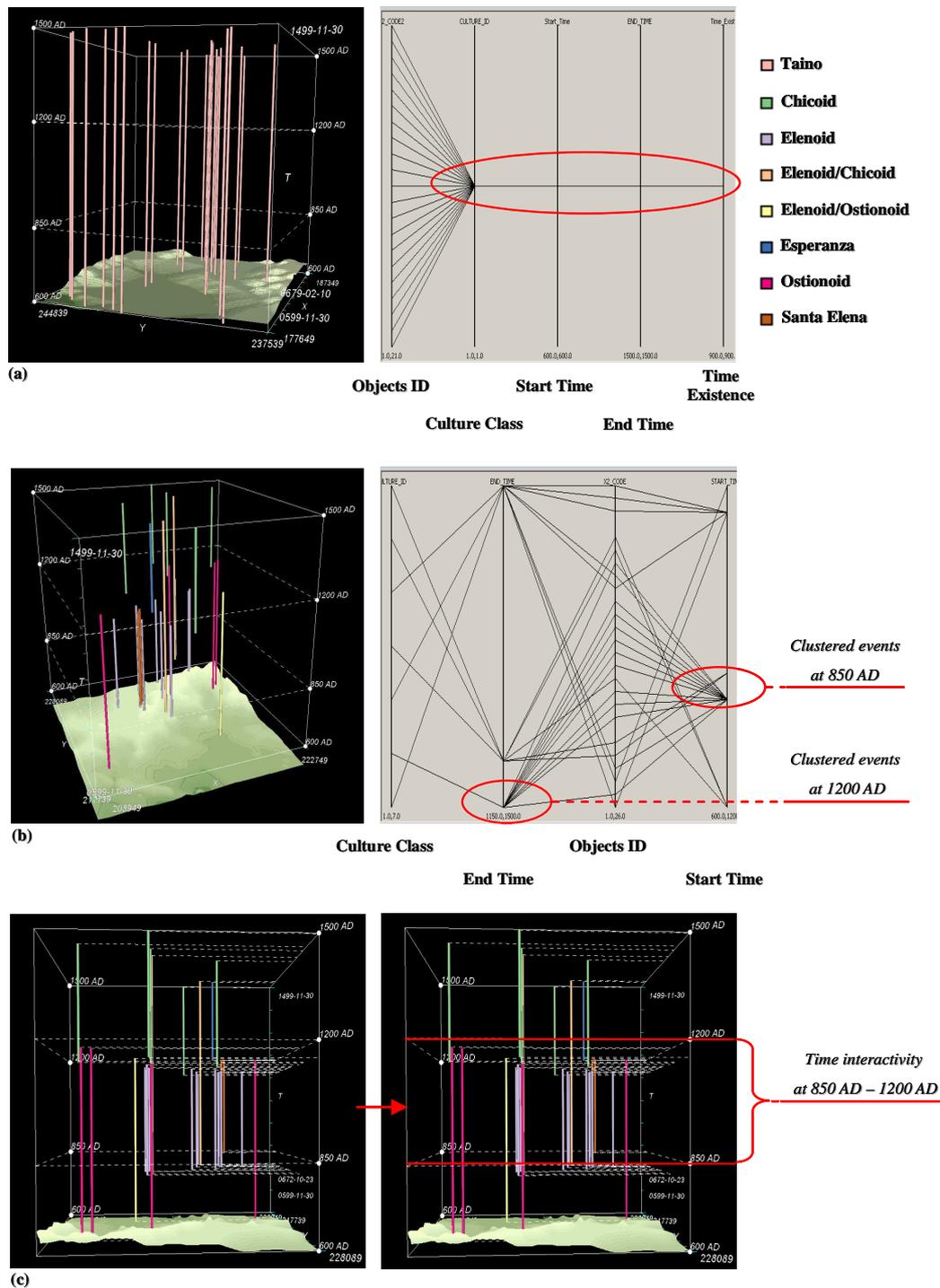


Fig. 4.13. Archaeological sites time exploration. (Fig. (a): Center Cluster (Stations and PCP), Fig. (b): South East Cluster (Stations and PCP), Fig. (c): South East Cluster (Front View).

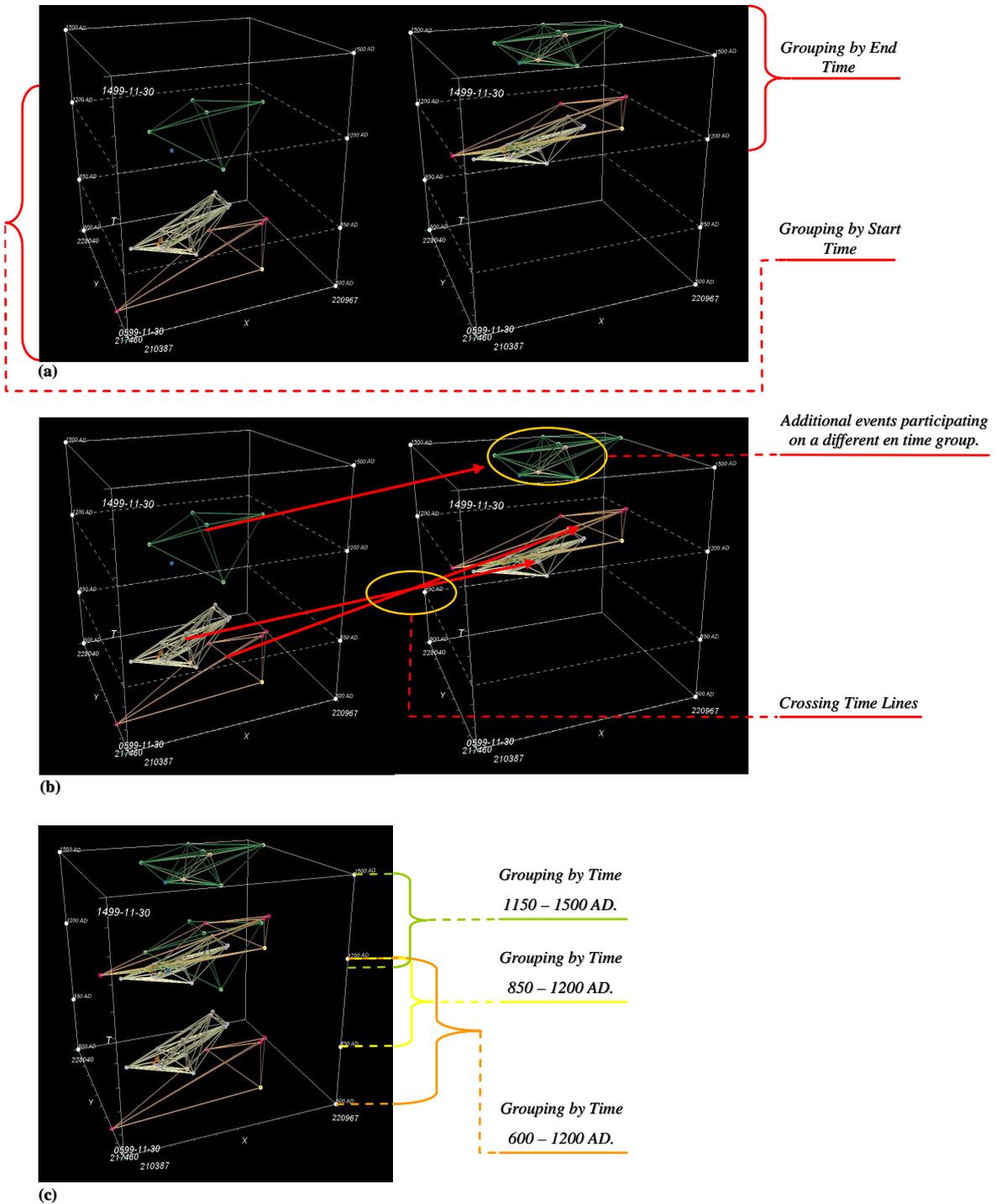


Fig. 4.14. Archaeological sites time exploration (South East Cluster). (Fig. (a): Grouping by Start Time, Fig. (b): Grouping by End Time, Fig. (c): Combined Cubes.

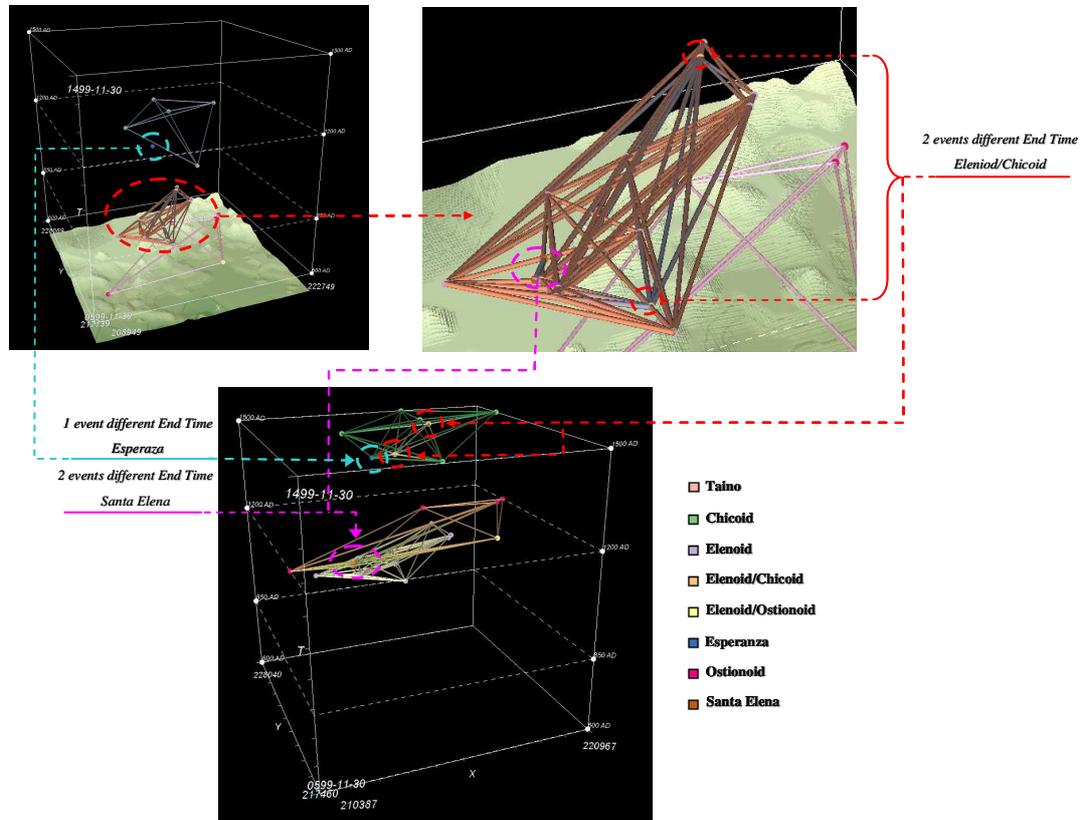


Fig. 4.15. Archaeological sites time exploration (South East Cluster). (Top Left: Grouping by Start Time coloring by End Time, Top Right: Second order Start Time group, Bottom: End Time group intrusion detection).

- *when + where* → *what*

Concerning the question what, the SMTC relies mostly on the interpretative capacity of the users and the available variables to be include in the semantic exploration. Using the existing graphic variables (e.g. colors, thickness, size, etc.) and the display options (e.g. stations, STP, voxels, points and lines) the user can set the desired combination/s to extract meaningful information or patterns about the events at hand. As mentioned before, the complexity of the SMTC environment especially at the time to explore the multivariate characteristics of the data make it difficult to handle for unskilled users. The multivariate environment of the SMTC is only limited by how well the users understand the data and their mental limitations to recognize the associations displayed. In order to prove and demonstrate the capabilities of the SMTC, an exploration of the multivariate characteristics of the Center Cluster will be presented.

The archaeological Center Cluster was selected for two important reasons. The first one is that all events resemble the same indigenous culture. Hence, massive relationships are stated by the SMTC environment. This advantage provides what is needed to explore other characteristics (e.g. geographical) on top of the semantic culture connections. Sorting and grouping the lines (STP) will lead to noticeable patterns using the remaining variables.

The second reason lays on the STP (updated version of the STC) simple connection of all events as a homogeneous (a simplify version of the semantic networks) path, facilitating a prior multivariate exploration of the events. Fig. 4.16(a) upper left, shows this effect. A total of 5 variables will be used. The STP is grouped by culture, colored by aspect and thickness by agrarian capacity. The nodes on the graph (archaeological sites) are colored by slope and elevation values as height of the cube (Z-axis).

Visualizing the archaeological events in this manner the following results can be stated. As mostly the display is saturated with visible lines it can be said that the study area concerning the archaeological events reflect rich soils for cultivation. The coloring of the STP (SMTC) reflects a strong likelihood for flat areas (green colors) as well the coloring of the node by slope (bright oranges). The position of the events regarding to the valleys on the area can be detected by the horizontal position of the STP connecting the events as well the vertical reference lines (not displayed for aesthetics purposes). The SMTC not only allows the user to detect and depict the multivariate characteristics of the data furthermore; corroborate the display retrieving the information involving those relationships. Fig. 4.16(b), show some of the results of the exploration inverting the colors to target the likelihood in slope values.

Due to time limitations for this research and the complexity involved, the exploration of the multivariate spatio-temporal characteristics for the South East and Center Clusters was not possible.

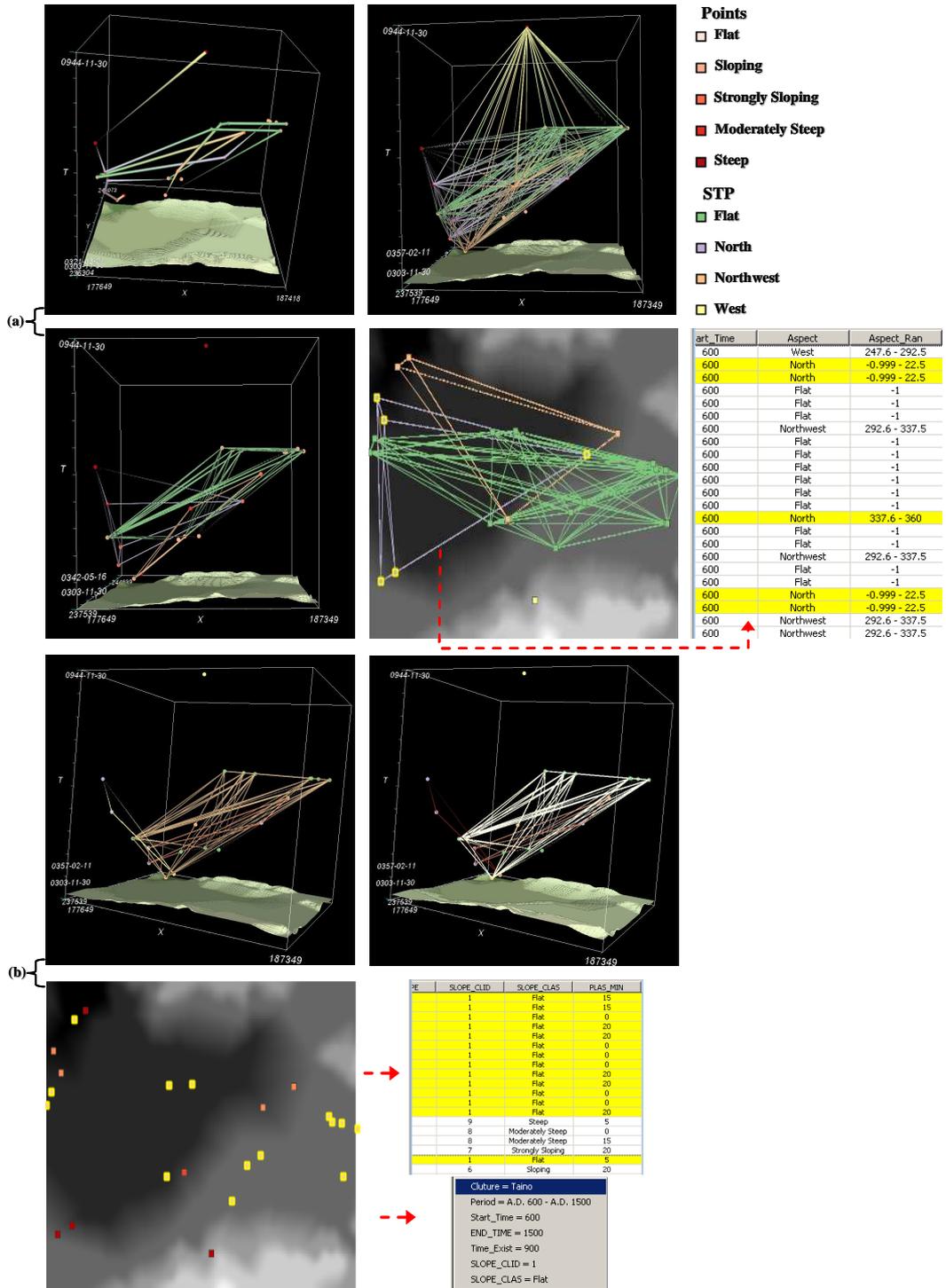


Fig. 4.16. Archaeological sites multivariate exploration (South East Cluster). Fig. (a): STC and SMTC depicting elevation, slope, aspect, agrarian capacity and grouping by culture. Fig.(b): Inverted colors for slope and aspect.

4.8.3. Limitations

As well as a range of significant benefits, there are also some significant issues which require discussion. The following lists present an overview of the problems and limitations of the actual version of uDig platform, STC extension that in some extent affects the research study.

Platform (uGig.) and STC:

- 1) Applying changes on the STC environment preferences options, previous colors settings and sizes resets.
- 2) Removing or changing class colors are not possible when a predefined color schema is used.
- 3) The size of the footprint (STC/SMTC) can not be combining with classification colors.
- 4) The compose time option box does not remember any previous settings and assigning randomly any numeric field.
- 5) Classification window (colors) does not remember any previous settings.
- 6) The background map (2D) color does not work always stays white.
- 7) Activation of the layers can only be done one by one.
- 8) Stations are only visible when the users place the layer on top of the others.
- 9) Wongs time measurements are computed if the Stations are called first. The software needs others layers to establish the computation for the Stations (End Time - Start Time = Duration).
- 10) Time clipping only works one time (first try).

PCP (plug-in):

- 1) All layers are activated; deactivation has to be done one by one.
- 2) Minimum and maximum value in some cases does not match the information of the layers.
- 3) Only one line can be selected. No way to select groups of lines.
- 4) The highlighting effects that's links the 2D map and PCP present confusion selecting additional objects on the 2D map.
- 5) No attribute information of the lines can be seen as well no references horizontal lines on the vertical axe.
- 6) Not possible to rearrange the vertical axes.

4.8.4. Summary

This chapter has presented and implemented the STC (updated version) and the SMTC prototype for the exploration and visualization of archaeological sites. A new environment, tools and functions have been developed in order to facilitate the investigation of such rich data set. However, an evaluation still needed to corroborate the usefulness of this prototype. The next chapter will touch this matter in detail.

5. Evaluation

5.1. Introduction to the focus group evaluation

In order to evaluate the implementation of the functions and tools discussed in the previous chapter, a focus group evaluation was held. A software evaluation allows the gathering of new ideas and develop the existence ones. Moreover, allows envision and enlighten to a software designer in order to satisfy users needs and create a more robust application. The greatness of a software only depends on the combination of its utility and its straightforward environment. That only can be achieved by a continuous development process and evaluation

“A focus group study is a carefully planned series of discussions designed to obtain perceptions on a defined area of interest in a permissive, nonthreatening environment. Each group is conducted with six to eight people by a skilled interviewer. The discussions are relaxed, and often participants enjoy sharing their ideas and perceptions. Group members influence each other by responding to ideas and comments of others.”

C. Abbott and P. Eubanks (2002).

Since the 1950s, focus groups have been targeted by the marketing field as a way to gain insight into the customer preferences and ideas. Not only focus groups has been prove his utility to design a perfect TV commercial for products (as we seen in today's television) that others non-related fields are taking the approach to validate or to evaluate their products. For example, in the field Information System one of the main objectives is to validate and test new software's before coming out to the public. The use of focus groups to performed specific task (such as a usability test) is a common method implemented nowadays. Almost every day a computer user finds that a new version of his software is available and free to test. Today's technology now allows the users to freely access the new product, test it, and submit their comments. Video conferences, groups gathering, and isolated web evaluations are the most typical form of focus groups used. On the other hand, other evaluation methods like questionnaires and discussions still sometimes form part of focus groups.

One of the major limitations for the current research is that the current state of maturity of the platform used for this prototype (the STC is implemented using the open-source GIS uDig) does not allow the carrying out of a usability test. Too many inconsistencies still exist (e.g. classification problems, lack of computational power, graphics card deficiencies, bugs, code errors, etc.) that will affect the users task presented.

Questionnaires are another suitable solution to gather user opinions. On one hand, participants are constrained by predefined choices and saturated questions that they in some cases do not clearly understand. On the other hand the time taken to fill the questions and comments tend to frustrate the

participant causing them to lose their interest. Since questionnaires are closely matched to the user's tasks, this method does not offer a suitable solution in this case.

Another method of focus group evaluation is the study of a series of discussions. A discussion group is characterized by retrieving the response of multiples groups into a form of a meeting session. In here the evaluator can group the participants in different categories according their expertise in order to provide a relaxed environment, a degree of trust between participants, homogeneity in opinions and as a result, deeper insights and responses. One of the main advantages of using a discussion session is that the evaluator can move forward to other issues when the participants are repeating the same ideas (saturation point). In that regard, the evaluator does not have to ask questions directly to each participant but, make short statements in relation of different opinions. This accelerates the evaluation session avoiding the loss of interest of the group.

For the reasons mentioned above, an evaluation discussion will be taken in consideration in order to get the impressions, comments and ideas about the prototype developed.

5.2. Discussion session and group members

The focus group evaluation took place at 11:00 AM – 12:00 AM in room 0.142 of ITC. The discussion plan was divided in three main parts:

- 1) First part; the evaluator gives a presentation for 20 minutes introducing the research and concepts (e.g. STC, Graphic Drawings, PCP, and SMTC prototype). The evaluator presents parts of the case study results, showing the main tools and options of the prototype (SMTC).
- 2) Second part; questions are made base on the results to retrieve participants opinions and interpretation. Discussions and opinions are recorded by the evaluator. An open discussion was allowed in order to let the participants share their opinions.
- 3) Third part; a live demonstration was given in order to let the participant to appreciate the flexibility of the prototype and change their previous opinions.

The group's members were separated in two main groups formed by GFM.2. MSc. candidates. The group was comprised of 10 participants, selected based on guidelines described in Abbot and Eubanks (2002). The first group was comprised of 5 students that to some extent have a similar geovisualization topic research (e.g. animation, interactive web mapping, STC, etc.). The second group was mainly formed by 5 students in different research areas (remote sensing and programming).

Forming the groups by different specializations can provide a diversity of opinions regarding the prototype to be evaluated. For example, is expected that the first group understand better the geovisualization environment and redirect their opinions on the effectiveness of the prototype to explore the information; assuming that they posses a more comprehensive understanding on geovisualization solutions and spatiotemporal data as well multivariate data. Regarding the second group, it is expected that they will focus their attention on the geovisualization environment itself and its ability to present or disseminate the information; assuming that they are limited by what they have been seen about geovisualization along the previews courses in the ITC Institute. As mentioned

before, in order to let the participants interact and share their opinions both groups were held in the same room but at different tables.

All information captured with a digital voice recorder, and transcribed and summarized at a later stage.

5.3. Discussion session objectives and questions

The discussion session was performed in order to evaluate the proposed prototype and cover one of the objectives of this research. The immediate objectives of this evaluation are:

- 1) To measure (by opinions) the visualization power of the SMTC environment.
- 2) To assess how well the SMTC answers multivariate spatio-temporal questions.
- 3) How well the participants understand the proposed model.
- 4) What are the participants preferences regarding the available tools and options.

Base on the objectives the following set of themes and associated questions were posed to the group members:

- 1) **First theme;** oriented on the STC environment itself in order to introduce the frame in where the SMTC takes place. Participants were asked about the appearances, disappearances, duration, and class cultures classification of the archaeological sites in the study area.
- 2) **Second theme;** using the previous example in conjunction with a liked view (PCP) participants were asked to associate the images (STC and PCP) and detect patterns. What was intended here is that the participants appreciate the multivariate behavior of the archaeological clusters in order to minimize the visualization impact of the SMTC for further questions.
- 3) **Third theme;** involve the question where, when and what. Participants where asked to identify groups of clusters (space and time) and using a specific attribute of the points (culture) describe the kind of interactions. This will provide the participants a basic idea of what kind of cultures (events by color) will participate in a grouping by time on the SMTC.
- 4) **Fourth theme;** focuses on a comparison of the STC and SMTC regarding their ability o visualize groups of events. Two pictures (STC and SMTC) of the same archaeological cluster grouped by culture were presented independently. The participants were asked to reflex about the frequency of the archaeological sites culture (for each picture) spreading around the landscape. Subsequently, the two picture where showed together and the participants where asked to state their likelihood. The intention here is to shows the participant how the SMTC establishes the associations between events as well the grouping effect.
- 5) **Fifth theme;** once the participants are familiarized with the SMTC grouping method; a question regarding time interaction was made. Two pictures (using the grouping function by start and end time) were shown in order to let the participants visualize possible interaction in time between the archaeological event groups.
- 6) **Sixth theme;** the last question concentrated mainly on the SMTC environment. Screenshots were presented to the participants in order to demonstrate how the STC and the SMTC are combined to explore the information. The intention here is to show the full exploratory potential of SMTC environment.

Once all the opinions based on the themes and questions were concluded; an open discussion between participants took place. The participants talked amongst themselves and shared their opinions. Moreover, a live demonstration of the actual prototype was presented to the participants. Prior settings were being prepared using various combinations in order to speed up the processes and present the SMTC environment. At the end of the demonstration, participants were asked for last remarks. The following section presents all the results.

5.4. Focus group results

5.4.1. General results by theme

As mentioned before, the most important issue in a discussion groups is the multiple perspectives of the participant regarding the presented topic (in this case a geovisualization prototype). The participants' comment, discuss, and reflect about the themes presented. Consequently participants offer their suggestions, ideas, and preferences for certain tools. Based on such responses and on the richness of the data collected, preliminary results are obtained that in some extent can help or attract the attention of others similar research. However, some limitations are still present. It is important to mention that running a discussion session, not necessarily covers all the requirements needed for a prototype evaluation (usability test). For example, groups' representatives and the number of questions are often too small to achieve a full evaluation within the session. Hence, the sample of participants from which one retrieves the information does not allow for any significant statistical analysis. On the other hand, with the information gained in the discussion session, the evaluator may have a good impression about the state of the prototype, allowing him or her to make clear statements to terminate or go on with the development.

In order to maintain the structure of this chapter, the results will be presented by their respective themes as in the previous section.

- 1) **First theme results;** regarding the first cluster (see Fig. 4.13(a)), the overall participants (both groups) appeared confused when events possess similar start, end, and duration time(s). None of the participants were able to draw concrete conclusions. Some of the participants mentioned that the screenshots appeared unfinished and that there might be some possible data errors. Conversely, when the second cluster was presented almost all the participants were able to clearly identify and understand the appearance and disappearance of the events. Eventually, they assimilated a lot of information about the duration of the events using the cube grids and establishing the differences in duration as well culture classes (using colors to differentiate the cultures) (see Fig. 4.13 (b) and(c)).
- 2) **Second theme results;** the first group quickly began to gather information about the multivariate behavior of the events at the first glance of the PCP (see Fig. 4.13 left side). They also mentioned problems in tracking the lines if the order of the vertical axes is not preserved. On the other hand, participants of the second group don't seem to understand well how the PCP works, although they did state clear comments about the absence of the specific values of the (PCP line) events as well the division of the classes into the vertical axes. The whole second group agreed on the inclusion of the colors lines into the PCP that matches the STC, and both groups agreed that the addition of the PCP colors and style symbol for the lines would be beneficial. After some minutes discussing the PCP and the

STC associations, both groups were able to identify patterns of connected lines (PCP) related to the STC environment.

- 3) **Third theme results;** in relation to the what, where and when questions the participants do not identify any major problems (see Fig. 4.13(c)). By this stage they are fully aware of the time, location and attribute domains of the cube, and even for partially occluded events, the participants easily associate the colors of the classification to gain information for those that they could not see clearly. Suggestions and comments mainly concerned the software (uDig) environment itself and the way to present the information. For example, some of the participants were attracted by the DEM presented and began trying to query the elevation values of the events, and wondered why the evaluation did not have an interactive presentation (animations, movies or series of snapshots) to visualize the events.
- 4) **Fourth theme results;** on the subject of the grouping method (see Fig. 4.10), the first group participants quickly identified the Elenoid culture as the events that dominate activity in the study area. Also, they noted that over time there was a spreading of the other event clusters. On the contrary, the second group of participants seemed to be mistaken by the extent of the lines, and gave erroneous statements about the frequency of the events. After an explanation of the results, the participants were encouraged to state their overall impression of the STC and SMTC grouping effects. In total, six of the participants (one from the first group and five belonging to the second group) chose the SMTC as the best environment to identify relationships and interactions between sets of objects. Four participants belonging to the first group chose the original version (STC). Two participants did state that both environments can depict the objects equally well, and that only the groups are influenced by how they are presented (e.g. grouped, colored, etc).
- 5) **Fifth theme results;** regarding the time grouping question (see Fig. 4.14), the first group identified the effect as well as the relationships in time involving the cultures in a matter of seconds. Temporal clusters were also identified and described. The group noted the spaces between time clusters and quickly drew concrete associations between duration and interactions in time using grouping, shape, and color effects. Comments about the inclusion of events in other groups (using end time grouping) were given, but no concrete interpretations were made. Suggestions about the inclusion of different symbols (for the point events) were made in order to better distinguish them. The second group still had problems understanding the geovisualization environment, many making vague comments about the cube perspective. This is thought to be partly due to their lack of geovisualization knowledge. This group made no significant comments about the temporal clusters or the absence or inclusion of some events in different time grouping graphs.
- 6) **Sixth theme and open discussion results;** mainly the participants were interested in the amount of variables that the SMTC can depict (see Fig. 4.15). In general there were positive comments about the interactivity, rotations, and 'stations' combined with semantic network (though only regarding the aesthetics). Participants were impressed by the way that the SMTC can filter large datasets and build a massive amount of relationships according to any attribute of interest in the dataset. The highlighting, selection of groups and similarities information box was considered satisfactory by the participants. Regarding the PCP, there were a lot of similar comments about the dislike of a saturated variable (on the vertical axis), consequently leading to massive lines spreading on the display. Conversely, the participants expressed their like when two or three variables were depicted on the PCP. In general there were positive comments about the hill shade (DEM) and the ability to extrude a feature using an attribute. The participants were interested in including a "layer

draped" over the surface (hill shade) and an automatic generation of an interpolated surface that resembles the attributes of interest in order to follow and track the values on the cube plane.

5.4.2. Synopsis

Based on the result of the focus group evaluation, both positive and negative comments were gathered. The overall results are that the SMTC as well as the STC now seems to possess what is needed to explore in some extent the multivariate spatio-temporal characteristic of the data. It is important to mention that only a subset of functions identified in Chapter 3 section 3.7. were implemented and of these only (stations, colors, grouping, and the network itself) were demonstrated/evaluated in the focus groups meeting. Others, like the combination of the stations and the network were not presented in order to avoid participants' confusions.

Participants generally put more emphasis on the ability of the SMTC to display and manage complex relationships (e.g. groups of objects and selection). However, given the feedback received, the SMTC environment does appear to be too complex for common users (those that lack knowledge in geovisualization solutions) to understand. Below the most significant issues are summarized:

- 1) The uses of colors classification (e.g. stations, points and lines) increase the levels of pattern recognition.
- 2) Linked view (e.g. PCP) seems to boost the user's patterns recognition (after familiarization).
- 3) PCP colors, lines symbols, and axes information must be implemented.
- 4) The inclusion of surfaces on the plane of the cube has a tremendous visual impact, and should be considered very helpful to information retrieval.
- 5) Additional options regarding point's symbology must be implemented (SMTC and STC).
- 6) The grouping effects of the SMTC surpass the capabilities of the STC to associate and differentiate objects (in this case in time).
- 7) The environment itself appears attractive (e.g. rotations, stations, 2D and 3D views) for all the users, capturing the participants' full attention and contributing to additional insights for the selected dataset.
- 8) Highlighting, similarities information box, filtering effects and group selections were evaluated as practical and useful extensions by the participants.
- 9) The SMTC has a greater flexibility to depict and manipulate multivariate spatio-temporal relationship than the standard STC.

5.5. Summary

This chapter has presented the results from a focus group evaluation of a new geovisualization environment (SMTC). Considering the results of the previous section, multivariate spatio-temporal data is well suitable for exploration by both environments (STC and SMTC). On the other hand, the SMTC has demonstrated his extraordinary capability to adapt, explore, and displays the multivariate characteristic of the data taking the advantages of the STC main functionalities. Based on the results it is expected that further researchers put more emphasis on investigating and implementing the remaining functionalities of the SMTC in order to complete and achieve a more robust application for data exploration.

6. Conclusions

6.1. Introduction

In the previous chapter a discussion of a focus group was held in order to evaluate the developed prototype. This research has propose and design a new geovisualization solution environment with the main objective of explore and retrieve meaningful information for multivariate spatio-temporal phenomena concentrating on the richness of the inquired data. Base on the actual technology and platforms (uDig.) available, the SMTC was build. During the research time, new ideas, problems and solution arises that in some extent affect the study research.

6.2. MSTC (Multivariate Semantic-Time Cube)

The SMTC is a geovisualization environment with a solid purpose to visualize the multivariate spatio-temporal relationships of the data. In the SMTC, users can state, visualize, explore and retrieve the relationship of geospatial information. The 2D and 3D environment as well the PCP linked view comprised the main visualization options of the SMTC. However, as mention in Chapter 4, the man capabilities of the STC are inherit in the SMTC providing the users the ability to combine both geovisualization environments. That for the simple fact, that the STC was created exclusively to explore the what, where and when component and the SMTC for multivariate spatio-temporal data. Due to the complexity involving semantic relationships (SMTC), the STC still needed to facilitate a prior exploration. For example, in the SMTC prototype; relationships between objects can be seen as lines interconnecting them. Applying any graphic variable available facilitates the user to indentify what characteristics are playing in the relation. Hence, the questions regarding the multivariate characteristics (what) lays on the connections and the object itself. Chapter 4 “Case Study”, corroborates this statement.

Advantages:

- 1) SMTC provide the necessarily flexibility to visualize, explore and retrieve information about multivariate spatio-temporal data.
- 2) SMTC surpasses the ability of the STC to associates events in time, location and context.
- 3) The SMTC exploration in time totally exceeds the STC environment. Particular or groups of events can be detected and compare more easily in different time terms.
- 4) Additional information can be seen faster with the use of the semantic information box.
- 5) The datamining method implemented as a new footprint with semantic information, enhance both environments (STC and SMTC) for exploratory geovisualization purposes.
- 6) The 2D now provides an alternative visualization displays to explore geospatial data (footprint exploration).
- 7) The addition of coloring classification and gradient effects (3D view) in both views increases the aesthetics of the geovisualization environment as well the exploratory capabilities. Different objects with different attributes are easily detected following the gradient change of the lines (STP).
- 8) The amount of combined variables is only limited by the user choices and computational costs.
- 9) Easy to combine multiples cube into one version (see, Chapter 4).

Disadvantages:

- 1) Not suitable for unskilled users.
- 2) The complexity of the environment is hard to understand if the user is not familiar with spatio-temporal, multivariate and drawing graphics concepts.
- 3) The SMTC prototype requires a tremendous computational power (e.g. graphics card, memory, etc.).
- 4) Occlusions problems appeared if too many relationships are visualized.
- 5) Not suitable for large data sets.

6.3. Further Work

After the realization of the SMTC prototype, certain tools and objective were not possible to complete. The following list presents an explanation of them as well additional ideas that are the product of this research.

- **Relationship Degree:** one of the sub-objective on this research was to establish some kind of relationship measurements regarding the semantic results (network). As mention before, the time limit for the research doesn't allow to execute this goal. On the other hand, considering the case study (archaeological sites) the appropriate computation of a true relationship degree, lays on the cultural characteristics (e.g. artifact, culture hierarchies, diet, time, ceramic style, etc.). That implies an extensive incursion of the archaeological records and archaeological history of Puerto Rico that totally deviate the geoinformation science and goes more deep into the archaeological field.
- **Distance to rivers relationship:** this sub-objective was achieve but not presented in this research. The following reasons where: the use of a Euclidean distance seems not suitable to target the exact location of the most nearest Perennial River for each archaeological sites location. As we are talking bout past peoples and the most comprehensive transportation method was on foot, it should be consider the use of a friction surface to calculate the most suitable path to the rivers. As different location can involve the same river the only reliable solution is to associates the events into a similar location in where possible remains can be found (rivers deposits) regarding the computed location. Again this will involve the uses of a trace analysis model in order to compute the rivers flow (upstream or downstream) and the uses of satellite mages to detect the rivers depositions areas.
- **Datamining:** the datamining sub-objective, but not totally exploited. For exploratory purposes this method have provide its usefulness on this research. Some applications that cloud take advantage of this implemented method of extraction are the graphic drawings techniques. As the lines (footprint) can easily be extracted any graphic drawing algorithm can be modify to uses the lines and adjust the results for further exploration as a non-planar graph. That could be an interesting solution to minimize the graphic performance of the SMTC by adding a graphic drawing linked view.
- **Temporal Interface:** due to the complexity involved to develop the SMTC prototype this sub-objective was not achieved. Other reasons influences as well. For example, the carefulness requirements on developing and designing such tool in order to don't let it fall as

a specific tool for the exploration of archaeological sites. And as one of the main notions taken on this research was to develop tools and options to help investigation the multivariate spatio-temporal as well archaeological data, but more concentrating on preserving the tools in order to create a most robust application for futures researchers, this tool was not implemented but still worthy to investigate.

- **Filtering:** this tool can provide the ultimate touch for the STC as well SMTC to explore the information. As a standard tool for visualization and geovisualization environment, the filtering option can ease user visual impact, increase the performance displays (fewer graphic to depict) and help the investigator to focuses in specifics objects for further exploration.
- **Lexis Stations:** and interesting idea that comes from this research study is the combination of the Lexis Pencil visualization with the Stations and the SMTC environment. As mention in Chapter 2, Lexis pencil depict time as height of the pencil, attribute information can be seen on the faces of the pencil and the location domain is targeted by the point of the pencil into the surface plane. The STC (update version) can uses stations to address the time characteristics of the data using information about the Starting and Ending time. Additionally, thickness and can color can be applied to show at least two variable values. On the other hand, the SMTC prototype is able to sate relationship between objects nourishing with the STC environment. The fact is, that the three environments resemblance similar characteristic on the geovisulaization environment and the way to display the information. Is not difficult to think that Station in the STC as well SMTC can be modify dividing the duration of the events by the amount of variable in use. Adding multiples points with different times in the same location can provide the effect needed to create a multivariate station depicting with color each sub-section according the available variables. Furthermore, the SMTC can generate for each individual multivariate station section a graph interconnecting them stating the relationships for each segment. The only needed tool that is not available for the actual prototype SMTC and STC is the filtering, that will serve as an instrument in order to focus on any particular connection or station sections for the incursion exploration. Is important to mention that the main idea was proposed by Tominski et al., 2003, but no indication of implementation and development has been made. With the actual technology available and the knowledge gathered in this research a prototype for this case can be achieve in no time.

7. Bibliography

- Abbott, C., and Eubanks, P., 2002. A Primer on Using Focus Groups in Technical Communication. <http://www.stc.org/confproceed/2002/PDFs/STC49-00058.pdf>. Access on 15/02/2008.
- Aditya, T., Kraak, M.J.p. and Ormeling, F.J., 2007. National Atlas as a Metaphor for Improved Use of a National Geospatial Data Infrastructure. PhD thesis Utrecht University; Summaries in Dutch and English. ITC, Enschede, 228 pp.
- Ahmad, W., Khokhar, A., 2007. cHawk: An efficient biclustering algorithm based on bipartite graph crossing minimization”, In: Proc. Workshop on Data Mining in Bioinformatics, 2007.
- Andrienko, N., Andrienko, G., Savinov, A., Voss, H., and Wettschereck, D., 2001. Exploratory Analysis of Spatial data Using Interactive Maps and data Mining. In: Cartography and Geographic Information Science, Vol. 28, No. 3, 2001, pp. 151 – 165.
- Andrienko, N., Andrienko, G., Gatalsky, P., 2003. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, Vol.14, No. 6, Dec., 2003, pp. 503-541(39).
- Andrienko, N., Andrienko, G., 2004. Parallel coordinates for exploring properties of subsets. In : Coordinated and Multiple Views in Exploratory Visualization, 2004. Proceedings. Second International Conference. Volume , Issue , 13 Jul. 2004, pp. 93 – 104.
- Bertin, J., 1983. *Semiology of Graphics*, Madison, WI: University of Wisconsin Press.
- Blok, C.A., 2000. Monitoring change: characteristics of dynamic geo-spatial phenomena for visual exploration. In: Ch. Freksa et al. (Eds), *Spatial Cognition II, Lectures Notes in Artificial Intelligence*, Vol. 1849, Springer, Berlin, Heidelberg, 2000, pp. 16 – 30.
- Blok, C.A., 2005. In: *ICC 2005: Proceedings of the 22nd international cartographic conference: mapping approaches into a changing world*, 9-16 July 2005, A Coruña, Spain. International Cartographic Association (ICA), 2005. ISBN: 0-958-46093-0. pp. 10.
- Blakley, R., http://bldgblog.blogspot.com/2006/04/assembling-north-america_11.html access at 11/ 2007.
- Cuadrado Landrau, V.M., 2002., *Accessing geospatial data to support research, monitoring and environmental management activities : case study, Island municipality of Vieques, Puerto Rico*, ITC, Enshede, 77pp.
- Davison, R. and Harel, D., 1996. Drawing Graphs Nicely Using Simulated Annealing. *AMC. Trans. Graph.*, 15(4), pp. 301 – 331.

- Di Battista, G.D., Tamassia, R., Eades, P. and Tollins, G.I., 1994. 1993. Algorithms for drawing graphs: An annotated bibliography. Preprint, Dept. of Computer Science, Brown Univ., Providence, R.I., Nov. To appear in *Comput. Geom. Theory Appl.*
- Dykes, J., MacEachren, A.M. and Kraak, M.J., 2005. Exploring geovisualization. In: *Exploring geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 3-19.
- Eades, P., 1984. A Heuristic for Graph Drawing. *Congressus Numerantium* 42. pp. 149-60. 1984.
- Edsall, R.M (1999). The dynamic parallel coordinate plot: visualizing multivariate geographic data. *Proc. 19th International Cartographic Association Conference, Ottawa, May 1999.*
- Edsall, R.M., 2003. The parallel coordinate plot in action: design and use for geographic visualization. *Computational Statistics & Data Analysis* 43(4): 605 – 619.
- Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life.* Times Books, New York.
- Ekman, P. and Friesen, W. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions.* Prentice Hall, Upper Saddle River, New Jersey.
- Fayyad, U. M., Grinstein, G.G., and Wierse, A., 2002., *Information Visualization in Data Mining and Knowledge Discovery.* Morgan Kaufman, San Francisco etc., 407 pp.
- Francis, B and Fuller, M (1996) Visualization of event histories. *Journal of the Royal Statistical Society Series A.* 159, 2, 301-8.
- Francis, B. and Pritchard, J., 1998. Bertin, Lexis and the graphical representation of event histories. *Bulletin du Comite Francais de Cartographie* (156). pp. 80-88.
- Gahegan, M., 2005. Beyond Tools: Visual Support for the Entire Process of GIScience. In: *Exploring geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 83-99.
- Gahegan, M.M., Harrower, M., Rhyne, T.M., 2001. The integration of geographic visualization with knowledge discovery in databases and geocomputation. *Cartography and Geographic Information Science*, Vol. 28, No. 1, Jan., 2001, pp. 29-44(16).
- Gatalsky, P., Andrienko, N. and Andrienko, G., 2004. Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on Volume , Issue , 14-16 July 2004 Page(s): 145 – 152.
- Hägerstrand, T., 1967. *Innovation Diffusion as a Spatial Process.* Chicago, Illinois: The University of Chicago Press.
- Hagerstrand, T. (1970). What about people in regional science? *Papers of the Regional Science Association* 24: 7-21.

- Haris, R.L., 1996. Information Graphics – A comprehensive Illustrated Reference. Management Graphics, Atlanta, Georgia, USA.
- Keim, D.A., 2001. Visual exploration of large data sets. *Commun. ACM*, 44(8): 38 – 44.
- Keim, D.A., Panse, C., Sips, M., 2005. Information Visualization: Scope, Techniques and Opportunities for Geovisualization. In: *Exploring geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 23-52.
- Koua, E.L, and Kraak, M.J., 2005. Evaluating Self-organization Maps for Geovisualization. In: *Exploring geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 627-643.
- Koua, E.L., McEachren, A.M. and Kraak, M.J., 2006. Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International journal of geographical information science = IJGIS*, 20(4).
- Koussoulakou, A. and Kraak, M.J., 1992. The spatio-temporal map and cartographic communication. *Cartographic Journal* 29 (2): 101 – 108.
- Kraak, M.J., 2003. Geovisualization and time - new opportunities for the space-time cube.
- Kraak, M.J., 2003. Geovisualization illustrated. *ISPRS Journal of Photogrammetry and Remote Sensing: official publication of the international society of photogrammetry and remote sensing - ISPRS*, 57(5-6).
- Kraak, M.J., 2005. Timelines, Temporal Resolution, Temporal Zoom and Time Geography. In: *ICC 2005 : Proceedings of the 22nd international cartographic conference : mapping approaches into a changing world, 9-16 July 2005, A Coruña, Spain*. International Cartographic Association (ICA), 2005. ISBN: 0-958-46093-0. 7 p.
- Kraak, M.J. and MadzudzoIn, P.F., 2007. Space time visualization for epidemiological research. *Proceedings of the 23rd international cartographic conference ICC: Cartography for everyone and for you, 4-10 August 2007 Moscow, Russia*. Moscow: International Cartographic Association (ICA), 2007. ISBN: 0-958-46093-0. 13 p.
- Kraak, M.J. and Koussoulakou, A., 2004. Visualization Environment for the Space - Time Cube. In: *SDH 2004: Proceedings of the 11th international symposium on spatial data handling: advances in spatial data handling II. : 23-25 August 2004, University of Leicester*. / ed. by P.F. Fisher. - Berlin etc.: Springer, 2004. pp. 189-200.
- Kraak, M.J. and MacEachren, A.M., 1999. Visualization for exploration of spatial data. In: *International Journal of Geographical Information Science*, 13(1999)4, pp. 285-287.

- Kraak, M.J. and Ormeling, F.J., 2003. *Cartography: Visualization of Geospatial Data*. Addison Wesley Longman, Harlow, 205 pp.
- Kristensson, P.O., Anundi, D., Björnstad, M., Dahlbäck, N., Gillberg, H., Haraldsson, J., Mårtensson, I., Nordvall, M. and Ståhl, J. 2006. The tradeoffs when using space time cube representation. In *Conference Compendium IEEE Symposium on Information Visualization (InfoVis '06)*. IEEE Press: 130-131.
- Krueger, R. A. & Casey, M. A. (2000). *Focus groups: A practical guide for applied research*. 3rd ed. Thousand Oaks, California: Sage Publications, Inc.
- Langran, G., 2002. *Time in geographic information systems*, London: Taylor & Francis.
- Lexis, W. (1875) *Einleitung in der Theorie der Bevölkerungsstatistik*. Trübner: Strassburg.
- Lloyd, R., 2000. Self organizing cognitive maps. *The Professional Geographer*, 52(3), 517-531. 2000.
- Lock, G., and Stancic, Z., 1995. *Archaeology and Geographic Information Systems: A European Perspective*. Taylor and Francis, Londond etc., 392 pp.
- Michels, W.J., 1972. Dating Methods. *Annual Review of Anthropology*, Vol. 1. (1972), pp. 113-126.
- Monomier, M. 1990. Strategies for the visualization of geographic time-series data. *Cartographica*, 27(1): 30 – 45.
- Moore, A.B., Whigham, P.A., Aldridge, C., Holt, A. and Hodge, K., 2003. Sport and Time Geography: A good match?. In: P A Whigham (ed), *Proceedings of the 15th Annual Colloquium of the Spatial Information Research Centre*, Dunedin, New Zealand: University of Otago.
- Moylan, B., 2001. *Mapping Landscape Change: Space Time Dynamics and Historical Periods*. University of Sydney, School of Geosciences and Archaeological Computing Laboratory.
- Müller, W. and Schumann, H., 2003. Visualization Methods for Time-dependent Data – An Overview. *Proc. Winter Simulation Conference*, New Orleans, 2003.
- Munzner, T., and Burchard, P., 1995. Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space. *VRML '95*, special issue of *Computer Graphics*, ACM SIGGRAPH, pp. 33-38. 1995.
- Peuquet, D.J., 1994. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*. 84(3):441-461.
- Peuquet, D.J. and Duan, N. 1994. An Event-Base Spatio-temporal Data Model (ESTDM) for Temporal Analysis of Geographic Data. *Forthcoming in International Journal of Geographical Information*.
- Pressat, R. (1961) *L'Analyse De'mographique*. Paris: Presses Universitaires de France.

- Rainio, P.A, and Kraak, M.J, 2005. Towards Multi-variate Visualization of Metadata Describing Geographic Information. In: *Exploring Geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 611-625.
- Robertson, G.G., Mackinlay, J.D., and Card, S.K., 1991. Cone Trees: Animated 3D Visualizations of Hierarchical Information, *Proc. CHI*, pp. 189-193, 1991.
- Roberts, J.C., 2005. Exploratory Visualization with Multiple Linked Views. In: *Exploring Geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 159-180.
- Rogers, P., 2005. Graph Drawing Techniques for Geographic Visualization. In: *Exploring Geovisualization* / ed. by J. Dykes, A.M. MacEachren and M.J. Kraak. Amsterdam: Elsevier, 2005. ISBN 0-08-044531-4. pp. 143-158.
- Rouse, I., 1993., *The Tainos: Rise & Decline of the People Who Greeted Columbus Cumberland, Rhode Island, U.S.A.* Yale Univ Pr.
- Sugiyama, K., Tagawa, S. and Toda, M., 1981. Methods for Visual Understanding of Hierarchical Systems Structures. *IEEE Trans. on Systems, Man and Cybernetics*, 11(2), pp. 109-125, February 1981.
- Tominski, C., Abello, J. and Schumann, H., 2003. Interactive Poster: Axes-based Visualizations for Time Series Data. *Proc. IEEE Symp. on Information Visualization*, Seattle, 2003.
- Tominski, C., Abello, J. and Schumann, H., 2004. Axes-Based Visualizations with Radial Layouts. *Proc. ACM Symp. On Applied Computing*, Nicosia, 2004.
- Tutte, W.T., 1963. How to draw a graph. In *Proceedings of the London Mathematical Society* 13, 743–768.
- Ulrik, B. and Dorothea, W., 1998. Using Graph Layout to Visualize Train Interconnection Data, *Proceedings of the 6th International Symposium on Graph Drawing*. Springer-Verlag.
- Wheatley, D., and Gillings, M., 2002. *Spatial technology and archaeology : the archaeological applications of GIS*. Taylor and Francis, Londond etc., 269 pp.
- Ware, C., 2004. *Information Visualization: Perception for Design*, 2nd ed. San Francisco, CA: Morgan Kaufmann.
- Yee, K.P., Fisher, D., Dhamija, R., Hearst, M., 2001. Animated Exploration of Dynamic Graphs with Radial Layout. *Proc. IEEE InfoVis*. pp. 43-50. 2001.
- Xia, L., 2005. “New Methods of Visualization of Multivariable Spatio - Temporal Data: PCP - Time - Cube and Multivariable - Time – Cube”, *ITC*, Enschede, 93 pp.

