Visual Analytics on Spatial Time Series for Environmental Data

Mithileysh Sathiyanarayanan, Vijayakumar Varadarajan, Pradeep .K .V

ABSTRACT--- Environemental science is a field of growing and vital importance utlizing cutting-edge research methods and techniques. In this paper we present a review of the visualisation techniques for spatial time series used in environmental research. Specifically, we focus on the following sub-domains in the field: resource depletion monitoring, climate change, environmental pollution and weather forecasting. These domains present their own idiosyncrasies in relation to data integrity and availability, answers to analytical questions sought, and therefore analytical approaches to visualisation when dealing with datasets referenced both spatially and temporally. We show that when data quality is poor (as in deforestation studies) most of the analytical processes target the data pre-processing stage of the visual analytics framework while the choice of visualization format aims for simplicity and visual impact via choropleth maps. When data quantity and quality are less of an issue (climate change, pollution) more sophisticated techniques are applied, by slicing the multi-dimensional dataset into more manageable parts. Typically, with time and space as references and multiple attributes considered, the visual analytical approach broadly splits two ways: either (1) considering the spatial variation of time series (via diagram maps and interactively linked displays) or (2) considering the temporal variation of spatial distributions (via the use of small multiples" and map animation). These visualisation techniques are an unvaluable tool to study the domain of interest. They help analyse attributes covariance, detect trends and anomalies, or more generally describe spatial, temporal and thematic relationships in the attributes-space. Finally, when data complexity becomes itself a challenge as in weather forecasting, in addition to traditional methods, more novel and experimental visualisation techniques are deployed to try and capture both spatial and temporal behaviour of given attributes in a single framework (3). 3D visualisations, space-time cubes, and more sophisticated approaches are here called upon to analyse model uncertainty and perform complex model comparison. Visual analytics therefore offers an extensive an evolving framework to support environmental research studies from beginning to end.

Index Terms — Spatial time series, visualisation methods, visual analytics, weather, pollution, climate change, deforestation,

INTRODUCTION

Environmental research is a constantly growing field reflecting the vital importance of understanding and managing the only ecosystem known to us to harbor life in the universe: planet Earth.

From descriptions of extreme pre-historic atmospheric events to modern automatic data collection, satellite imagery and state-of-the-art descriptive or forecasting models, the need to understand the functioning of our biophysical environment has evolved into an imperative to (i) monitor changes, man-made or not, (ii) provide accurate short term and long term forecasts, (iii) devise appropriate policies and solutions to manage dwindling resources and finally (iv) mitigate potential long term changes that could threaten our very existence.

The complexity of the task at hand first presents itself in the data. Environmental data is routinely big data, multivariate and typically time-varying over different spatial distributions. Researchers need to interact with this data typically over four distinct modes of inquiry: Exploration, Explanation, Prediction and Planning (Peuquet, 1994).

A visual analytics framework aims to support these different modes of inquiry by helping answer specific analytical questions as summarized by Blok (2000):

- How to characterize ongoing processes
- How to highlight anomalies, deviations from 'normal'
- How to identify trends
- How to describe spatial, temporal or thematic relationships
- How to highlight causes

In this paper, we look at how the visual analytics framework is used within this context when the environmental data is represented as spatial time series "i.e. values of numeric attributes referring to different moments in time and locations in space" (Andrienko and Andrienko, 2005).

These visual analytics approaches/views are split 2-ways depending on how the dataset is considered (Andrienko et al., 2015):

- (a) as a sequence of spatial distributions of attribute values in different time units
- (b) as a collection of local time series of attribute values in different locations

Those views are mostly one or two dimensional but increasingly a third category using 3D visualization (c) is being used, particularly for very high dimensionality data sets.

We focus on 4 specific subfields of environment research: resource depletion monitoring (deforestation in this case), climate change (long term weather changes), environmental pollution (here air quality monitoring) and weather (short and medium-term forecasts).

We order these domains in terms of increasing data complexity and highlight for each, the principal methodological approaches for visualization used to answer specific research questions or attain specific outcomes. Here, we define "data density" or "data complexity" by both its relative quality and absolute quantity.

Revised Manuscript Received on May 15, 2019.

Mithileysh Sathiyanarayanan, City University of London, U.K (s.mithileysh@gmail.com)
Vijayakumar Varadarajan, VIT University, Chennai, India. (vijayakumar.v@vit.ac.in)
Pradeep .K .V, VIT University, Chennai, India. (Pradeep.kv@vit.ac.in)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

1. LOW DATA DENSITY: DEFORESTATION

In this section of the report, we look at the how visual techniques have been applied to deforestation studies, highlighting the particular issues and challenges. unique to this domain.

Deforestation can occur for many reasons, for instance conversion of the land for farming or industrial use of the trees (timber of fuel) or due to natural changes in the environment.

The process is concerning from an ecological point of view for a number of reasons, including damage to habitat, biodiversity loss and aridity. However, perhaps the main concern is its adverse impact on bio sequestration of atmospheric carbon dioxide.

1.1 Key themes, questions and challenges

In reviewing existing literature on mapping deforestation, the key themes that emerge are data availability, quality and interpretability.

Historically, it has been challenging to obtain accurate data in order to map areas where large scale deforestation was occurring. For instance, the deforestation of the Amazon rainforest, which supplies around 20% of the world's oxygen is of major concern to researchers but is located in a very remote area of the world and the size of the area to monitor is vast.

Until recently, data-gathering involved a variety of "on-site" collection methods overly sensitive to sampling bias since only a limited proportion of the areas of interest could be accessed. A few monitoring programs have been set up to help this process, for instance PRODES in the Amazon rainforest (Souza et. al, 2013).

However, since the creation of the Brazilian Institute for Space Research (INPE) in the 1970s, remote sensing using satellite imagery has become central to how deforestation in the Brazilian Amazon and other areas are monitored and measured (Monteiro and Rajão 2017)

Using satellite imaging has a number of advantages, including reduced cost, increased coverage and accessibility to researchers as well as transparent and reliable data collection methods (Qamer et al. 2016). As these methods have become more common, imaging quality has improved and the data has been made available in the public domain (for example Google Earth Images), their use has become widely adopted and now represent a real alternative to traditional approaches.

Yet the use of satellite imagery has brought with it its own issues: satellite images are very hard to read and interpret. They need to be pre-processed to adjust for cloud cover, variable illumination and geometric distortions (Qamer et al. 2016). This pre-processing can be very resource-intensive and time-consuming. Figure 1 shows the steps needed to map deforestation in the Amazon rainforest using satellite imagery.

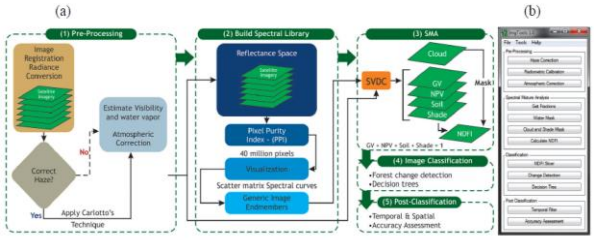

Figure 1: Pre-processing steps for Satellite imagery of rainforests (Souza et. al, 2013).

Once the data has been pre-processed, it needs to be interpreted and classified in order to be used for further analysis and/or visualization. This task has often been carried out manually, thus limiting the availability and accuracy of the data and here also potentially biasing it. Recently some researchers have applied machine learning methods to help interpret the satellite images. Souza et al (2013) use decision trees to classify the cells of images into different deforestation areas. They report that using this approach led to the same level of accuracy in determining areas of deforestation as the traditional PRODES approach.

1.2 Visual analytics approaches

Issues with data collection, the low cardinality of the data itself, and the main objectives of resource monitoring studies in general have oriented most of the visualization techniques towards variations of approach (a) and (b) to illustrate and communicate changes of attribute values over large areas via the use of choropleth maps.

1.2.1 Use of single choropleth maps to represent changes

The majority of the papers on deforestation reviewed chose to illustrate the spatial time series data as a single map rather than using time displays.

This makes sense considering the primary focus of most of the research in the domain is to analyze how the level of forest cover has changed over time in a given area. Figures 2 and 3 show typical examples of the most common visualization approaches to the issue of conveying the main message: i.e. the change in forest cover between two time periods. This approach is meaningful because over time some areas show positive or negative changes as to the level of deforestation or use of land.

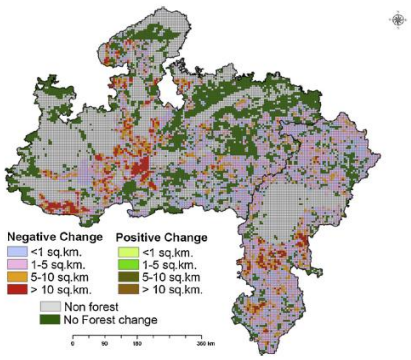

Figure 2: Forest cover change map of Central India between 1935 and 2010 (Sudhakar Reddy, C. *et al.*, 2015)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

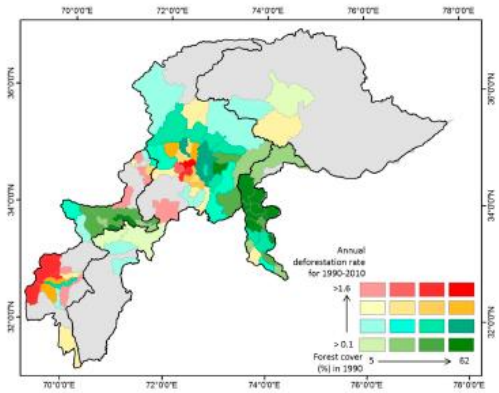

Figure 3: Forest cover change map of Central India between 1935 and 2010 (Sudhakar Reddy, C. *et al.*, 2015)

1.2.2 Use of sequential maps (mini-maps)

The other common approach is the use of multiple choropleth maps or even actual images to illustrate changes of multiple attributes over time. See Figure 4 below showing details of terrain types and their temporal evolution.

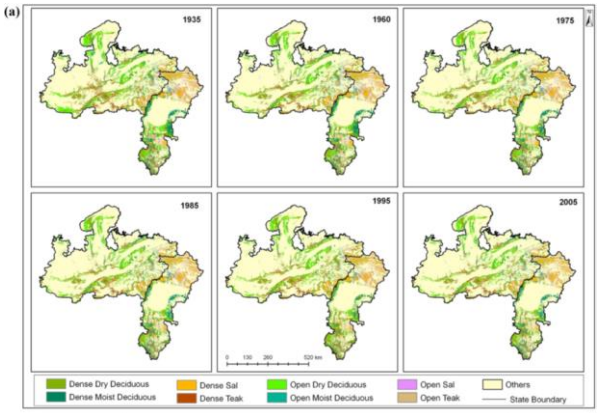

Figure 4: Forest cover maps of Central India between 1935 and 2005 (Sudhakar Reddy, C. *et al.*, 2015)

These approaches are very common in all the other domains as well and we show other examples later in this paper.

1.3 Conclusions for low data density

In this section, we have highlighted a common situation where when data quality is low, pre-processing is usually the most analytically intensive stage while visualizations of the final data are relatively simple and aim at maximizing visual impact by highlighting changes. Nevertheless, visual analytical approaches can also form an important part of the pre-processing stage as shown in figure 2.

2. MEDIUM DATA DENSITY: CLIMATE CHANGE

In this section, we focus on climate change and how visualization techniques are used in the research process when the data is represented as spatial time series

Climate change is one of the biggest challenges faced by humanity in 21st century. In the fourth IPCC report, Barros *et al.* (2014) define climate change as the "change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties, and that persists for an extended period, typically decades or longer" and conclude that global mean temperature has increased greatly due to the emission of greenhouse gases largely due to human activity.

While related to weather and climate research, which have very high-density data sets (see section 4), climate change research tends to focus on long term macro averages and therefore deals with medium density data sets.

2.1 Key themes, questions and challenges

There are two main challenges in climate change research (or climate impact research) that are supported by a visual analytics approach.

First is anomaly detection. Researchers in the field are constantly looking for indications of deviation to the normal, historical, or forecasted climate parameters. Indeed, despite mounting evidence, there is hardly a consensus as to what is really happening to the Earth's climate. There is even greater uncertainty as to what should be done about it, as evidenced by the US pulling out of the Paris climate agreement earlier this year.

Second is conveying the results of research to decision makers and the wider public. The issue is of such great importance to everyone that findings need to be presented in a manner that is easily accessible to non-academic readers.

2.2 Visual analytics approaches

2.2.1 Multiple linked choropleth map displays

Moving away from the simplicity of the single choropleth maps while keeping their full visual impact, Nocke et al (2008a) present an interactive visual tool allowing the non-academic user to compare different climate scenarios and their related impacts on many biosphere variables and from different perspectives.

This allows to convey the complexity of the underlying data and the inherent uncertainty of any analysis over long time scales and spatial variation, while delivery a strong identifiable message (Figure 5)

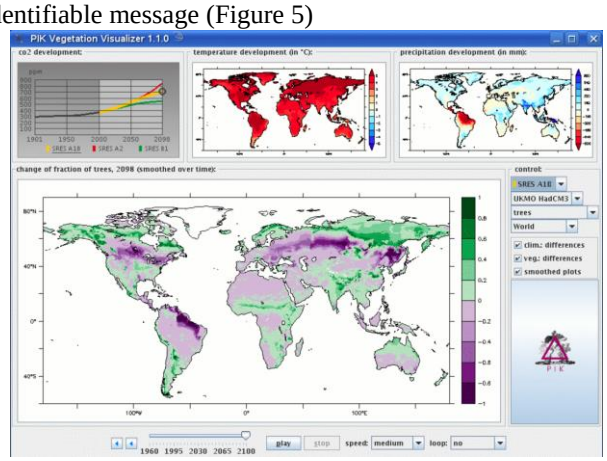

Figure 5: PIK Vegetation Visualizer, allowing the user to choose different CO2 emission scenarios, climate models, vegetation variables and different time steps. (Nocke et al. 2008a)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

2.2.2 Interactive Visual Analysis

In climate change research, most visualization techniques are typically applied after the data has been post-processed using statistical methods (Nocke et al, 2008b), but recent approaches have sought to use visual analytics as a research tool in its own right.

Li et al., 2015 present Vismate, an interactive visual analytics tool for spatial time series. The systems integrate three interrelated visualization techniques (Figure 6):

- A global radial map surrounded by cluster rings to help identify the overall state of climate changes
- Time-series discs detailing cluster features
- An interactively linked PCA scatter plot to identify outliers in the data.

The tool is applied to analyze surface air temperature in mainland China between 1975 and 1989.

The authors show how the complementary views can be used to analyze trends and detect anomalies. They also highlight the potential to generalize the same visualization framework to other datasets of spatial time series

2.3 Conclusions for medium data density

As data density increases, so does the opportunity for visualization techniques to move beyond the purely descriptive stage and onto a more explanatory role.

Choropleth maps are complemented by more sophisticated spatial time series visualization methods that helps identify and visualize trends and discover anomalies.

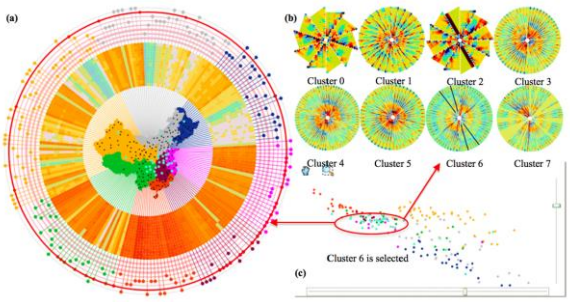

Figure 6: Vismate dashboard showing (a) radial map with cluster rings, (b) Time series discs for each cluster and (c) interactively linked PCA scatter plot (Li et al., 2015)

3. HIGH DATA DENSITY: MAN-MADE POLLUTION

Moving towards greater data complexity, we now look at time series data from studies on air pollution.

Deterioration of air quality in major cities around the world is becoming an increasing concern. Making predictions based on modelled data can be helpful in many ways including reducing potential human exposure, reducing respiratory health effects and also in informing sound policy decision.

Air quality monitoring is done through the deployment of multi-modality and large-scale sensor networks. As an illustration, there are around 300 monitoring sites in the UK, each recording concentration of over 10 attributes relating to air quality such as PM10, PM2.5, CO, NOx. Mean concentration and real-time data for these attributes are recorded and available for download (DEFRA, 2017). Data density is therefore high and can be streamed in real-time.

Pollution levels measured at each station can be influenced by a range of factors including: proximity to pollution sources (both static and moving), dispersion of pollution by wind, removal processes such as rain and deposition, temperature, sunlight and weather conditions.

Therefore, spatiotemporal modelling of air quality data is a multivariate, multidimensional problem that analytical and numerical techniques alone are unable to solve. Visualization techniques are often combined with numerical techniques to get an insight into air quality data and unveil hidden patterns and trends. Tools used to model and predict air quality data are required to be capable of handling a range of scalar and vector attributes such as the concentration of sources, temperature, weather conditions, sunlight and wind.

3.1 Key themes, questions and challenges

There have been numerous studies on establishing the most effective way to analyze and visualize spatially and temporally distributed air quality data (Qu *et al.*, 2007; Fecht *et al.*, 2016; Du *et al.*, 2017; Zhou *et al.*, 2017).

These studies tend to focus on establishing the correlation between API (Air Pollution Index) and various attributes concerning air quality and then utilizing this information to detect spatial and temporal patterns of API.

Key questions the visual analytic research on air pollution attempt to answer are (Qu *et al.*, 2007):

- How to evaluate correlation between API and related attributes
- How visual analytics can be best utilized to establish similarities and differences of pollution levels both in terms of space and time
- How visualization can be used to pinpoint pollution sources
- How time series information on pollution data can be used to detect trends.

One of the major challenges in analyzing and visualizing pollution data is the multidimensional nature of its attributes. According to (Qu *et al.*, 2007), weather and pollution data often consist of more than 10 dimensions. Since these attributes are a mixture of both scalar and vector quantities, traditional methods used in processing multivariate scalar quantities only cannot effectively deal with pollution data. Therefore, novel visualization techniques that can encode both vector and scalar quantities are needed.

3.2 Visual analytics approaches

3.2.1 Spatial variation of the temporal behaviour

Maps produced for spatiotemporal pollution data usually need to be scalable so that the data can be visualized at different levels of granularity. In Du et al. (2016), the authors present an interactive visual analysis system (AirVis) that combines multi-dimensional visualization, spatiotemporal analysis and multi-scale methods.

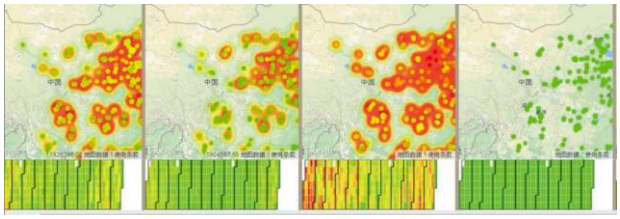

Figure 8: Multidimensional view in AirVis (Du et al., 2016)

The system consists of a collection of different views: map, calendar and trend. An interactive, scalable heat map is used in this system to visualize overall air quality data of an area. Two trend lines are used to display trends in different regions and the calendar view is for detecting long term or seasonal changes. The system also provides tools to combine the three views to gain a multidimensional oversight. Multi-scale time series visualization map-based view presented here allows the user to view air quality data on a map at various levels of granularity (Figure 8).

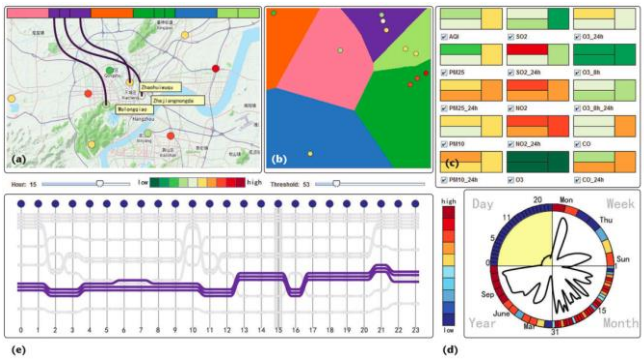

Figure 9: Visual Analytics system proposed by Zhou et. al. (2017). (a) Geographical of the map of Hangzhou city, China with Beziers highlighting the areas selected for visualization. (b) A Voronoi diagram (c) A tree map showing concentration of pollutants for the selected cluster (d) A navigation map that captures periodic changes and (e) Temporal variation of pollutant levels on a story line.

The work proposed by Zhou et al. (2017) focuses on visual analytics of spatial clusters. Spatial clusters from air quality data are initially established by analyzing the similarities and differences of various monitoring parameters. Visual analytic techniques are then applied for selected individual clusters to detect patterns. Multiple views (see Figure 9) are used to combine and assess the correlation between various attributes at each spatial location. An interactive navigation chart with a range of different scales allows the user to select and assess the variation of concentration of pollutants over space and time.

3.2.2 Temporal variation of the spatial behaviour

Due to multidirectional nature of air quality data, temporal analysis of air quality data also presents both linear and cyclic variations.

Although there are a number of different techniques to visualize time series, most of these techniques require user input to manually select time intervals to detect trends. Du et al. (2017) present an approach where the system selects the granularity of the presentation automatically based on the sensed data. Time series information is presented in two trend views, one on short term variations and the other on long term trends, and in both the views, the system selects the level of granularity based on observed nature of the temporal variation. Interactive multidimensional views are available in this system, allowing comparison between concentration levels of various attributes both spatial and temporal dimensions. However, the system can take only a limited number of parameters and the analysis becomes complicated when more attributes are considered. The system needs to be expanded further to simulate real world situations where factors such as wind and temperature need to be considered.

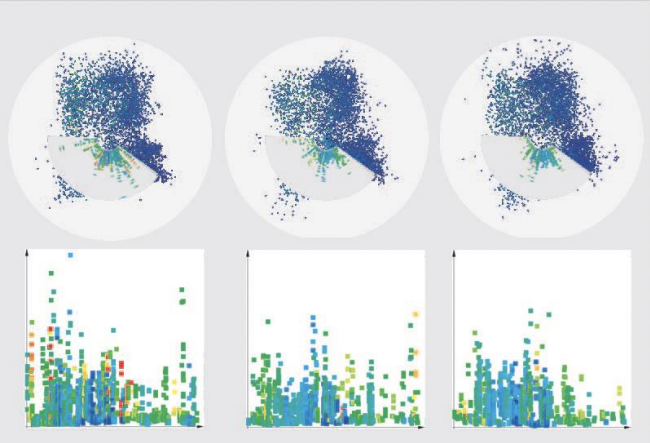

Figure 10a: Time Series polar plots for Kwai Chung station focusing on the impact of local pollution from the southwest direction. (Qu et al. 2007)

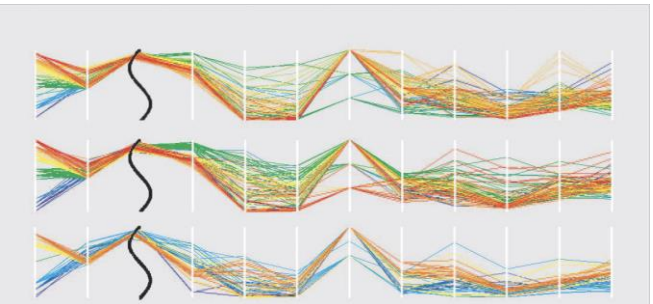

Figure 10b: Time series data for Yuen Long district on a parallel coordinate with time axis. (Qu et al. 2007)

In Qu et al. (2007), time series data and pollution data in Hong Kong has been investigated using polar plots (see Figure 10a and 10b). Here, a method that combines polar and parallel coordinates has been used to sample and select data for further analysis through parallel coordinates. By combining the two systems, unnecessary data is filtered leaving effective time series data to be processed through the parallel coordinates. Then the parallel coordinates are used to assess correlation among various attributes and trends. A special s- shaped map is used to encode the vector quantity (wind direction) information on a 2D scalar map.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

3.2.3 3D visualisations

Existing 2D visualization methods are still insufficient to cope with complex data sets such as pollution data due to the multi-dimensional nature of its attributes. There are attempts to develop 3D visualization systems to address these issues such as ViNeu (Kreuseler, 2000). ViNeu supports both 2D maps and 'real' 3D landscape representations. This system has been developed for investigating marine ecosystems and it provides visualization tools for rendering multidimensional data within virtual 3D scenes. Although a similar approach could be implemented to visualize pollution data in 3D, the system needs to be further developed to incorporate challenging issues posed by pollution data. Furthermore, air quality measurement systems deployed today are still too sparse to predict three-dimensional pollution information at street level. There are, however, some effort to extend existing 2D maps to 3D for selected attributes to enhance visualization.

3.3 Conclusions for high data density

High density datasets present unique challenges that can defeat purely computational approaches without a tool to seed the calculations. Visualization is increasingly applied as such a tool when the data dimensionality has reached a certain threshold. Visual analytics then becomes an important part of the modelling process.

In addition to trend and anomaly detection, a correctly posed visual analytical framework can help dive more into the analysis to highlight correlations, compare situations in time and space and help establish causal links between attributes.

4. VERY HIGH DATA DENSITY: WEATHER FORECASTING

As one of the most powerful expression of natural forces, weather is important to all our lives. Extreme weather can impact the livelihoods of large populations and weather damage can cost billions of dollars.

Whilst the accuracy of weather forecasting has improved significantly, with the availability of more data measurements and more powerful analysis techniques, the desire to improve forecasting has only increased.

Weather-related datasets have very high dimensionality, are fed real-time from a huge number of satellite and ground-based stations into computing super-clusters to run some of the most complex calculations done in modern scientific research.

4.1 Key themes, questions and challenges

Machine learning and predictive weather models are central to meteorology today but forecasters need to be able to both understand the reasons underlying the machine-learned forecasts and communicate those forecasts to a wide audience (Anadiotis, 2017). The design and use of the right visualization framework is critical to address both of these demands.

Meaningful visualization of weather data is complex for a number of reasons.

Firstly, weather is inherently an ever-evolving 3-dimensional phenomenon and therefore weather data typically exists over 3D space and time.

Secondly, the data is likely to incorporate many different measurements such as temperature, pressure, rainfall etc. Accurate forecasting often needs very granular information, requiring an extremely high-density dataset.

Finally 'forecasters need tools to detect and analyze trends and anomalies, especially model errors, to assess forecast uncertainty and to produce more accurate predictions' (Diehl et al., 2015). Visualization of that uncertainty is a key challenge for meteorology.

4.2 Visual analytics approaches

4.2.1 Spatial variation of the temporal behaviour

There is limited scope to extend the traditional use of small multiples without simply squeezing more information on the screen. However, intelligent use of very small multiples can still support a more detailed analysis as part of an interactive dashboard approach, particularly with the use of visually simplified map icons as illustrated in the 'mini-map' example in Figure 11.

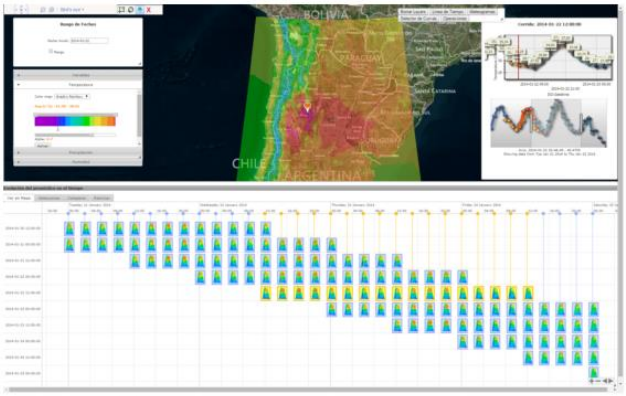

Figure 11: 'Minimaps' showing the use of granular small multiples,with each line representing a different forecast initialization time (Diehl et al., 2015).

Computational operations can be applied to a range of forecasts selected across horizontal and/or vertical axes to highlight different characteristics within forecast data (Diehl et al., 2015). For example, between different models, visualizing a subtraction operation can highlight differences between forecasts over time and visualizing standard deviations helps the user to estimate forecast uncertainty.

4.2.2. Temporal variation of the spatial behaviour

Several proposed enhancements can be made to the standard diagram map visualization and two different but potentially complimentary approaches are possible.

One approach is to minimize the size of diagram maps in order to present time data across a more granular map grid and to enhance interpretability with the use of stylized icons ('glyphs') that highlight a particular feature of the data.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Another approach is to use non-standard graphs to highlight different aspects of weather data, particularly of a seasonal or cyclical nature using techniques such as time spirals.

Wickham et al. (2012) have proposed a detailed approach for visualizing climate data that combines both of these approaches using a variety of 'Glyph-maps' that are designed to highlight different aspects of complex multivariate data and 'allow the discovery of both local and global structure' as illustrated in Figure 12. The proposed Glyph-maps have been designed to be both computationally efficient and to provide visual clarity, recognizing that limits to the resolution of human vision 'prevent perfect perception of large spatial domains or seasonality of many years of data'.

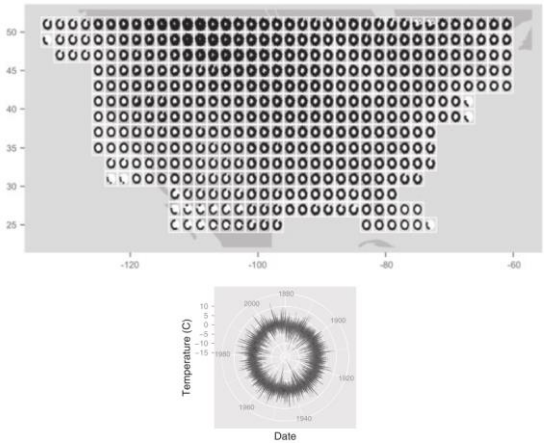

Figure 12: The use of stylized graphical icons ("glyphs") proposed as a way of visualizing more detailed time series data over a map area that still provides global structure to the data. (Wickham et al. 2012)

Whilst multi-variate weather data can be highly complex, changes over space are typically smooth such that neighboring locations exhibit similar data. Glyph-maps take advantage of this, as they give patterns and trends that are exhibited over the map "the appearance of a textured landscape" (Wickham et al. 2012) and therefore can help to make these trends and any local anomalies easy to detect by eye.

4.2.3. Space time cubes and 3D visualisations

Whilst the approaches discussed so far focus on ways to make visual analysis possible once more data is "crammed in", they still do not readily address the specific challenges of 3-dimensional physical space and forecast uncertainty where a large number of forecast models are available. We find that these challenges have been best addressed with 3-dimensional visualizations and space-time cubes combined with sophisticated computational approaches such as clustering.

Several authors adopt the approach of identifying patterns of similar behavior in 2D or 3D space as an 'iso-contour' and then visualizing its progression over time as illustrated in Figure 13.

The visualization of iso-contours is perhaps the simplest example of clustering similar data points together. Clustering is the most widely used technique for visualizing highly complex data. This approach lends itself to 3D visualizations where clustering can be done in time, space or both.

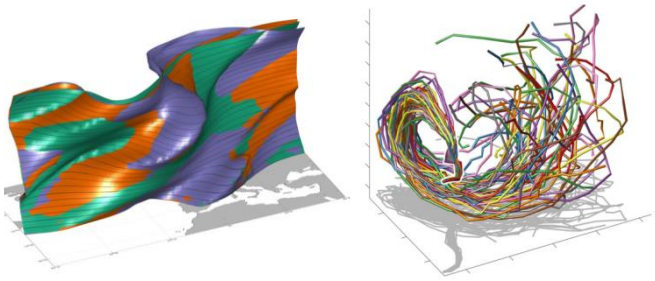

Figure 13: Iso-contours and their development over time: (left) in 2D space, with time on the z-axis (colors representing clusters) and (right) in 3D space with time-based coloring (Ferstl et al. 2017).

Clustering is particularly powerful when trying to differentiate between weather forecast models. For example (Ferstl *et al.*, 2017) a 'time-hierarchical clustering' approach which first groups forecasts together using contour clustering and then merges diverging clusters together in time-reversed order before projecting these color-coded clusters onto space-time cubes as in Figure 14.

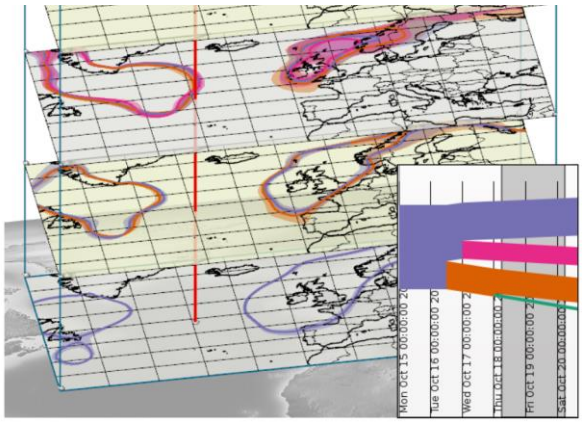

Figure 14: An example of forecast clustering - stacked time-cuts with variability plots and the corresponding time-hierarchical cluster tree (bottom right) (Ferstl et al. 2017).

The obvious disadvantage of any of the 3D visualizations highlighted above is that occlusions are often introduced, with some data obscuring other potentially meaningful information. However, 3D rendering is a more intuitive way of visualizing complex data and graphical techniques such as transparency and shading are widely used to mitigate the occlusion problem.

4.3 Conclusions for very high data density

The inherent complexity of weather and weather forecast data means that there are no easy solutions to presenting the analysis to the user in the most visually comprehensive, insightful and intuitive way.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Unsurprisingly, a combination of different visualizations continues to be the favored approach and much software innovation is focused on creating dashboards that provide an intuitive interface to allow the user to combine and switch between different views. This approach has the benefit of bringing together tools such as 'data analytics, ensemble analysis, uncertainty quantification, metrics computation, and visualization components' (Diehl et al., 2015) into one place as illustrated in Figure 11, allowing direct model comparison and visualization of model uncertainty profiles.

5. DISCUSSION AND CONCLUSION

Spatial time series data exhibit an inherent conceptual complexity which calls for visual representation support to deal with even the lowest density dataset. Therefore, visual analysis is very often a critical part of the analytical process applied to data referenced both in space and time.

Environmental research has a very wide and far reaching subject area. Its various subdomains each present unique features, challenges, approaches and desired outcomes. In this paper, we have highlighted how the visual analytical framework, in relation to spatial time series datasets, interacts with specific subdomains of environmental research ordered in terms of data complexity or data density.

Our findings are summarized in Table 1 below. We show that, understandably, the complexity of visual analysis tools and techniques selected increases as the data density increases. Although simpler tools can be (and often are) used at any higher stage, the full benefits of a correctly posed visual analytical framework can be reaped when complexity levels are matched.

Ultimately, instead of simply being a tool or support to the research process, the most successful approaches integrate a visual analytics framework into their core.

Data Density Level	Optimal VA Techniques	Main Target Outcome
Low	Choropleth Maps	Visual Impact
Medium	Dynamic-linked Displays	Trend and Anomaly Detection
High	Clustering and Projections	Correlation and Causality Analysis
Very High	Space-Time Cubes and 3D Displays	Model Comparison and Uncertainty Analysis

Table 1: Visual analytics according to data density and target outcomes

We believe that these findings can be easily generalized to other subject domains, as they are data-source agnostic and instead highlight an interplay between data typology and generic visual analytics techniques.

There does however appear to be room for further innovation at the cutting-edge of environmental data visual analytics. Virtual reality has been considered by some authors (Helbig et al., 2014) (see Figure 15) but the benefits are inconclusive and further research is needed in order to determine how this technology can best be exploited in practice.

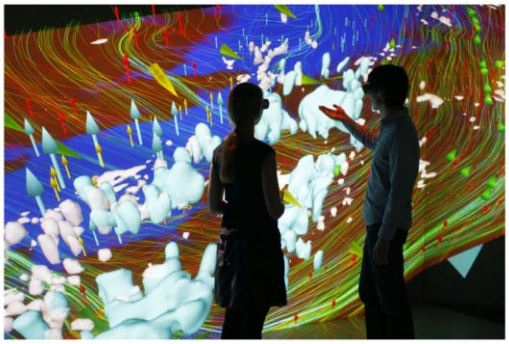

Figure 15: Concept visualization utilizing high definition graphical displays and virtual reality googles (Helbig et al., 2014)

Furthermore, techniques to viewing otherwise-occluded areas of a visualization (for example, see Figure 16) have been considered in other fields and we believe that these could have significant benefits in the weather domain as well, justifying further research.

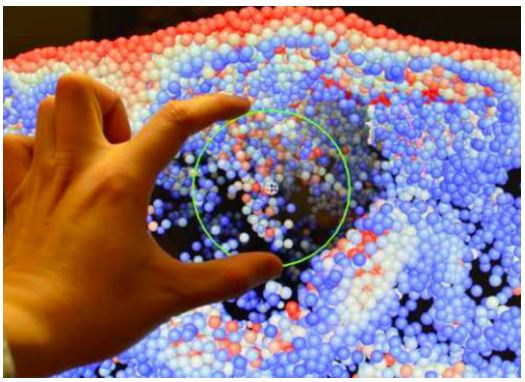

Figure 16: Illustration of a technique for managing occlusion with a user-controlled 'GlyphLens' (Tong, Li and Shen, 2017).

ACKNOWLEDGMENTS

The authors wish to thank City University and the lecturers on the Visual Analytics module of its Data Science MSc class of 2017/18 for providing much of the inspiration for this paper.

REFERENCES

1. 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*. Taylor & Francis, Ltd.Association of American Geographers.
2. Blok, C. (2000). Monitoring Change: Characteristics of Dynamic Geo-Spatial Phenomena for Visual Exploration 1, 16–30.
3. Andrienko, G. and Andrienko, N. (2005) 'Visual Exploration of the Spatial Distribution of Temporal Behaviors', in *Ninth International Conference on Information Visualization (IV'05)*. IEEE, pp. 799–806.
4. Andrienko, G. *et al.* (2013) *Visual Analytics of Movement*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

5. Souza, Jr, C. *et al.* (2013) 'Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon', *Remote Sensing*. Multidisciplinary Digital Publishing Institute, 5(11), pp. 5493–5513.
6. Monteiro, M. and Rajão, R. (2017) 'Scientists as citizens and knowers in the detection of deforestation in the Amazon', *Social Studies of Science*, Vol. 47(4) 466 –484
7. Qamer, F. M. *et al.* (2016) 'Mapping deforestation and forest degradation patterns in Western Himalaya, Pakistan', *Remote Sensing*, 8(5), pp. 1–17.
8. Sudhakar Reddy, C. *et al.* (2015) 'Multi-source and multi-date mapping of deforestation in Central India (1935-2010) and its implication on standing phytomass carbon pool', *Ecological Indicators*. Elsevier Ltd, 57, pp. 219–227.
9. Barros, R. *et al.* (2014) 'WORKING GROUP II CONTRIBUTION TO THE FIFTH ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE WG II Summary for Policymakers SPM'. Cambridge University Press.
10. Nocke, T. *et al.* (2008a) 'Visualization of Biosphere Changes in the Context of Climate Change', in *International Conference IT and Climate Change, Berlin*.
11. Nocke, T. *et al.* (2008b) 'Visualization of Climate and Climate Change Data: An Overview', *Digital earth summit on geoinformatics 2008: tools for global change research*, 1999(August 2014), pp. 226–332.
12. Li, J., Zhang, K. and Meng, Z. P. (2015) 'Vismate: Interactive visual analysis of station-based observation data on climate changes', *2014 IEEE Conference on Visual Analytics Science and Technology, VAST 2014 - Proceedings*, pp. 133–142.
13. DEFRA (2017) 'Air Quality A Briefing for Directors of Public Health'.
14. Qu, H. *et al.* (2007) 'Visual Analysis of the Air Pollution Problem in Hong Kong', *IEEE Transactions on Visualization and Computer Graphics*, 13(6), pp. 1408–1415.
15. Fecht, D. *et al.* (2016) 'Spatial and temporal associations of road traffic noise and air pollution in London: Implications for epidemiological studies', *Environment International*.
16. Du, Y. *et al.* (2016) 'A Visual Analytics Approach for Station-Based Air Quality Data', *Sensors*, 17(1), p. 3.
17. Zhou, Z. *et al.* (2017) 'Visual Analytics for Spatial Clusters of Air-Quality Data', *IEEE Computer Graphics and Applications*, 37(5), pp. 98–105.
18. Kreuseler, M. (2000) 'Visualization of geographically related multidimensional data in virtual 3D scenes', *Computers & Geosciences*. Pergamon, 26(1), pp. 101–108.
19. Anadiotis, G. (2017) *Raiders of the storm: The data science behind weather prediction | ZDNet*, *ZDNet.com*. Available at: http://www.zdnet.com/article/raiders-of-the-storm-the-data-science-behind-weather-prediction/
20. Diehl, A. *et al.* (2015) 'Visual Analysis of Spatio-Temporal Data: Applications in Weather Forecasting', *Computer Graphics Forum*, 34(3), pp. 381–390.
21. Wickham, H. *et al.* (2012) 'Glyph-maps for visually exploring temporal patterns in climate data and models', *Environmetrics*, 23(5), pp. 382–393.
22. Ferstl, F. *et al.* (2017) 'Time-Hierarchical Clustering and Visualization of Weather Forecast Ensembles', *IEEE Transactions on Visualization and Computer Graphics*, 23(1), pp. 831–840.
23. Helbig, C. *et al.* (2014) 'Concept and workflow for 3D visualization of atmospheric data in a virtual reality environment for analytical approaches', *Environmental Earth Sciences*, 72(10), pp. 3767–3780
24. Tong, X., Li, C. and Shen, H.-W. (2017) 'GlyphLens: View-Dependent Occlusion Management in the Interactive Glyph Visualization', *IEEE Transactions on Visualization and Computer Graphics*, 23(1), pp. 891–900

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication