International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Geovisual analytics of Satellite Image Time Series

Raul Zurita-Milla, Connie Blok, and Vasilios Retsios.
Department of Geo-Information Processing.
Faculty of Geo-Information Science and Earth Observation (ITC).
University of Twente. 7500 AE Enschede. The Netherlands
zurita-milla@itc.nl

Abstract: Satellite image time series provide valuable information on the Earth's dynamics at a variety of spatio-temporal scales. Progress on information and communication technologies has greatly improved the access to such time series. For instance, the GEONETCast system freely distributes near real time raw satellite images and higher level products to end-users all over the word. This explains the ubiquitous use of satellite image time series in a wide range of environmental applications. However, geospatial software and, in particular, geographical information systems (GIS), generally lack robust functionality to visualize and to analyze large time series of data. This paper presents ongoing work to enhance an open source GIS software package, ILWIS, by developing an open source toolbox for the visual and analytical exploration of satellite image time series. Presently, the toolbox offers advanced functionality for animating and interacting with time series, for filtering the data based on area of interest or on attribute values and for aggregating the data to different temporal granularities. Three examples are used to illustrate the functionality. Finally, we show that there is a lot of potential in linking this toolbox with the R open source programming language so that advanced geo-statistical and data mining functionality can also be used during the analysis.

***Keywords**:* Geovisual analytics; animations; data mining; GEONETCast; ILWIS.

1 INTRODUCTION

Most geographic phenomena are dynamic in nature. The location and/or the attributes of geographic entities change through time. Satellite image time series (SITS) provided by past and present sensors allow the study of these changes at a variety of spatio-temporal scales. This explains the ubiquitous use of SITS in a wide range of environmental applications such as land cover mapping (Zurita-Milla et al., 2011), studying of vegetation seasonal dynamics (Zurita-Milla et al., 2009) or the monitoring of glaciers (Ballantyne and Long, 2002). In fact, relatively recent developments on data policy, data processing and data distribution have enormously facilitated the access to SITS by both geoscientists and by the general public. For instance, the GEONETCast system (Mannaerts et al., 2009) freely distributes near real time raw satellite images and higher level products to end-users all over the world.

However, making sense of SITS is far from trivial because we are confronted with vast amounts of data that must be simultaneously analyzed in space and in time. Moreover, we believe that there is a lack of specific tools for the analysis of this type of data because most geospatial software was designed to work with "snapshots" rather than with a "time continuous" view of reality. Because of this, both remote sensing (RS) and geographical information systems (GIS) software still lack robust functionality to visualize and to analyze time series. More generally, one could say that there is a need to develop and to incorporate temporal data models and data analysis methods and techniques to geospatial software and that need can only be expected to grow with increasing attention towards environmental change. This paper presents ongoing work to enhance an open source GIS software package by developing a toolbox for the visual and analytical exploration of SITS (Blok et al., 2011).

The rest of the paper is set up as follows: section 2 describes the relatively new discipline of geovisual analytics which provides the theoretical foundations for our work. Section 3 provides more details on ILWIS and on the implementation of the SITS toolbox. Section 4 contains three examples to illustrate the functionality of the toolbox and, finally, section 5 summarizes the main findings and introduces our future research work.

2 GEOVISUAL ANALYTICS

Extracting information from SITS requires a combination of geovisual and geoanalytical approaches as well as human interpretative skills linked to disciplinary knowledge. With respect to geovisual approaches, these are typically applied to complex spatio-temporal data and/or problems as they usually lead to abductive reasoning which, in turn, results in the formulation of one or more hypotheses that can be further investigated (Blok, 2005). For time series data, geovisualization offers three main possibilities (Kraak, 2000):

- Single map where change events are represented by graphic variables and symbols,
- Series of maps (also known as “small multiples”) where sequential display of successive snapshots is used to depict changes over time, and
- Animated map where change events are deduced from real movements or changes in location and attribute values on the map itself

Out of these three possibilities, animations are preferred to analyze SITS because they easily get user’s attention and can handle long time series. It is relatively straightforward to use animations to qualitatively explain processes and/or to identify spatio-temporal patterns and trends (Turdukulov, 2007). However, animations also have limitations. One of them is known as “change blindness” which is “a striking failure to see large changes that normally would be noticed easily” (Simons and Rensink, 2005).

In an attempt to improve visual exploration of large and complex datasets, Keim and his colleagues (Keim et al., 2006) modified the classical exploration mantra “overview first, zoom and filter, details on demand“ (Shneiderman, 1996) and transformed it into: analyze, show the important, zoom/filter, analyse further, details on demand. This new mantra (paradigm) led to the development of a relatively new discipline named geovisual analytics, GVA (Keim et al., 2008; Andrienko et al., 2007). GVA integrates knowledge, methods and techniques from geovisualization, spatio-temporal data management, spatio-temporal data analysis, human-computer interaction and human perception issues (Andrienko et al., 2007; Andrienko et al, 2010). In short, GVA approaches fit well the requirements for extracting information from SITS.

With respect to the analytical approaches, GVA advocates, amongst others, the use of clustering methods. Clustering is both a pre-processing step that eases the exploration of massive amounts of data and a fundamental data analysis method to extract knowledge (Keim et al., 2008). This is because clustering methods help to reduce the complexity of spatio-temporal data by organising it into different classes.

Other analytical tools like filtering attribute values and temporal aggregation are also useful for analyzing SITS. The former operation can be used to identify anomalies in SITS while the latter one is required to explore geographic phenomena that change at different rates. A flexible temporal aggregation helps user to identify the most appropriate temporal unit to study a phenomena.

3 ILWIS, GEONETCast and the SITS toolbox

The SITS toolbox is implemented in the Integrated Land and Water Information System (ILWIS), which is a RS and GIS software that combines image, vector and thematic data in one desktop package. ILWIS provides much of the functionality that users expect from both RS and GIS packages. Present functionality includes, digitizing, editing, (re-)projection, reformatting, map calculator, spatial analysis, production of cartographically sound map as well as import/export support for the most common RS and GIS file formats.

ILWIS was started in the 1984 by the ITC as a project to support multidisciplinary research projects in developing countries. This means that it was designed to be low-cost, low-entry level and application oriented. After successive improvements leading to new versions, ILWIS was made available as an open source software package (GNU GPL) in 2007. The source code and windows executables can be downloaded from the 52°North website (http://52North.org/). The latest ILWIS version is 3.8 beta 2, which was released in October 2011.

ILWIS was selected for the SITS toolbox because it provides easy access to and management of the various satellite and environmental data and products delivered by the GEONETCast system via a dedicated toolbox. More precisely, it contains a format conversion library that is able to perform the required pre-processing for all the data formats used by the system. Additionally, it incorporates functionality from freeware software to access specialized data formats (e.g. the meteorological BURF and GRIB/GRIB2 formats) and/or to process mission specific data files (e.g. BEAM, BRAT for data from the ENVISAT satellite and VGT-Extract for data from the SPOT-VEGETATION sensor).

The SITS toolbox, written in C++ like ILWIS, is based on a research prototype named aNimVIS (Blok, 2005). This prototype allowed the animation of SITS and added various types of interactions to common media player type of interactions, but there was no analytical support. Since 2010 we work on expanding aNimVis functionality and on integrating it into a comprehensive RS and GIS software package. New functionality includes advanced controls for animating and interacting with time series, for filtering the data based area of interest and on attribute values and for aggregating the data to different temporal granularities.
Most of the rendering of raster images and the execution of the animations is done with OpenGL, which is an industry standard for visualizations commonly installed on desktop computers. This choice increases the performance although the available graphics RAM in the computer determines the final results.

4 EXAMPLES

Three examples are shown in this section to illustrate the functionality of the SITS ILWIS toolbox. For these examples we downloaded SPOT-VEGETATION data using the GEONETCast ILWIS toolbox. The VEGETATION sensor delivers daily images in four spectral bands and at a 1 km spatial resolution since April 1998. More precisely, we downloaded 10-day Normalized Difference Vegetation Index (NDVI) images. These composite images are produced using the maximum value composite method to minimize the impact of clouds. NDVI images were selected because this vegetation index (combination of two or more spectral bands) is one of the most common ones to detect vegetation vigour and “greenness” (Dorigo et al. 2007) The NDVI is computed as follows:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{R}}{\mathrm{NIR} + \mathrm{R}} \quad (1)$$

Where NIR and R are the reflectances in the Near InfraRed and Red spectral bands

4.1 Change blindness

A method to reduce change blindness based on the Gestalt figure-ground principles (MacEachren, 1995) was implemented in the SITS toolbox. The user can select an area or a range of attribute values of interest (the 'figure') and the toolbox will subdue the rest of the image (the 'ground'). This reduces the cognitive load, and enhances to focus attention on relevant aspects. At the same time, changes happening in the surroundings of the selection remain visible, so the context is not lost. This will facilitate the exploration of the data. Figure 1 illustrates a selection of area and of attribute values. A combined selection of area of interest and range of attribute values is also possible.

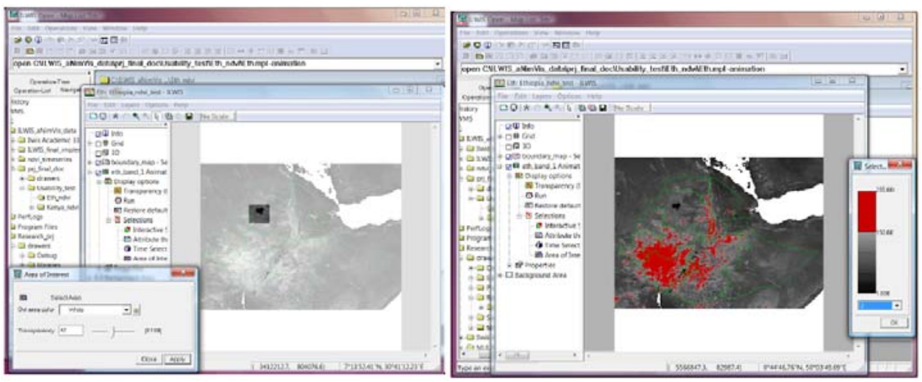

Figure 1. Selection of an area of interest (left) or of a range of NDVI values (right).

The selection of the area of interest is implemented using transparency options so that the analyst can focus his/her attention without losing the geographical context. Transparency can be properly implemented using a layering technique that minimizes computer processing costs and avoids having to pre-process the SITS before their visualization. A preliminary usability test of this new functionality was conducted (Mekonnen, 2011) in which the participants who made use of these interactive tools performed their tasks better (faster and with a higher degree of satisfaction) that those without access to them.

4.2 Time granularity

As briefly explained at the end of Section 2, geographic phenomena change at different rates. However, the decision to study a given phenomena mostly depends on the data availability (Dewi, 2012) whereas the exploration of SITS at different temporal intervals (daily, weekly, monthly, etc) results in significantly different patterns (Harrower et al., 2000). This justifies the development of time aggregation functionality in our toolbox. The implemented functionality allows regular time aggregation, for example, make one image out of three. Alternatively, irregular time selection is also possible. This is, displaying the images fulfilling a certain condition, for instance NDVI values larger than a given threshold. Figure 2 shows an example of a regular aggregation of the SPOT-VEGETATION data from its original 10-day interval to monthly aggregates. This helps to reduce or overcome issues with residual pre-processing errors or missing data due to cloud coverage.

Moreover, temporal aggregation also reduces the number of images that must be visualised; therefore helping to focus the user's attention.

Temporal aggregation can also be useful to explore and compare different types of vegetation because they typically present different phenological dynamics in a season. For instance, Figure 3 shows the locations of high NDVI values at a given moment of the growing season. This was done using the irregular time selection functionality.

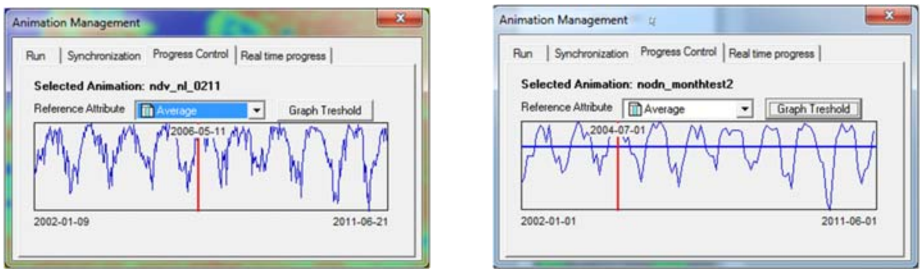

Figure 2. Average NDVI profile over The Netherlands at the original temporal granularity (10-day composites; left) and the aggregated monthly values (right).

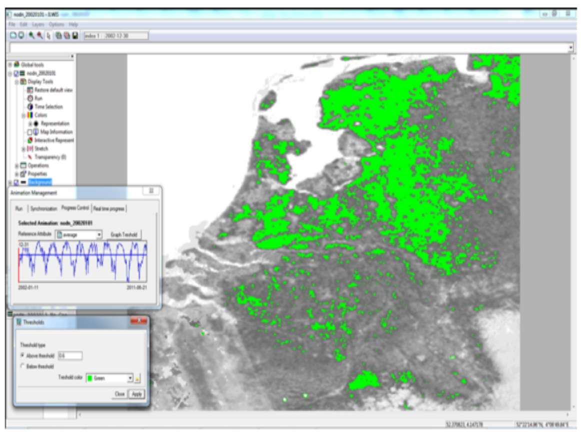

Figure 3. shows the visualization of attribute values fulfilling a condition (a user-defined NDVI threshold) as 'figure' and the evolution of other NDVI values as subdued 'ground'.

4.3 Phenological patterns

Finally, we present an application where analytical support for the analysis of SITS is drawn from the R open source statistical software (R Development Core Team, 2011). Many time series packages are available for this software and this is why we presently study how to couple ILWIS and R. Here we present an example based on the *Kohonen* package. This package contains an efficient implementation of Self-Organized Maps (SOMs). SOMs are often used in exploratory data analysis and data mining studies because they can identify patterns and clusters in the input data in an unsupervised fashion. While doing this, they map the (multi-dimensional) input data into a two dimensional space in which we can visualize the results. An interesting characteristic is that this two dimensional space preserves the topology of the input data.

In this example, SOMs are used to identify the main spatio-temporal patterns in the Kruger National Park, South Africa (Zurita-Milla et al., 2012). To do so, we trained a SOM with 36 neurons using 13 years (April 1998 till March 2011) of VEGETATION NDVI data. The training process resulted in a set of topologically ordered NDVI states and in a mapping of the input data into such states as small multiples. Figure 4 shows the main NDVI states found after processing 468 images (13 years * 36 images/year).

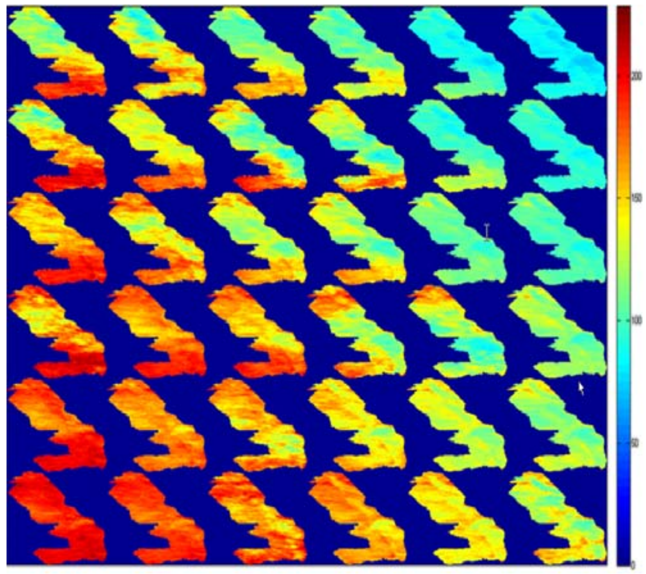

Figure 4. Main 36 synoptic NDVI states found in the Kruger National Park, South Africa.

Figure 5 illustrates the states that occurred in the first complete vegetation season (October 2004 till September 2005). These states offer a simplified representation of the dynamics as seen by the SPOT-VEGETATION sensor. Such a representation facilitates, for instance, the study of inter-annual differences in vegetation phenology as it reduces the cognitive load of visualizing vegetation dynamics in SITS. In this case, it is straightforward to see that the season at hand was very dry as it did not reach any of the high NDVI synoptic states shown in the lower left corner of Figure 4.

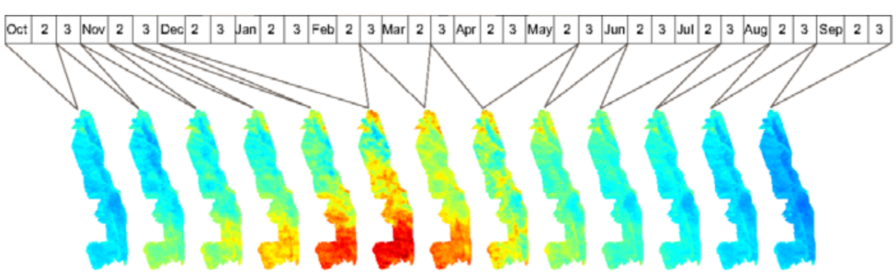

Figure 5. Synoptic states that occurred in the vegetation season October 2004 till September 2005.

5 CONCLUSIONS

This paper reports ongoing work to implement and expand a geovisual analytics-inspired toolbox for the exploration of large time series of satellite imagery. The toolbox is implemented in an open source RS and GIS software package named ILWIS. This offers two main benefits: on one side, the users benefit from existing RS and GIS functionalities and from an easy access to SITS via the GEONETCast system. On the other side, users can create their own community to expand and improve visual and/or analytical functionalities.

The geovisual analytic paradigm was chosen because proper analysis of SITS requires not only a visualization, but also the use of geocomputational methods and the interaction of experts with domain knowledge. Three examples were presented to illustrate the possibilities of the toolbox as an independent package (examples in 4.1 and 4.2) and to further analyse data coming from external specialised software (example in 4.3). In particular, the small multiples produced by the SOM algorithm reduce the cognitive overload created by long SITS (e.g. from 468 to 36 images in our example) and the vegetation dynamics are better shown by (aggregated) animations. Hence, we conclude that the toolbox has a high potential to extract valuable information from SITS as tools can be selected depending on the aim, type of user questions, and available data.

At the same time, we recognise that the toolbox has some limitations and further work is needed to overcome them. For instance, we need to reduce the number of floating dialogs because they might interfere with the visual data mining process. We want to minimise split attention problems, which occurs when one tries to visualise the animation whilst simultaneously consulting the attribute values of the legend. Additionally, we would like to explore smoothing techniques to make the animations appear more “natural” after aggregating them in time. And, we need to identify a generic way to extract time from the metadata or filename of the SITS. This can be complicated since there is no universal way to provide time amongst the different SITS providers. Finally, we plan to improve options for comparing and synchronizing time series and we would like to investigate ways of integrating specialised software (e.g. the R open source statistical package) into our toolbox.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Mekonnen Metaferia, Ms. Rani Charisma Dewi, Dr. Ulanbek Turdukulov and Mr. Martin Schouwenburg for their valuable input on the development of the open source toolbox described in this paper. Dr. Anton Vrieling is also thanked for providing the pre-processed NDVI time series used in the phenology example.

REFERENCES

Andrienko, G., N. Andrienko, P.Jankowski, D. Keim, M.-J. Kraak, A. M MacEachren and S. Wrobel. Geovisual analytics for spatial decision support: setting the research agenda. *International Journal of Geographical Information Science*, 21(8), 839–857, 2007.

Andrienko, G., N. Andrienko, J. Dykes, M.-J. Kraak and H. Schumann, GeoVA(t) – geospatial visual analytics: focus on time. *Journal of Location Based Services* 4(3), 141–146, 2010.

Ballantyne, J., and D.G. Long. A multidecadal study of the number of Antarctic icebergs using scatterometer data, paper presented at IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Toronto, Canada, June 24-28, 2002.

Blok, C. Dynamic visualization variables in animation to support monitoring of spatial phenomena. PhD thesis, Universiteit Utrecht, ITC, Utrecht, Enschede. ITC Dissertation 119, 2005.

Blok, C.A., U.D. Turdukulov, R. Zurita-Milla, V. Retsios, M. Schouwenburg, and M. Metaferia, Development of an open-source toolbox for the analysis and visualization of remotely sensed time series. *Cartographica* 46(4), 227–238, 2011.

Dewi, R.C. Supporting visual exploration of animated time series with modifiable temporal unit. MSc thesis. Faculty ITC. University of Twente. 2012.

Dorigo, W.A., R. Zurita-Milla, A.J.W. de Wit, J. Brazile, R. Singh, and M.E. Schaepman, A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modelling, *International Journal of Applied Earth Observation and Geoinformation*, 9(2): 165-193, 2007.

Harrower, M., A. MacEachren and A. L. Griffin. Developing a geographic visualization tool to support earth science learning. cartography and geographic information science. 27(4): 279-293, 2000.

Keim, D., G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer and G. Melançon, Visual analytics: definition, process and challenges, Information Visualization – Human-Centered Issues and Perspectives, Lecture Notes in Computer Science, Springer, 154-175, 2008.

Kraak, M.-J., Visualisation of the time dimension. Time in GIS: Issues in spatio-temporal modelling, 47:27–35, 2000.

Mannaerts, C.M., B.P.M. Maathuis, M. Molenaar and R. Lemmens, The ITC GEONETCast toolbox: a GEO capacity building component for education and training in global Earth observation and geo-information provision to society. Paper presented at IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Cape Town, South Africa. July 13-17, 2009.

MacEachren, A.M. How maps work: representation, visualization, and design. Guilford, New York, 1995.

Metaferia, M.T., Visual enhancement of animated time series to reduce change blindness. MSc thesis. Faculty ITC. University of Twente, 2011.

R Development Core Team, R: a language and environment for statistical computing, R Foundation for Statistical Computing (URL http://www.R-project.org/), Vienna, Austria, 2011.

Simons, D. J. and R. A. Rensink, Change Blindness: Past, Present, and Future, *Trends in Cognitive Sciences*, 9(1), 16–20, 2005.

Shneiderman, B., The eyes have it: a task by data type taxonomy for information visualization, paper presented at IEEE Symposium on Visual Languages, Boulder, Colorado. 336–43, 1996.

Turdukulov, U.D., Visualizing the evolution of image features in time-series: supporting the exploration of sensor data. PhD thesis, Universiteit Utrecht, ITC, Utrecht, Enschede. ITC Dissertation 149, 2007.

Zurita-Milla, R., G. Kaiser, J.P.G.W. Clevers, W. Schneider, and M.E. Schaepman, Downscaling MERIS FR data for land cover mapping and vegetation monitoring of heterogeneous landscapes, *Remote Sensing of Environment*, 113 (9): 1874–1885, 2009.

Zurita-Milla, R., L. Gomez-Chova, L. Guanter, G. Camps-Valls and J.G.P.W. Clevers, Multi-temporal unmixing of medium spatial resolution satellite images: a case study using MERIS images for land-cover mapping, *IEEE Transactions on Geosciences and Remote Sensing*, in press, 2011.

Zurita-Milla, R., J.A.E. van Gijsel, N.A.S. Hamm, P.W.M, Augustijn and A. Vrieling, Exploring spatio-temporal phenological patterns and trajectories using self organizing maps, *IEEE Transactions on Geosciences and Remote Sensing*, under review, 2012.