Workshop on Visualisation in Environmental Sciences (EnvirVis) (2016)
A. Middel, K. Rink, and D. Zeckzer (Editors)

Short Paper

Towards Visual Analytics for Multi-Sensor Analysis of Remote Sensing Archives

D. Eggert[1], M. Sips[1,2] & P. Köthur[1]

[1] GFZ German Research Centre for Geosciences, Potsdam, Germany
[2] Humboldt University Berlin, Germany

Abstract

To better detect and study processes on the Earth's surface, scientists want to combine various satellite data and extract potentially interesting patterns from the combined data. This analysis approach is called multi-sensor analysis. In this paper, we present an interactive visual exploration solution for the first important step of multi-sensor analysis: the assessment and selection of remote sensing scenes. This solution is the first step towards a larger Visual Analytics (VA) approach that turns multi-sensor analysis into a transparent and interactive analysis method. We conduct our research in the context of GeoMultiSens, which is an interdisciplinary research project between remote sensing, computer science and VA experts. To demonstrate the utility of our visual exploration solution, we use a real-world scenario: the assessment and selection of scenes in order to study the change of forest cover in Europe. The application example indicates that interactive visual exploration facilitates a structured assessment of the quantity and quality of remote sensing scenes and enables scientists to exclude low-quality scenes from subsequent multi-sensor analysis.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—

1. Introduction

Climate change, population growth, water scarcity, and loss of biodiversity are pressing global challenges. Satellite remote sensing systems can observe processes on the Earth's surface associated with these challenges. Therefore, the continuous monitoring of these processes through satellite-based remote sensing missions plays a key role in the development of strategies to address these challenges.

Science organizations and commercial companies have conducted more than 300 satellite-based remote sensing missions to monitor the Earth's surface [Liu15]. Several satellite operators provide public access to the mission data via web portals. The mission data consist of satellite images called scenes. Scenes from different archives often differ in the spatial resolution, acquisition times and spectral channels of recorded scenes [AW09]. To better detect and study a variety of Earth surface processes, scientists want to combine the scenes of different archives and extract potentially interesting patterns from the combined scenes. Scientists call this concept multi-sensor analysis. Scientists conduct multi-sensor analysis to answer a particular question for a specific region of interest (ROI).

One approach to get such an answer is to implement a fully automated method that performs all processing steps as a black box without user feedback. However, such an approach has major disadvantages. It does not provide information about the quantity and quality of the remote sensing scenes. The remote sensing scenes available for a ROI may be unevenly distributed in space and time. Scenes may also contain missing values caused by cloud cover. To be able to assess the scientific merits of multi-sensor analysis results, scientists must have detailed knowledge about such deficiencies. Furthermore, multi-sensor algorithms utilize a variety of strategies to cope with incomplete and unevenly distributed input data, e.g. spatial or temporal interpolation. To fully exploit the available data, scientists need to be aware of these strategies and be able to refine algorithmic decisions where necessary. Finally, scientists need to explore the extracted patterns to decide whether these are associated with certain processes on the Earth's surface.

Visual Analytics (VA) has the potential to increase user involvement in analytical computations [Fis11,FPDs12,MPG*14,SPG14]. Thus, our long-term research focus is to develop a VA approach that enables transparent multi-sensor analysis of remote sensing scenes. More precisely, we want to enable steering of multi-sensor analysis through VA. We conduct our research in the context of GeoMultiSens [Geo16], which is an interdisciplinary research project between remote sensing, computer science and VA experts. One aim of GeoMultiSens is to develop a VA approach that supports (a) assessment of quantity and quality of remote sensing scenes, (b) exploration of partial results during the execution of multi-sensor algorithms and refinement of algorithmic decisions, and (c) exploration and interpretation of the analysis results.

DOI: 10.2312/envirvis.20161100

In this paper, we present an interactive visual exploration solution for the first important step of multi-sensor analysis: the assessment and selection of available remote sensing scenes. It comprises three main steps. First, the available satellite missions and corresponding number of scenes are retrieved for a user-specified ROI. Our interactive exploration solution provides a visual overview of this information. The second step is the assessment of the spatial and temporal distribution of the scenes provided by user-selected satellite missions. A heat map and a temporal histogram view [DMK05] enable scientists to obtain an overview and detect time intervals or geographic regions with poor or high data density. The third step is the filtering of low quality scenes to increase overall input data quality. Scientists may focus on a specific time interval or adjust certain quality criteria of the scenes, such as the maximum cloud cover allowed in a scene. Interactive linked views allow scientists to explore the impact of their decisions for the ROI.

The contributions of this paper are the following:

- We present an interactive visual exploration solution for the assessment and selection of the input scenes to multi-sensor analysis. It is the first step towards a larger VA approach that turns multi-sensor analysis into a transparent and interactive analysis method.
- Our exploration concept facilitates a structured assessment of the quantity and quality of tens of thousands of scenes. This structured assessment enables scientists to increase the overall input data quality.
- We present the utility of our solution in a real-world scenario: the assessment and selection of scenes to study the change of forest cover in Europe.

2. State of the Art

As indicated above many operators provide public access to their satellite data via web portals. Most portals provide basic query functionality to support users in finding the data they need. These web portals also provide some visual representations of the data. In contrast to our approach, these visualizations do not support a comprehensive exploration and assessment of the spatial distribution and temporal distributions of tens of thousands of scenes.

The Earth Resources Observation and Science (EROS) center, as part of the U.S. Geological Survey (USGS), provides access to one of the biggest archives for satellite remote sensing data. Users can access the data through the EarthExplorer website [USG16]. The archive includes more than 250 datasets representing various satellite missions and derived data products. Another portal is the Scientific Data Hub [ESA16], which hosts the satellite data of the ESA's Sentinel missions. The main purpose of both portals is to provide basic query functionality for users. They focus on finding scenes for a particular ROI and support the workflow of setting and refining the query parameters. The query parameters include time interval and additional quality constraints, such as the maximum cloud cover allowed in a scene. However, they only present the query result to users as a list of scenes. Assessment of the spatial and temporal distribution of a large number of scenes is not sufficiently supported, which also renders it very difficult to judge the impact of changing quality parameters on the data distribution.

A more holistic approach is Google's Earth Engine [Goo16] portal. Earth Engine is a cloud-based platform hosting many of the freely available mission archives like Landsat and Sentinel-1. The Earth Engine not only provides access to data but also allows users to conduct computations. However, it does not support comprehensive assessment of the spatial and temporal distribution of the input scenes for these computations.

3. Visual analytics for interactive assessment of remote sensing scenes

In this section, we discuss the requirements for our interactive exploration solution and demonstrate the utility of our solution in a real-world scenario.

3.1. Requirements

We applied a user- and task-based design approach [DKS*10] in our ongoing research project between VA and remote sensing experts. It enabled us to elicit design requirements (DRs) for the interactive assessment of remote sensing scenes: As mentioned in Section 1, scientists must be able to study the spatial and temporal distribution as well as the quality of the available satellite data. The DRs for supporting this task are:

DR1 Enable intuitive definition of a region of interest.
DR2 Provide overview of available data sets (satellite missions with specific sensors).
DR3 Provide information about the number of available scenes for each data set.
DR4 Enable visual assessment of the spatial and temporal distribution of chosen scenes.
DR5 Allow for interactive filtering of scenes based on quality constraints.

3.2. Approach and Application Scenario

In this section, we describe our approach and demonstrate its benefits with a specific application example: exploring and assessing the quality of remote sensing scenes to study forest cover change in Europe between 2010 and 2014. Studying forest cover change is important. For example, Foley et. al. [FDA*05] discuss the negative impacts of forest cover loss on the environment, like undermining the capacity of ecosystems to maintain freshwater and regulate climate and air quality. Subsequently it has motivated the development of various analysis methods; e.g., Hansen et. al. [HPM*13].

We implemented our interactive exploration solution as a web-based tool, which queries online mission data archives to obtain important meta-data of available scenes, e.g., spatial extent, acquisition date and cloud cover. We use these meta-data to support the visual assessment of remote sensing scenes.

To define the region of interest (ROI), we simply have to draw a rectangle around Europe in the interactive map element of the tool (Figure 1, (1)) (**DR1**). To enable fluent exploration, a list of available mission data sets for Europe appears adjacent to the ROI (Figure 1, (2)) (**DR2**). The list also depicts the number of scenes for each data set (**DR3**). Note that the data set list is sorted in ascending order according to the number of available scenes. This enables

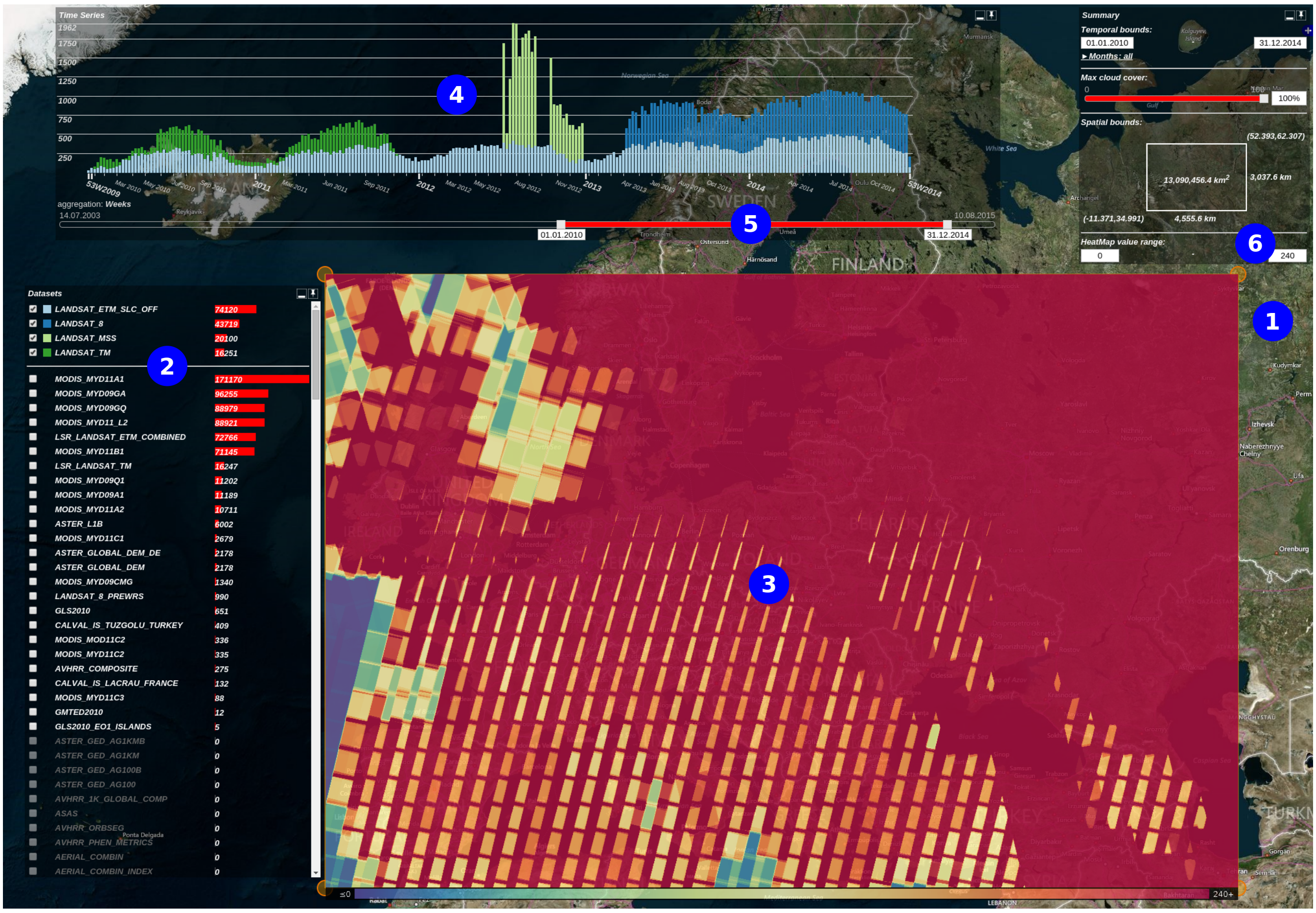

Figure 1: *Visual interface of our interactive exploration solution, including ① interactive map element for intuitive ROI definition, ② list of data sets contained in the ROI, ③ heat map view providing overview over the spatial distribution, ④ histogram view visualizing the temporal distribution of scenes, ⑤ date range slider to constrain the relevant time period and ⑥ input field setting the heat map's color scale.*

us to assess the potential contribution of particular data sets to the input data volume.

The spatial resolution of the scenes plays a crucial role when studying forest cover change [HPM*13]. Therefore, we select the freely available Landsat-8 and Landsat-7 data sets since they provide scenes with a good spatial resolution. When data sets are selected, our approach provides two linked visualizations to support assessment of the data distribution (**DR4**). A heat map, superimposed on the ROI, indicates the number of available scenes for each pixel in the ROI (Figure 1, (3)). A histogram view visualizes the temporal distribution of scenes (Figure 1, (4)). The histogram view uses stacked bars to depict the contribution of individual data sets to the total amount of scenes. Using the provided date range sliders (Figure 1, (5)), we constrain the time period to 2010-2014 in order to meet our study interval. The inspection of the histogram view shows that Landsat-8 data is only available from 2013 on. To improve the data density in the years 2010 to 2012, we add the freely available Landsat TM and Landsat MSS data sets to our selection.

We consider 240 scenes for each image pixel – an average of four scenes per month – as a good basis for studying forest cover change in Europe. We set the maximum value of the heat map's color scale to 240 using the corresponding input field (Figure 1, (6)). As a consequence, the red regions represent areas in the heat map that meet our initial requirements regarding data density. We can observe that although the density of scenes varies, the major parts of Europe are covered by 240 or more scenes. For our scenario, we conclude that the chosen data sets provide a good starting point for forest cover change analysis.

Moreover, the good spatial and temporal data density allows us to reduce the maximum cloud cover allowed per scene. The lower the cloud cover, the lower the number of missing values in the satellite data. Therefore, excluding low-quality input data from subsequent analysis may contribute to more reliable analysis results. We use the corresponding slider to adjust the maximum cloud cover allowed per scene from 100% to 50% (Figure 2, (1)) (**DR5**). The impact of this filtering step on the data distribution can be assessed in the linked heat map and histogram view. In the latter, the status before the filtering step is indicated by gray bars in the background (Figure 2, (2)), which represent the number of scenes for the unconstrained case (maximum allowed cloud cover of 100%). The

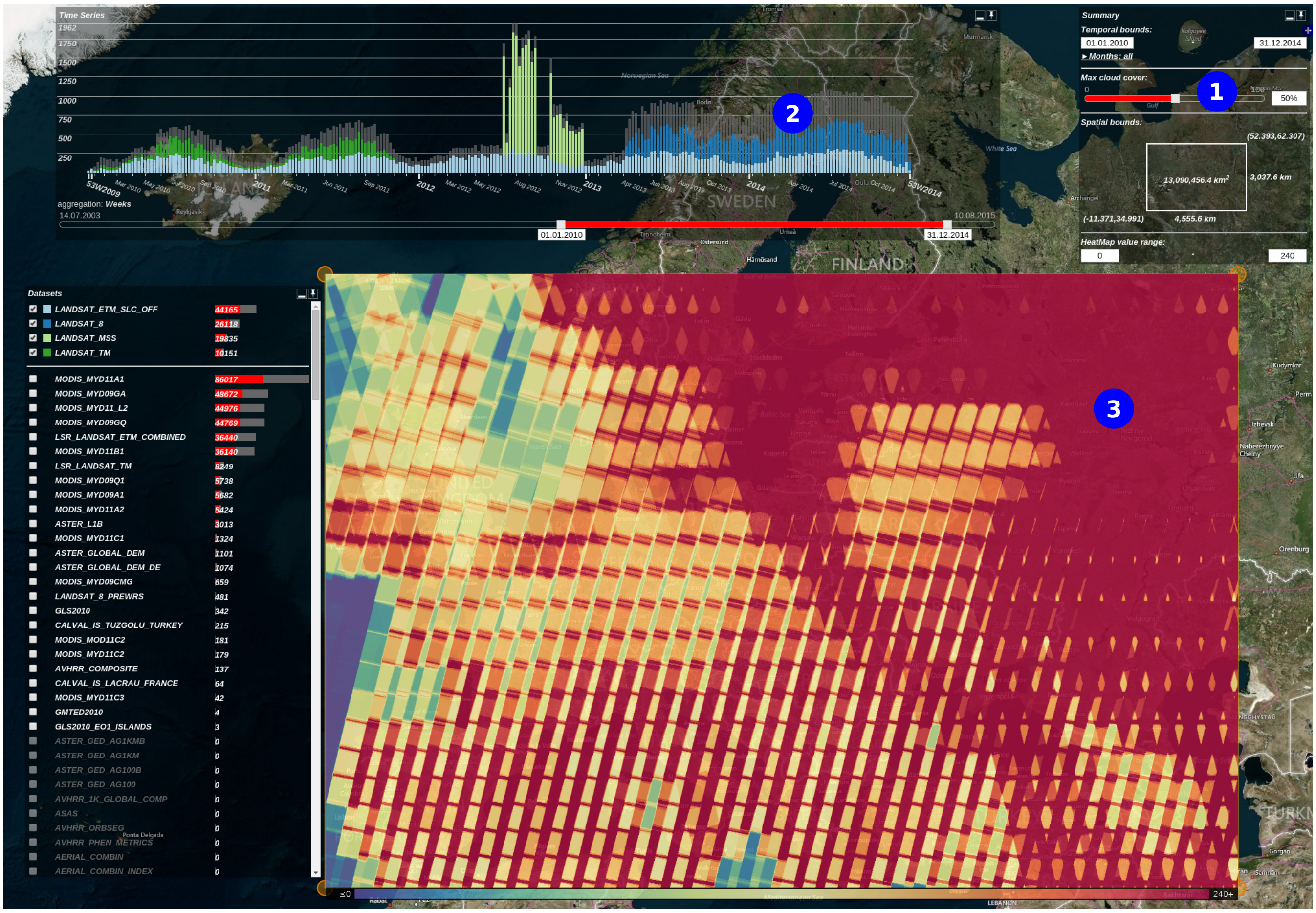

Figure 2: *Visual interface of our interactive exploration solution with adjusted maximum cloud cover parameter, including ① slider to adjust the maximum cloud cover allowed per scene, ② gray background bars indicating the number of scenes for the unconstrained case and ③ densely covered region revealed by the heat map.*

histogram view indicates that the number of available scenes per week is reduced by up to 50%. The assessment of the heat map reveals that the overall data density is reduced and that the eastern part of Europe is more densely covered (Figure 2, (3)). Nevertheless, we consider this an appropriate trade-off between quantity and quality. Moreover, we became aware of potential deficiencies regarding data density through the visual exploration.

4. Conclusion

The trend over the recent years shows that the major agencies NASA, USGS, and ESA will provide open access to future mission data. Therefore, multi-sensor approaches will become increasingly important to analyze processes associated with changes on the Earth's surface. This development will increase the need to turn multi-sensor analysis into a transparent analysis method.

In this paper, we took the first step towards a larger VA approach that contributes to a robust multi-sensor analysis and a reliable interpretation of the analysis results. In particular, we presented an interactive visual exploration solution for the assessment and selection of remote sensing scenes. We demonstrated the utility of our solution with a real-world scenario: assessment and selection of scenes for forest cover change analysis in Europe. The results indicate that interactive visual exploration facilitates a structured assessment of the quantity and quality of input data and enables scientists to exclude low-quality scenes from subsequent multi-sensor analysis.

The results obtained in the real-world scenario encourage us to continue our research with the remaining crucial components of VA for multi-sensor analysis. We plan to develop solutions that enable (a) interactive visual exploration of partial results and steering of multi-sensor algorithms, and (b) the assessment of analytical results. Once all system components are developed and integrated we are going to conduct further evaluation studies of the overall system.

5. Acknowledgement

We thank our partners in the GeoMultiSens consortium for fruitful discussions and suggestions. This research is funded by the German Federal Ministry of Education and Research (BMBF project GeoMultiSens, 01IS14010A).

References

[AW09] ALBERTZ J., WIGGENHAGEN M.: *Guide for Photogrammetry and Remote Sensing*. Wichmann, 2009. 1

[DKS*10] DRANSCH D., KÖTHUR P., SCHULTE S., KLEMANN V., DOBSLAW H.: Assessing the quality of geoscientific simulation models with visual analytics methods – a design study. *International Journal of Geographical Information Science 24*, 10 (2010), 1459–1479. doi:10.1080/13658816.2010.510800. 2

[DMK05] DYKES J., MACEACHREN A. M., KRAAK M.-J.: *Exploring Geovisualization*. Elsevier, 2005. 2

[ESA16] ESA: Scientific data hub, 2016. [Online; accessed 3-February-2016]. URL: https://scihub.copernicus.eu/. 2

[FDA*05] FOLEY J. A., DEFRIES R., ASNER G. P., BARFORD C., BONAN G., CARPENTER S. R., CHAPIN F. S., COE M. T., DAILY G. C., GIBBS H. K., HELKOWSKI J. H., HOLLOWAY T., HOWARD E. A., KUCHARIK C. J., MONFREDA C., PATZ J. A., PRENTICE I. C., RAMANKUTTY N., SNYDER P. K.: Global consequences of land use. *Science 309*, 5734 (2005), 570–574. doi:10.1126/science.1111772. 2

[Fis11] FISHER D.: Incremental, approximate database queries and uncertainty for exploratory visualization. In *IEEE Symposium on Large Data Analysis and Visualization (LDAV'11)* (2011), IEEE. 1

[FPDs12] FISHER D., POPOV I., DRUCKER S., SCHRAEFEL M.: Trust me, i'm partially right: Incremental visualization lets analysts explore large datasets faster. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)* (New York, NY, USA, 2012), ACM, pp. 1673–1682. 1

[Geo16] GEOMULTISENS: Project website, 2016. [Online; accessed 3-February-2016]. URL: http://www.geomultisens.gfz-potsdam.de. 1

[Goo16] GOOGLE: Earth engine, 2016. [Online; accessed 3-February-2016]. URL: https://code.earthengine.google.com/. 2

[HPM*13] HANSEN M. C., POTAPOV P. V., MOORE R., HANCHER M., TURUBANOVA S. A., TYUKAVINA A., THAU D., STEHMAN S. V., GOETZ S. J., LOVELAND T. R., KOMMAREDDY A., EGOROV A., CHINI L., JUSTICE C. O., TOWNSHEND J. R. G.: High-resolution global maps of 21st-century forest cover change. *Science 342*, 6160 (2013), 850–853. doi:10.1126/science.1244693. 2, 3

[Liu15] LIU P.: A survey of remote-sensing big data. *Frontiers in Environmental Science 3*, 45 (2015). doi:10.3389/fenvs.2015.00045. 1

[MPG*14] MUHLBACHER T., PIRINGER H., GRATZL S., SEDLMAIR M., STREIT M.: Opening the black box: Strategies for increased user involvement in existing algorithm implementations. *Visualization and Computer Graphics, IEEE Transactions on 20*, 12 (2014), 1643–1652. 1

[SPG14] STOLPER C. D., PERER A., GOTZ D.: Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec 2014), 1653–1662. 1

[USG16] USGS: Earthexplorer, 2016. [Online; accessed 3-February-2016]. URL: http://earthexplorer.usgs.gov/. 2