Spatio-Temporal Discussion Board

Gustavo A. García-Chapeton
Faculty of Geo-information
Science and Earth Observation
ITC, University of Twente
Enschede, The Netherlands
g.a.garciachapeton@utwente.nl

Frank O. Ostermann
Faculty of Geo-information
Science and Earth Observation
ITC, University of Twente
Enschede, The Netherlands
f.o.ostermann@utwente.nl

Abstract

In this paper, we address the particular challenges for collaborative, multi-disciplinary analysis posed by today's data rich environments, by designing the Spatio-Temporal Discussion Board, a method for collaborative long-term analysis in Geo-Visual Analytics environments. Our hypothesis is that having a clear defined context and boundary for discussions will help in focusing analysts' attention, and elicit sensible information. For this aim, the method enables users to define explicit context and boundaries for map-based discussions, which is particularly useful for data spaces that have a large spatial and temporal extent. Within these boundaries, analysts can build knowledge by exchanging pieces of information in the form of written messages, playing roles such as background information, question, hypothesis, evidence and conclusions. Additionally, this messages can be enriched with annotations over the visual products, and attached files. As proof of concept, a prototype was developed and the results show that currently available web-technologies are suitable to implement the proposed method.

Keywords: Geo-visual Analytics, Collaborative Analysis, Long-term analysis, Spatio-temporal data, Discussion board.

1 Introduction

This paper addresses the particular challenges for collaborative, multi-disciplinary analysis posed by today's data-rich environments. It proposes a novel approach to improve collaborative analysis of large and varied spatio-temporal data by adding a means of spatially and temporally anchoring information exchange between collaborating analysts. The work builds on theories of information exchange through argumentation and research on collaborative geo-visual analysis to develop an original method to enrich distributed geo-visual analysis environments with a spatio-temporal discussion board.

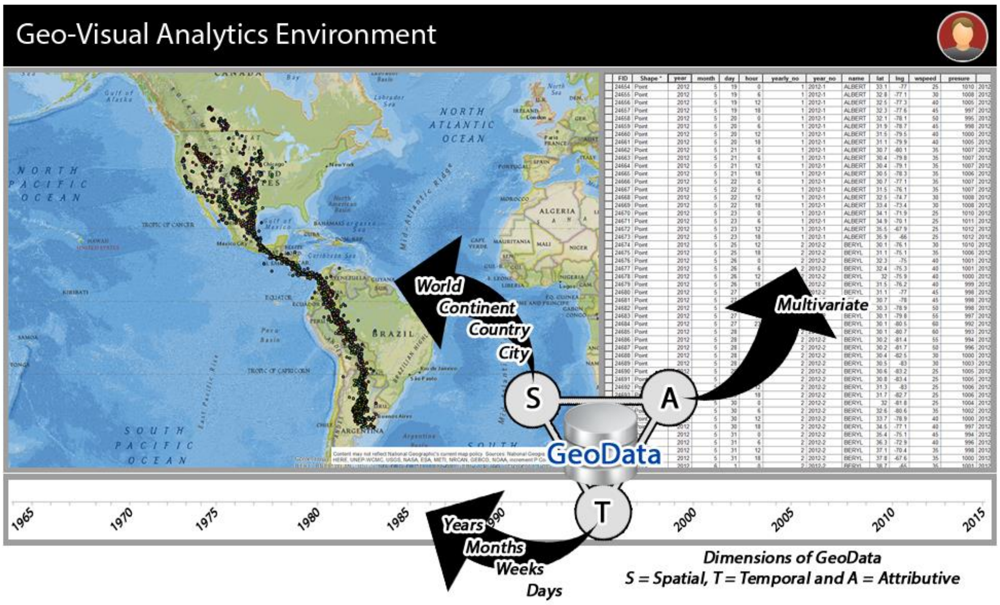
The rationale for this research are the increasing volume of available spatio-temporal data, the heterogeneity of this data, and the complexity and multi-disciplinary character of many current research challenges, which require a team of researchers, practitioners, and also stakeholders who may be laypersons, to team up and work together. It is a well-established fact that several ground breaking advances in geospatial technologies (e.g. small GPS-enabled devices, high-resolution remote sensors, linking of geo-web services) have led to more and larger geospatial datasets. While this is clearly an opportunity to increase knowledge and understanding of natural and socio-economic processes, it also means that it has become common for analysts to face the challenging task of analysing datasets composed of multivariate data covering a large spatial and temporal extent (See Figure 1). Examples are: a 150 years of U.S. census data used in [6]; the OECD regional database from 1960 to the present, containing around 50 indicators for 1700 sub-regions in the 34 OECD countries [12]; the hurricane dataset of UNISYS from 1851 to the present, covering the Atlantic, Indian and Pacific oceans [16]; and the varied user generated geographic content (UGGC) available through public application programming interfaces (API) of social media platforms, or dedicated citizen science project web pages.

Analysing such datasets is a complex task that requires both human and machine input and contribution. Computers are capable to store and process large datasets, and use computational methods to identify recurrent patterns, trends and outliers. However, only humans are currently capable to formulate hypothesis, provide evidence, draw conclusions, discuss strategies, and take decision based on the data [1]. One approach to leverage this combined intelligence is geo-visual analytics (GVA), which aims to produce a true synergy between human analytical skills and computer storage and processing power, and to support effective understanding, reasoning and decision making on the basis of large and complex spatio-temporal datasets. Since phenomena in geographic space occur or evolve in time, GVA has put special emphasis on this relationship between space and time [2].

This research focuses on advancing the support for collaborative and long-term analysis in GVA environments. Collaboration is one of the grand challenges for GVA, because analysts increasingly face large, complex, ill-defined, and broadly scoped problems [1, 9, 15]. Additionally, analysis is often a complex, multi-staged and dynamic task that resembles a long-term process [9, 10].

In this paper, we introduce the Spatio-Temporal Discussion Board (STDB) as a method for collaborative long-term analysis of spatio-temporal phenomena. This method aims to support users in performing analysis through discussions bounded in space and time. Providing in this way a mechanism to better organize discussions in GVA environments, which can be of particular interest when the dataset covers a broad spatial and temporal extension.

Figure 1: A birds' migration dataset covering the American continent, over several years [3] as an example of a multivariate dataset with a large spatial and temporal extent, i.e. up to thousands of variables, spatially distributed over cities, countries, and continents, and spanning in time from several days to decades.



Source: the authors.

Although in principle being application independent, we focus our method on the context of species' populations' dynamics. This application domain is of great importance because a better understanding of species' populations' dynamics can for example: lead to develop eco-friendly and cost-effective pest controls strategies, with direct impact in food security; or, help in designing effective conservation policies for the protection of endangered species; among others.

2 Related work

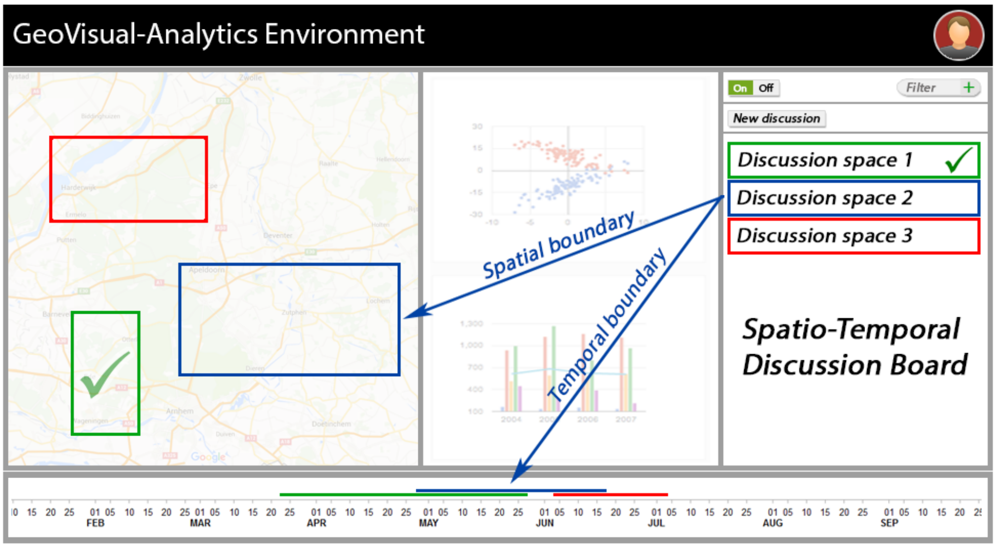
The design of the STDB is based on the principles of Distributed Cognition (DC) and Argument Mapping Theory (AMT). DC provides a framework in which cognition is conceived as a social process, involving many human actors as thinking entities and artifacts as means for knowledge exchange and shared memory, e.g., visual products and textual messages. Additionally, it recognizes that cognition can be distributed over time, and therefore, provides theoretical support for long-term cognitive systems [8]. AMT provides understanding on the exchange of pieces of information playing specific roles in the process of argumentative knowledge construction, e.g., hypothesis and evidence [14]. These principles have been applied in the design of methods for discussion and annotation, which are the fundamental building blocks of the STDB.

Discussion is a general-purpose method for collaborative knowledge-building, through the exchange of pieces of information in the form of oral or textual messages. In GVA, methods for discussion can already be found in the form of chat (e.g., GeoJabber [4]), discussion board (e.g., Big Board [5]) and geo-located discussion (e.g., Twitter-based Geocollaboration system [13]). To effectively support analysis of geographic phenomena, these methods include features to provide geographic context to the discussions. For example, GeoJabber allows analysts to exchange special messages containing a Snapshot[1] of their analytical environment, therefore allowing them to share their perspective of the data. In Big Board and the Twitter-based system, messages are linked to a coordinate in the map, which provides explicit geographic context. However, these systems lack features to explicitly set the context of a discussion, which should include spatial, temporal and thematic components. Additionally, they rely mainly in simple messages list which cannot capture the structure of a discussion, and analysing how a conclusion was reach becomes a complicated task. Finally, messages can be enriched with semantic annotations enabling the system to build semantic links between and within discussions, which can constitute a knowledge base for future discussions.

In collaborative systems, annotations are artifacts that provide a flexible way to share information pieces. Depending in the format of the annotations (e.g., textual or graphical), they can serve different purposes [7]. We are particularly interested in the role of annotations as boundary objects, which can provide a way to explicitly set the scope of an analysis (or part of it) to a specific geographic area. For example, in the context of species conservation, an icon can be used to pinpoint a location of particular interest for the species reproduction; a line can be used to represent a

[1] A capture or memorize state of analytical environments in a given moment.

Figure 2: STDB offers functionality to analyze features of interest through discussions bounded in space and time.



Source: the authors.

migration route of birds; and a polygon can be used to bound an area for which conservation policies need to be developed. However, this still lacks the crucial information when the phenomenon occurred.

3 Spatio-Temporal Discussion Board

The main assumption behind the STDB is that in data spaces with a large spatial and temporal extent, features of interest such as patterns and outliers may occur in different places and times. Therefore, allowing analysts to define discussion boards bounded in space and time offers several advantages: it focuses the attention of analysts in features of interest, while keeping the features of interest within their original spatio-temporal context, thereby improving organization of analytical discussions within GVA environments (see Figure 2).

Based in the principles of DC, we conceive map-based discussions as a socio-technical system in which artifacts provide analysts with means for information exchange. However, this information exchange can only be effective if it takes place within a clearly defined context. The lack of this context can lead analysts to assume that a discussion extends over the whole data space, and further to the exchange of irrelevant information. Our hypothesis is that having a clear defined context and boundary for discussion will help in focusing analysts' attentions, and elicit sensible information, and ultimately meaningful knowledge. To address this, we design the annotations as spatio-temporal bounding elements, which enable analysts to define explicit spatial and temporal context, and boundaries for map-based discussions. Additionally, analysts can provide background information, which constitutes the thematic context for the discussion. Finally, having explicit boundaries for discussions offers the possibility of better organize them in the analytical environment, and facilitate the search for related and relevant context information.

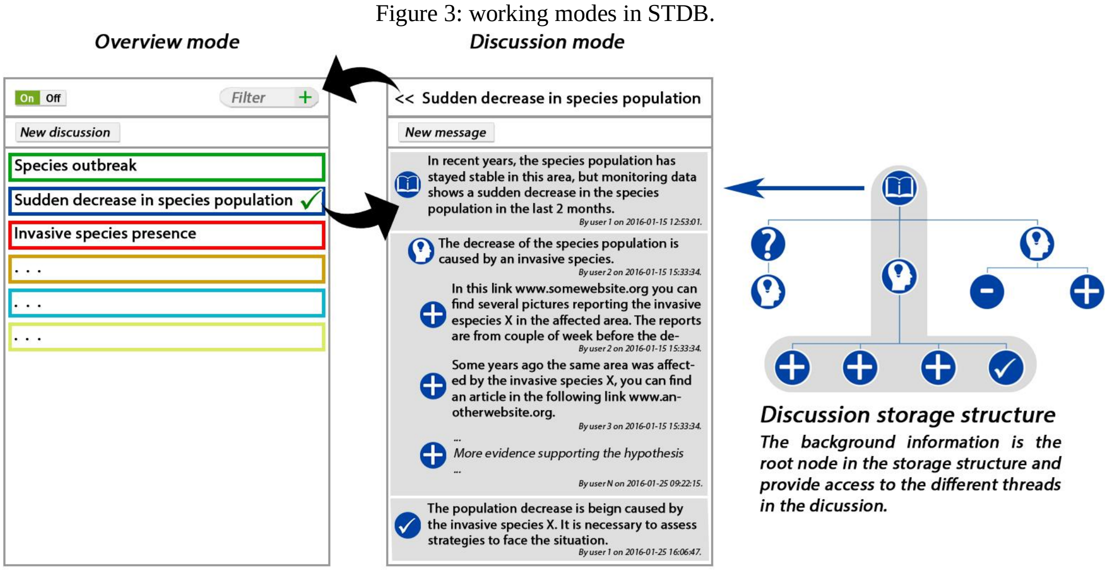
The STDB is a concept that can be implemented as a component in any GVA environment. For this reason, implementations of the method should allow the user to easily activate and deactivate it at any time, for example by including an on/off button (see Figure 2). When deactivated, the discussion spaces are hidden, and therefore the user can't access them. This aims to avoid interrupting the user tasks that don't require the usage of STDB. However, any analysis tasks performed within the spatio-temporal boundaries of a discussion will be added to the discussion and will be accessible later if required. The conceptual framework of the STDB can thus also work in the background.

When activated, the component offers two working modes: overview and discussion. Moving from one to another mode requires a single click in either directions. As shown in Figure 3, the user can access a discussion by clicking in its title, and going back to overview mode by clicking in the "<<" (going back) button. In the following paragraphs, the working modes are described.

The overview mode offers options to search for existing discussions and define new ones (see Figure 2). In this mode, the discussions are shown in a list identified by colors, which are also used to display their spatial and temporal extensions. Additionally, the title and spatial extension can include an icon indicating the status of the discussion (i.e., active, inactive and concluded). The user can search a discussion by entering keywords in the filter box, or use the advanced filtering options which include spatial and temporal bounds, participants and discussion status. The advanced options are accessible through a button located aside the filter box. Advanced filtering will be based in semantic annotation of the discussions and usage of an ontology to build semantic links within and between discussions.

The analyst starts a new discussion simply by clicking on the "New discussion" button (See Figure 3), and providing

Figure 3: working modes in STDB.



Source: the authors.

title, background information, bounding box and time-range. The title and background information serve the purpose of informing the aim of the discussion to other analysts, while the bounding-box and time-range define the spatial and temporal boundaries for the discussion respectively. The bounding-box and time-range can be defined by clicking over the map and time-line.

In the discussion mode, the user can read the messages in the selected discussion and if its status is active, contribute to it. The discussion is developed by exchanging messages in an asynchronous manner. To contribute, the user can click in the “New message” button or over another message to reply to it. Based in the principles of AMT, we decided to decompose discussions in background information, question, hypothesis, evidence and conclusion. The contributions are shown accompanied with semantic cues that represent the type of contribution, and helps in understanding the structure of the discussion space (See Figure 3 and Table 1). Additionally, the contributions can be enriched with links to external sources, annotations over the visual products and attached multimedia files.

STDB uses a tree-based structure as shown in Figure 3 to store the contributions and the links between them. This structure aims to facilitate the display and exploration of the discussion space in a hierarchical way (i.e., tree graph or threaded discussion). Additionally, this structure facilitate to analyze how a conclusion was reached, and keep stored the discarded questions and hypothesis, which can be of importance in future discussions boards.

As mentioned earlier, discussions can be in three states: active, inactive and concluded. A discussion is automatically set to active when it is created and keeps this status until a conclusion is written, or the space is disable by its creator either because it is not relevant anymore or a low level of activity.

Table 1: Contribution types. Based on [11].

Type	Icon	Description
Background information		Relevant known information (thematic context).
Question		Guide the discussion and call for hypothesis generation.
Hypothesis		Propose answers and call for evidence provision.
Evidence (pro)		Evidence to support a hypothesis.
Evidence (con)		Evidence to reject a hypothesis.
Conclusion		Conclusion drawn upon the questions, hypothesis and evidence.

Source: the authors.

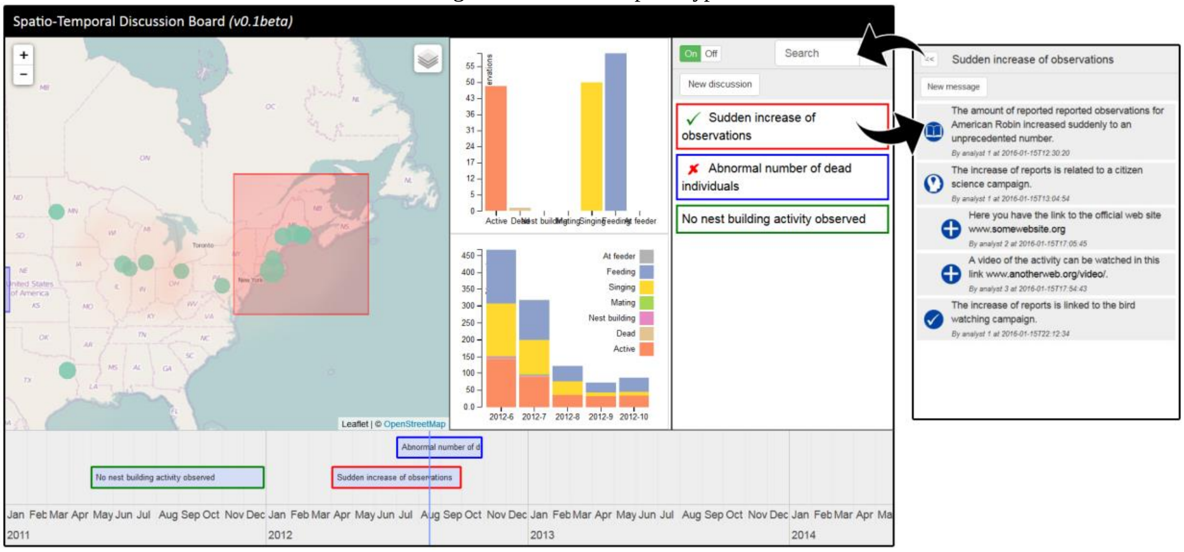
4 Prototype implementation[2]

Following the description provided in the previous section, we developed a simple prototype as proof-of-concept for basic functionality. The development effort aimed to: identify libraries that can be used to implement STDB in a web-based GVA environment; and identify potential interaction issues caused by the integration of STDB in the GVA environment. Figure 4 shows the interface of the developed prototype. The development stages are briefly summarized in the following paragraphs.

The first step in developing the prototype was to build a basic GVA environment that allow us to display and interact with the spatial, temporal and attribute dimensions of the data

[2] Currently, an improved version is under development.

Figure 4: web-based prototype.



Source: the authors.

space[3]. For this aim, we used Leaflet, Vis and D3[4] respectively.

The second step was to implement the overview mode. For this aim, we needed to show a list of discussions identified by colors, this was done with basic JavaScript programming. Additionally, to address the need of displaying the spatial and temporal bounds, functionality provided by the libraries was used. For the spatial bound, Leaflet's Rectangle class was used, and for the temporal bound, Vis offers the option to define element's styling using Cascade Style Sheets (CSS).

As third step, we implemented the discussion board. This stage was restricted to the development of a simple feature to exchange messages, however, given the functionality offered by the selected libraries, we don't foresee any difficulty on developing the full discussion board.

Finally, we tested the prototype with a mock case study using a dataset of American Robin observations [17]. This case study aimed to identify potential design issues, and to draw conclusions on the objectives stated at the beginning of this section.

5 Conclusions and future work

The results reported in this paper show that currently available web-technologies offer a suitable platform to implement the STDB. However, in order to draw empirically grounded conclusions on its functionality or usability, we plan to conduct user experiments.

Since the STDB requires to lay over information on the map and time-line to show the spatial and temporal extensions of the discussion space, the user interaction with the GVA environment is affected.

Following our conclusions, we are developing a full prototype to conduct functional and usability tests. Once completed, the prototype will be tested in a real-world analysis environment. The selected scenario is the monitoring and control of the Olive Fruit Fly (*Bactrocera oleae*) in Málaga, Spain. This case was selected because as a long-term analysis process, it can benefit from the definition of analytical sub-spaces. Additionally, the analysis is already performed as a collaborative effort through face-to-face meetings, which restricts the participation to specific location and time. Therefore, an asynchronous and distributed collaborative method will allow more stakeholders to be engaged.

References

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

[2] G. Andrienko, N. Andrienko, D. Keim, A. MacEachren and S. Wrobel. Challenging problems of geospatial visual analytics. *Journal of Visual Languages & Computing*, 22(4):251-256, 2011.

[3] M.R. Fuller, W.S. Seegar, and L.S. Schueck. Routes and Travel Rates of Migrating Peregrine Falcons Falco peregrinus and Swainson's Hawks Buteo swainsoni in the Western Hemisphere. *Journal of Avian Biology*, 29(4):433-440, 1998.

[4] F. Hardisty. GeoJabber: enabling geo-collaborative visual analysis. *Cartography and Geographic Information Science*, 36(3):267-280, 2009.

[3] Data were provided by the USA National Phenology Network and the many participants who contribute to its *Nature's Notebook* program.

[4] http://leafletjs.com/, http://visjs.org/ and http://d3js.org/

[5] J. Heard, S. Thakur, J. Losego and K. Galluppi. Big Board: Teleconferencing Over Maps for Shared Situational Awareness. *Computer Supported Cooperative Work (CSCW)*, 23(1):51-74, 2014.

[6] J. Heer, F. Viégas, and M. Wattenberg. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Francisco, California, 2007.

[7] S. Hopfer and A.M. MacEachren. Leveraging the potential of geospatial annotations for collaboration: a communication theory perspective. *International Journal of Geographical Information Science*, 21(8):921-934, 2007.

[8] J. Hollan, E. Hutchins and D. Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2):174-196, 2000.

[9] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, M. Kwan-Liu and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization* 10(4):310-326, 2011.

[10] D. Keim, J. Kohlhammer, G. Ellis and F. Mansmann. *Mastering the Information Age: Solving Problems with Visual Analytics*. Florian Mansmann, 2010.

[11] K. Luther, S. Counts, K. B. Stecher, A. Hoff and P. Johns. Pathfinder: an online collaboration environment for citizen scientists. *In proceedings of Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, 2009.

[12] OECD. *OECD data*. Retrieve on 2016-01-04, from https://data.oecd.org/, 2015.

[13] G. Rojas and V. Muñoz. On Principle and Application Progress in Location-Based Services. In C. Liu, editor. *Twitter-Based Geocollaboration: Geovisualization and Geotagging of Microblogging Messages*, pages 181-198. Springer International Publishing, 2014.

[14] O. Scheuer, F. Loll, N. Pinkwart, B. McLaren. Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, 5(1):43-102, 2010.

[15] J. Thomas and K. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2005.

[16] UNISYS. *Hurricane/tropical data*. Retrieved 2016-01-04, from http://weather.unisys.com/hurricane/, 2014.

[17] USA National Phenology Network. *Plant Phenology Data for the United States*, 2011-2015. USA-NPN, Tucson, Arizona, USA. Retrieved 2016-01-08 at http://www.usanpn.org/results/data, 2016.