Proceedings of the 3rd Open Source Geospatial Research & Education Symposium

Edited by Ari Jolma, Pekka Sarkola, and Lassi Lehto

Aalto University

SCIENCE +
TECHNOLOGY

CONFERENCE
PROCEEDINGS

Proceedings of the 3rd Open Source Geospatial Research & Education Symposium OGRS 2014

Ari Jolma, Pekka Sarkola, Lassi Lehto (editors)

441   697
Printed matter

Abstract

The third Open Source Geospatial Research & Education Symposium (OGRS) was held in Helsinki, Finland, on 10 to 13 June 2014. The symposium was hosted and organized by the Department of Civil and Environmental Engineering, Aalto University School of Engineering, in partnership with the OGRS Community, on the Espoo campus of Aalto University. These proceedings contain the 20 papers presented at the symposium. OGRS is a meeting dedicated to exchanging ideas in and results from the development and use of open source geospatial software in both research and education.

The symposium offers several opportunities for discussing, learning, and presenting results, principles, methods and practices while supporting a primary theme: how to carry out research and educate academic students using, contributing to, and launching open source geospatial initiatives. Participating in open source initiatives can potentially boost innovation as a value creating process requiring joint collaborations between academia, foundations, associations, developer communities and industry. Additionally, open source software can improve the efficiency and impact of university education by introducing open and freely usable tools and research results to students, and encouraging them to get involved in projects. This may eventually lead to new community projects and businesses. The symposium contributes to the validation of the open source model in research and education in geoinformatics.

CONTENTS

Preface

The third Open Source Geospatial Research and Education Symposium (OGRS 2014) follows two successful symposiums: the first OGRS in Nantes, France in 2009 and the second OGRS in Yverdon-les-Bains, Switzerland in 2012. We hope that we have been able to meet the high standards set by the two previous symposiums and that the third edition delivers interesting and high-quality contents to participants and to the open source geospatial community at large.

The open source geospatial world moves along and develops in a very fast pace. The FOSS4G conference series is becoming a staple global event with regional conferences taking place regularly. The academic world is becoming more and more interested, as the growth of the open source geospatial laboratory network testifies. Participating in open source initiatives can potentially boost innovation as a value creating process requiring joint collaborations between academia, foundations, associations, developer communities and industry. Additionally, open source software can improve the efficiency and impact of university education by introducing open and freely usable tools and research results to students, and encouraging them to get involved in projects. This may eventually lead to new community projects and businesses.

OGRS 2014 originally received 27 extended abstracts, which were reviewed by at least two independent referees. The authors of the accepted submissions were given time to implement the referees' comments. This book contains 20 papers that have been organized into three tracks: educational use of open source geospatial, natural resources research, and urban research with open source geospatial. The papers on education have perspectives ranging from curriculum development to individual courses and from specific software tools to teaching materials. The papers on natural resources deal with, among others, development of various types of decision support systems and advanced terrain analysis. The urban research papers discuss urban mobility, urban planning, and biodiversity in urbanizing regions. The diversity of the papers emphasizes the multidisciplinary nature of geospatial research.

The symposium would not have been possible without the OGRS community, its steering group, the local organizing committee and especially its unselfish student members, the referees, the keynote speakers, the workshop organizers, and, finally, all the authors. The organizers are also grateful for the financial support received from the Federation of Finnish Learned Societies and the hospitality of the City of Espoo and Aalto University. We want to thank you all.

Ari Jolma, Pekka Sarkola, and Lassi Lehto
Espoo, May 2014

The symposium program

Tuesday 10.6.2014

Parallel workshop: Monteverdi 2.0 - Software for Remote Sensing Image Analysis
Parallel workshop: Road Analysis with H2Network

Lunch

Opening session

Keynote - Prof. Peter Baumann (Jacobs University Bremen)
Spatio-Temporal Big Data - the rasdaman approach

Session *Educational applications*

Wednesday 11.6.2014

Keynote - Prof. Roger Bivand (Norwegian School of Economics)
Why depending on the contributions of others while contributing yourself makes sense

Session *Educational applications* continues

Lunch

Session *Urban analysis and applications*
Parallel workshop: Processing LiDAR data with LAStools
Parallel workshop: Managing and Sharing Sensor Time-Series Data with istSOS

Thursday 12.6.2014

Keynote - Prof. Matti Rossi (Aalto University School of Economics)
Open data services

Session *Natural resources analysis and management*

Lunch

Session *Natural resources analysis and management* continues
Parallel workshop: Big Data Analytics with rasdaman, the raster data manager
Parallel workshop: Applied Spatial Data Analysis with R

Friday 13.6.2014

Keynote - Prof. Matt Duckham (University of Melbourne)
Open access publishing in GIScience

Session *Data processing and standards*

Closing Session

Committees

The OGRS Community Steering Group

- **Erwan Bocher**, *Institute for Research on Urban Sciences and Techniques, France*
- **Olivier Ertz**, *HEIG-VD, University of Applied Sciences Vaud, Switzerland*
- **Ari Jolma**, *Aalto University, Finland*
- **Stéphane Joost**, *EPFL, Geographic Information Systems Laboratory, Switzerland*
- **Ivan Marchesini**, *IRPI, Italy*
- **Sergio J. Rey**, *Arizona State University, USA*

The organizing committee

- **Anas Altartouri**, *Aalto University, Finland*
- **Erwan Bocher**, *Institute for Research on Urban Sciences and Techniques, France*
- **Olivier Ertz**, *HEIG-VD, University of Applied Sciences, Vaud, Switzerland*
- **Ari Jolma**, *Aalto University, Finland*
- **Stéphane Joost**, *EPFL, Switzerland*
- **Jiri Kadlec**, *Aalto University, Finland*
- **Lassi Lehto**, *Finnish Geodetic Institute, Finland*
- **Pekka Sarkola**, *Gispo Oy, Finland*
- **Hanna Vellonen**, *Aalto University, Finland*
- **Aura Salmivaara**, *Aalto University, Finland*
- **Kari Salovaara**, *Finland*

The referees

- **Paula Ahonen-Rainio**, *Aalto University, Finland*
- **Erwan Bocher**, *Institute for Research on Urban Sciences and Techniques, France*
- **Flora Branger**, *Irstea, France*
- **Maria Brovelli**, *Politecnico di Milano, Italy*
- **Massimiliano Cannata**, *SUPSI, Switzerland*
- **Eduardo Corbelle Rico**, *University of Santiago de Compostela, Spain*
- **Arie Croitoru**, *University of Alberta, Canada*
- **Pierluigi De Rosa**, *University of Perugia, Italy*
- **Olivier Ertz**, *HEIG-VD, University of Applied Sciences, Vaud, Switzerland*
- **Gilles Gesquières**, *University of Provence, France*
- **Carlos Granell**, *European Commission Joint Research Centre, Italy*
- **Manuel Grizonnet**, *CNES, France*
- **Eric Grosso**, *IGN, France*
- **Frédéric Hubert**, *Laval University, Canada*
- **Ari Jolma**, *Aalto University, Finland*
- **Jiri Kadlec**, *Aalto University, Finland*
- **Christian Kaiser**, *University of Zurich, Switzerland*
- **Tomi Kauppinen**, *Aalto University, Finland*
- **Stefan Keller**, *University of Applied Sciences of Eastern Switzerland, Switzerland*
- **Mikko Kolehmainen**, *University of Eastern Finland, Finland*

- **Ilkka Korpela**, *University of Helsinki, Finland*
- **Barend Köbben**, *University of Twente, Netherlands*
- **Ari Lauren**, *The Finnish Forest Research Institute, Finland*
- **Nathalie Long**, *La Rochelle University, France*
- **Ivan Marchesini**, *Research Institute for Geo-hydrological Protection (IRPI), Italy*
- **Helena Mitasova**, *North Carolina State University, USA*
- **Judicaël Picaut**, *IFSTTAR Institute, France*
- **Christine Plumejeaud-Perreau**, *CNRS, France*
- **Stefan Steiniger**, *University of Calgary, Canada*
- **Martin Vermeer**, *Aalto University, Finland*
- **Li Xiang**, *East China Normal University, China*
- **F.Javier Zarazaga-Soria**, *University of Zaragoza, Spain*

The scientific committee

- **Paula Ahonen-Rainio**, *Aalto University, Finland*
- **Dan Ames**, *Brigham Young University, USA*
- **Suchith Anand**, *University of Nottingham, UK*
- **Flora Branger**, *Irstea, France*
- **Maria Brovelli**, *Politecnico di Milano, Italy*
- **Christophe Claramunt**, *Naval Academy Research Institute, France*
- **Eduardo Corbelle Rico**, *University of Santiago de Compostela, Spain*
- **Arie Croitoru**, *University of Alberta, Canada*
- **Pierluigi De Rosa**, *University of Perugia, Italy*
- **Jérôme Gensel**, *University Pierre Mendès, France*
- **Gilles Gesquières**, *University of Provence, France*
- **Carlos Granell**, *European Commission Joint Research Centre, Italy*
- **Manuel Grizonnet**, *CNES, France*
- **Eric Grosso**, *IGN, France*
- **Frédéric Hubert**, *Laval University, Canada*
- **Christian Kaiser**, *University of Zurich, Switzerland*
- **Tomi Kauppinen**, *Aalto University, Finland*
- **Stefan Keller**, *University of Applied Sciences of Eastern Switzerland, Switzerland*
- **Mikko Kolehmainen**, *University of Eastern Finland, Finland*
- **Ilkka Korpela**, *University of Helsinki, Finland*
- **Barend Köbben**, *University of Twente, Netherlands*
- **Ari Lauren**, *The Finnish Forest Research Institute, Finland*
- **Nathalie Long**, *La Rochelle University, France*
- **Ivan Marchesini**, *IRPI, Italy*
- **Helena Mitasova**, *North Carolina State University, USA*
- **Judicaël Picaut**, *IFSTTAR Institute, France*
- **Sergio Rey**, *Arizona State University, USA*
- **Tapani Sarjakoski**, *Finnish Geodetic Institute, Finland*
- **Charles Schweik**, *University of Massachusetts Amherst, USA*
- **Hans-Jörg Stark**, *University of Applied Sciences and Arts Northwestern Switzerland, Switzerland*
- **Stefan Steiniger**, *University of Calgary, Canada*

- **Ann van Griensven**, *UNESCO-IHE, Netherlands*
- **Martin Vermeer**, *Aalto University, Finland*
- **Kirsi Virrantaus**, *Aalto University, Finland*
- **Elizabeth A. Wentz**, *Arizona State University, USA*
- **Li Xiang**, *East China Normal University, China*
- **F.Javier Zarazaga-Soria**, *University of Zaragoza, Spain*

Teaching spatio-temporal analysis and efficient data processing in open source environment

Giuseppe Amatulli[1], Stefano Casalegno[2], Remi D'Annunzio[3], Reija Haapanen[4], Pieter Kempeneers[5], Erik Lindquist[3], Anssi Pekkarinen[3], Adam M. Wilson[1], and Raul Zurita-Milla[6]

[1]Department of Ecology and Evolutionary Biology, Yale University, New Haven, USA
[2]Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, UK
[3]FAO Forestry Department, Rome, Italy
[4]Haapanen Forest Consulting, Kärjenkoskentie 38, 64810 Vanhakylä, Finland
[5]Flemish Institute for Technological Research (VITO), Mol, Belgium
[6]Faculty of Geo-Information Science and Earth Observation, University of Twente, The Netherlands

Abstract: Geo-data are getting larger and complex; consequently, their analysis requires new capabilities both in computer hardware and programming skills. We describe the effectiveness of teaching several open source programming languages to analyze spatio-temporal using Linux OS installed in a Virtual Machine (VM) and/or as LiveUSB. Courses organized in developed and developing countries haven shown positive results including improving students' motivation and engagement, encouraging experimentation and collaboration, and utilizing the allocated time for classes more effectively. Beside the tremendous effort of the developers to maintain and improve the functionality open source geo-spatial software more energy should be invested in organize courses to enlarge the use of this tools.

Keywords: Open source programming, teaching programming, big data

1 Introduction

In recent years there has been an explosion of geo-datasets derived from an increasing number of remote sensors, field instruments, sensor networks, and other GPS-equipped "smart" devices. Geo-datasets are available at increasingly fine

spatio-temporal grains, which facilitate addressing old and new research questions across larger domains with increased resolution and unprecedented rigor. However, many of these datasets, especially those labeled as "big data", are complicated and their analysis requires data processing and analysis skills beyond those taught in basic courses (Donovan 2008). Furthermore, some data sets can only be efficiently analyzed using cluster and/or cloud computing facilities, such as Google Earth Engine, or NASA NEX. Access to these increasingly powerful systems generally requires knowledge of at least one programming language.

Indeed, as many of the potential users of these new data sets have limited access to often expensive proprietary software as well as to high speed internet and efficient hardware, they use desktops and laptops with limited processing power for their daily work. Free and open-source software under the Linux Operating System (OS) can address this gap, offering a powerful alternative to proprietary software which can be used in desktop PCs, laptops, low-cost computing hardware such as the RaspberryPi, and is also the standard OS on most scientific computer clusters. Thus proficiency in free and open-source software enables increased hardware efficiency, access to freely accessible code, and freedom from licensing constraints. These have the combined effect of increasing research transparency and reproducibility and advancing the ability of the geo-data community to analyze complex data sets.

Open source software also enables broad and relatively inexpensive access to scientific workflows, which is especially important in developing countries (Camara & Fonseca 2007). The scientific community in many areas is limited by technical skills, modern computer hardware, and software availability. Proprietary software are typically expensive and may not work in old hardware. Internet connections can be slow and inconsistent and frequent power outages may occur. In this situation, carrying out the computations in a cloud environment is a powerful option. Also, sending a script (text file), which performs the desired image processing tasks is lighter and less prone to cuts in electricity and internet than doing the processing in-house (Roy et al. 2010). For example, cloud-processing would solve the problem of downloading and processing Landsat from countries with low-bandwidth connections (Roy et al. 2010).

This paper presents our experiences in teaching spatio-temporal analysis and big data processing of geo-spatial data using multiple open source programming languages integrated in a Linux OS installed in a Virtual Machine (VM) and/or as LiveUSB. We describe our experience in teaching in academic institutions, research centres, and national or international organizations in developed and developing countries. We describe the effectiveness of teaching several programming languages, including course expectations and outcomes, in courses ranging from three hour workshops to one/two week intensive training to full semester-long traditional academic courses.

Section 2 illustrates the importance and the difficulty of learning programming skills to develop geo-spatial applications. Section 3 describes the design,
www.ogrs-community.org

organization, teaching approach and the technical aspects required to organize courses in open source geo-spatial programming. Section 4 reports on feedback from surveys collected during the courses and pros/cons of the implemented teaching methods with special attention paid to each participant's learning experience and how they relate to a priori expectations.

2 The importance and difficulty of programming in environmental sciences

Mastering skills in analyzing spatio-temporal data is fundamental for most environmental and engineering studies. This is why academic institutions have incorporated or increased basic GIS and remote sensing courses in their traditional and online curricula (Govorov & Gienko 2013). Nevertheless, these courses are mostly based on proprietary and/or graphical user interface-based (GUI) software. Graduate students in environmental sciences are rarely trained in the programming skills required to design and implement (complex) geo-spatial workflows. There are few example of academic institutions with courses in programming applied to environmental analysis which address only one programming language (e.g. Python and/or R) at introductory level. In the absence of targeted courses, learning these skills often requires several years of independent learning and strong motivation where personal experiences are accumulated and assimilated.

Learning programming languages enhances students' practical problem-solving skills, potentially inspires new research questions and ideas stimulating critical, analytical and high-level thinking and promotes creative synthesis to get a first draft of task (Britto & Sá-Soares 2014). Programming allows the automation of repetitive procedures in a research workflow, but can also be used to test new hypotheses "on the fly". For example, experienced developers frequently write new functions and algorithms years before they are assimilated into graphical user interfaces by major software companies or become operational additions to open source software. In this way, learning programming skills enables researchers to use and share custom and state-of-the-art analytical algorithms. Thus, programming enables an interactive approach to learning, where algorithms and new ideas can be immediately explored and tested by changing the code. As students advance, they often discover more sophisticated methods of analysis and workflow organization, iteratively upgrading their script routines with more complex geo-functions and improving processing speed and use of available computing resources. Thus, learning a programming language is much more than learning how to write code. In some ways, it teaches students how to think scientifically and develop transparent scientific workflows.

The inherent difficulty in learning programming is not new in the literaure. Studies have discussed several problems in teaching programming (Cook 2008), the inherent difficulty of programming (Teague & Roe 2008) and controlling what kind

of mistakes students are more likely to make (Spohrer & Soloway 1986). Generally speaking, a programmer needs to instantiate abstract programming concepts and techniques to solve particular computational problems (Britto & Sá-Soares 2014). This is considered one of the most difficult aspects of programming. Applying abstract conceptualization for solving environmental and engineering issues/studies demands active experimentation with writing the code and reflective observations of the obtained results, keeping in mind the theory beyond the environmental problem (Giraffa et al. 2014). In other words you have to be able to write the code and also assess the results and the general data trend based on your expert knowledge of the subjects.

Many graduate students, especially those in environmental studies, do not have any programming experience or knowledge in how to deal with "big data" prior to starting a PhD research program or an environmental career requiring knowledge in programming (Tan et al. 2009, Donovan 2008). Many have the impression that programming is difficult to learn which may unconsciously prevent students from getting started (Tan et al. 2009). Besides the psychological dimensions, there are technical aspects which further discourage student's ability to start learning. Indeed, even if the online documentation is exhaustive in terms of quality and quantity, the users may encounter problems in learning an open-source language, including: 1) shifting to a new OS (e.g. Linux) requires time and effort, 2) learning new patterns and commands required to perform previously normative work is difficult, 3) installing necessary software is not always straightforward, 4) software updating, likewise, is not always simple.

3   Methods: developing open source geo-spatial programming course

To bridge the gaps in training and skills that are critical to perform cutting edge analyses of complex geo-data, we began organizing intensive training courses for graduate and postgraduate researchers and technicians. The courses target those who already possess basic knowledge in geospatial sciences therefore they focus on applied processing with no introductory GIS/RS lessons.

3.1   Main aspects of the course

Our primary tool in open source geo-spatial programming courses is a standardized Linux OS environment with pre-installed software packages and readily available tutorials and exercises. In order to facilitate organizing training courses and to expedite the learning process all the documentation and exercises are freely accessible via the internet (through wikis at www.spatial-ecology.net and www.openforis.org/wiki). The tutorials and exercises are contributed by researchers and experts who use open source tools in their daily work. The approach to teach data analysis is unique, integrating multiple programming languages such as

www.ogrs-community.org

AWK, BASH, Python, R, GRASS, and other software to build workflows. Our initiative complements existing communities, such as OSGeo by providing specific tutorials and tools that help users to analyze their data and summarize results. The courses focus on automated processing and teach basic programming concepts for using command-line utilities to process large data sets. With simple scripts, we show how to automate necessary tasks and adapt existing programs for specific problems to gain higher efficiency and improved performance. From an outside perspective, the course program may be perceived as covering only technical aspects (e.g. computer programming or code writing). Nevertheless, scientific aspects are integrated as soon as the participants start to use the tools to solve specific tasks related to environmental issues.

Most scripts we use are based on powerful geo-tools that act as modular building blocks for the task at hand. Rather than focus in one programming language or software package, we give an introduction to several tools and languages and integrate them according to their features and strengths. Each language/tool is explained using a common structure: 1) language syntax including specification of various flags and options, 2) accessing available documentation, 3) identifying and explaining the typical problems in data processing and the procedures to solve them (debugging), 4) structuring a script to connect various tools and/or languages, 5) working with the output 6) assessing the results and evaluating them based on a full knowledge of the problem. Furthermore, cloud computing and remote server access are considered major milestones in processing large amounts of data, while multi-core computing allows several processes to run simultaneously. Near the end of the course, and only in case of outstanding course participants, we also give an introduction for using the tools in a cluster computing environment by transforming a simple "for loop" into a "multicore for loop" to allow simultaneous processing of massive data processes. We address this topic using specific R libraries (*foreach, parallel*) and in BASH (*xargs, qsub*).

Providing to each course participant a version of "ready-to-usue" Linux OS including all necessary software and course materials ensures: 1) minimum software installation effort 2) standardized working environment 3) easy portability to other computers, and 4) a gentle transition from the host OS to guest OS which does not intimidate beginning users. Furthermore, participants are able to create their own workspace for data and scripts within the Linux OS VM/LiveUSB, with the ability to precisely reproduce the computing environment for completing class experiments at home or on any other computer. From the trainers' point of view, using a common cloned Linux OS on each participant's computer facilitates group problem-solving, debugging, and testing commands during the training. Furthermore, solutions tested on one machine can be easily transferred to other students. Specifically, for training we use an ad-hoc customization of the Ubuntu distribution with free and open source remote sensing, GIS, and statistics software pre-installed, including sample geo-data, scripts, and exercises directly linked to the web content.

Selection of the teaching environment depends mainly on the available hardware resources. If efficient PCs with more than 4GB of memory are available, VMs solutions can be efficiently applied. VMs are commonly used by developers and advanced programmers (Nieh & Vaill 2006) but not commonly used for teaching purposes. With older hardware and less memory, a Live system, in which the entire OS and required software is contained either in a USB or CD format, is a more robust solution as it can utilize all the hardware resources of the host computer yet does not use valuable and likely scarce hard-disk space for additional installations. However, the VM approach may be recommendable for users with sufficient hardware as it allows them to use two systems simultaneously. This may facilitate the transition from proprietary software to open source. The Linux OS VM/LiveUSB, in combination with the resources on the wiki web application, can also aid self-taught programmers. Indeed, users can download the Linux OS from the web application and use the scripts, data and web tutorials for self-learning even if they do not have access to an organized course. The materials, data, and software are all published under a Common Public License agreement, so that they can be redistributed or installed on other computers.

In the longer courses, we also mentor participants in personal projects typically related to their PhD/Master thesis. We select the subset of programming tools, idioms, and procedures that would best serve the participants' needs and suggest proper resources (e.g., sample code, blog posts, online tutorials, etc). We request the participants to build up a procedure for their own case study using the tools learned in class. This encourages the participants to focus on the coding and workflow aspects of their projects, having already assimilated the scientific aspect. Moreover, working on their own project increases the motivation in learning something that can be extremely relevant for their career and personal development.

3.2   Software tools and support material

In this section we describe several of the software tools incorporated into the course. In addition to commonly used geo-spatial software such as GRASS (reference) and QGIS (reference), we also incorporate several command-line utilities (such as the Open Foris Geospatial Toolkit and pktools) and the R statistical language which has a rapidly growing library of geo-spatial functionality. We recognize that mastering a programming language often represents a long-term learning process. The use of a variety of new tools help students to understand that the same task can be carried out in many ways, often using a different combination of tools. Exposing the participant to a number of new tools helps them also to understand that becoming an expert user does not require committing everything to memory. In fact, learning how to look for information as required is much more important than mastering a single software package. However, using the tools developed by some of the authors of this paper has allowed us to have the best

www.ogrs-community.org

possible response time as new functionality can be added on-the-fly during the course.

The tools we have chosen to include in our Linux OS distributions are mean operative processing of geo-spatial data sets, primarily for remote sensing, GIS application and spatial statistical analysis. Furthermore, many geo-tools are implemented primarily in a compiled programming language (such as C or C++) with high performance (e.g. processing speed and memory management) while providing a relatively simple user interface via the command line. In fact, these and similar tools have been used, for example, to process first high resolution forest maps for Europe (Kempeneers et al. 2011, Pekkarinen et al. 2009). In addition, some of them are currently in use in more than 20 countries by various institutions and individuals and are being used by a number of developing countries to map their natural resources. During the courses, the tools are used to address participants' problems in such a way that they can see how learning new skills benefits their performance and productivity. We also emphasize that they are free to use the tools for their own purposes including private consulting. In particular, we teach the use of Open Foris Geospatial Toolkit, pktools and R/BUGS for computational modeling. Here below we describe their main features.

Open Foris Geospatial Toolkit is part of the FAO's Open Foris initiative aimed at helping developing countries assess and monitor their forest resources, especially forest cover change. The tools are designed for efficient and automated interpretation of remote sensing data, and can be used for common image processing tasks such as radiometric normalization, image algebra, and image segmentation and classification. It contains modules also for change detection, extraction of image statistics, image filtering, feature extraction and gap-filling. The toolkit also allows interaction with R, which provides access to more sophisticated classification algorithms such as Random Forest. The tools are divided into stand-alone C-programs and scripts. Many of the stand-alone programs and scripts use GDAL/OGR libraries and command line utilities. Open Foris Geospatial Toolkit wiki has been visited more than 8000 times (Mar 2014) and the toolkit has users in more than 20 countries.

Pktools is another open source software toolbox that is implemented in a similar way as the GDAL/OGR utilities. This means all programs are executed from the command line using various options. Though some overlap exists between their functionalities, pktools is more oriented towards remote sensing and image processing applications than GDAL/OGR. Typical tasks include: image (morphological and other) filtering, compositing multiple images according to different rule sets. Typical machine learning algorithms are also implemented for image classification or data regression, including feature selection algorithms, artificial neural networks and support vector machines. Like GDAL/OGR utilities, maintaining a simple command line interface allows for a modular approach that can easily be integrated in other scripting languages.

We also introduce spatial statistics using with R and BUGS. We have developed several modules on spatial statistical modeling to enable inference of patterns and processes from the data. The growing size and complexity of geo-spatial datasets has facilitated increasingly sophisticated models of environmental processes. In order to evaluate and critically apply these models, it is vital that complete workflows are documented and repeatable. Our approach uses geo-spatial software mentioned above for data pre-processing and organization followed by R and various BUGS languages to fit statistical models and produce summary output. Examples of this type of approach are available in a public GitHub repository.

Beside these main geo-tools we also mention the potential use of other open source software such as CDO, Orfeo Toolbox, ImageMagick, and gnuplot pointing out their features and their aspects for solving specific tasks. These widen the view of the open source capability without adding new languages and commands. By scripting each stage of the analysis, from the 'raw' data to graphical and tabular summaries, we teach how to develop a transparent analytical workflow that is also easy to share and repeat. In this way, we provide students with the tools they need to understand, run, and share their complete analysis.

4   Results and discussion: course evaluation

Since 2008, more than 400 students and technicians have participated in trainings organized in by the authors of this paper. These training/courses took place in Italy, Spain, Denmark, Peru, Brazil, Paraguay, Mexico, Ecuador, Uganda, The Republic of South Sudan, UK, the Netherlands, Tanzania and the USA, in academic institutions, research centres, and in national and international organizations. The background of the course participants included some exposure to GIS, remote sensing and/or database software but the Linux OS was new to most of them.

Though the initial learning curve to the approach followed in our courses is somewhat steeper than that of proprietary software packages with graphical user interfaces, we are convinced this effort is well spent. The purpose of the courses is to guide geo-data users through basic programming tasks so that they are quickly able to perform similar tasks independently. At the end of each course we asked the participants to fill in a questionnaire so that we could better understand the impact of the course. The questionnaire consisted of four main parts: 1) general information about the student's motivation, 2) the capacity of the students to independently continue to develop skills in this area, and 3) the utility of using new tools (OFGT, pktools) for their work, and 4) detailed feedback about course organization and effectiveness. We also collected valuable information from personal comments. Participants had the option to remain anonymous and to skip questions so the number of respondents varies for each question.

The primary motivation for completing the course was to acquire new expertise in geo-spatial data analysis (88%), followed by the desire to improve pre-existent programming knowledge (31%). The results show that 87% of participants are
www.ogrs-community.org

interested in continuing to use open source software after the course. In particular, 83% found the tools related to their professional work; 14% have little professional interest and 3% no professional interest (see Table 1). About 45% of the course participants described Linux OS as indispensable and the remaining participants see its value for daily work. Concerning the self learning attitude, an average of 85% of the participants felt able to independently improve and learn more about open source tools (see Table 2). In particular, the command line tools most assimilated was BASH (82%) probably due to its extensive usage during the exercises including those based on GDAL and GRASS. About 72% of the participants felt they were able to continue to independently improve their knowledge following the course. In the same line, 86% reported they would likely keep using the open-source command line. It is also notable that those willing to continue using QGis is a bit less (75%) compared to the command-line option (see Table 2). This demonstrates that the course was able to transmit the importance and power of the command-line versus graphical user interface to software. With questions in Table 3 we evaluate the use of OFGT and pktools to perform remote sensing analysis. For both of the new tools we report a positive attitude in their use for different image analysis tasks. Finally, while considering the quality of the pedagogic efficiency, 78% participants felt it was well adapted to their needs, skills, and knowledge; 18% believed it was too difficult and 4% found the training boring and/or too basic (see Table 4).

Also, during trainings in developing countries, the government officials and technicians received the training well. There, we focused on understanding the users' operative needs and expectations. In the recent trainings organized in Peru and Uganda, 100% of the participants replied that they are "likely to use" the presented tools in their future work and that they would recommend the course for their colleagues. The feeling after the training is that more effort have to be done to explain GRASS, R, and BASH as compared to AWK, Gdal or to insist on the theoretical aspect of the scientific questions addressed. The supervision effort and the trainers preparation were positively rated by more than 97% of participants. At the same time, the most commonly listed drawback was the lack of time and therefore the participants wished to receive additional training on the subject.

To address this problem, a relatively innovative course design was tested in the Netherlands between December 2013 and January 2014. There we offered a week of intensive face-to-face training (lectures and practicals), followed by a month of self-study where course participants could assimilate the course contents and apply them to datasets/problems related to their own research. This put the participants in control of their own learning and allowed them to focus on tools and techniques directly related to their work. Finally, a second intensive face-to-face training was organized to present specialized tools (e.g. pktools). During this week, we also scheduled time for each participant to present the work done during the self-study period and after that both participants and course teachers could provide formative feedback to consolidate knowledge and enhance the learning process.

<div align="right">OGRS 2014</div>

Our long term follow-up with some of the students reveal that the impact of the training outside of the academia depends on the institutional commitment. Student working in environments that are unfriendly to open-source software are less likely to adapt the tools and use them to solve their every-day problems because of the lack of software and hardware support. In other environments, it seems to be relatively easy to switch from proprietary software to open-source tools. In cases where that has happened, the use of open source tools has drastically increased among the colleagues of these students as soon as they demonstrated their efficiency.

Table 1: Student's expectations and interest in learning open source geo-spatial programming

Why did you attend a course? (42 participants replied to this questionnaire)			
Acquire new expertise 88%	Improve my skills 31%	Discuss related issues 9%	No similar course found 9%
Do you see any interest in using open source tools in your current and future job? (92 participants replay to this question))			
Yes - 83%	Little - 14%	No - 3%	
How do you rate the use of Linux OS in your daily work? (11 participants replied to this questionnaire)			
Indispensable 45%	Valuable 55%	Of general interest 0%	Not relevant to my activity 0%

Table 2: Ability of participants to independently learn and progress to use open source tools

Do you feel to be able to independently improve and learn more about open source tools?						
Tools	BASH	AWK	GDAL	GRASS	Qgis	R
Answers	73	43	69	71	69	69
Yes (%)	88	79	77	96	97	72
No (%)	12	21	23	4	3	28
Will you use and progress using open source tools in future?						
Answers	74	41	74	77	87	76
Yes (%)	94	92	89	97	75	86
No (%)	6	8	11	3	25	14

Table 3: Overall tools (OFGT, pktools) evaluation considering their use and their implementation in remote sensing analysis

Were your expectation met with regards to the following image processing lessons? (11 participants replied to this questionnaire)					
Image analysis using OFGT	Met entirely (Likely to Use)	Met to some degree (Likely to Use)	Met to some degree (Not likely to use)	Not met at all (Likely to use)	Not met at all (Not likely to use)
General imagery processing (clipping / stacking)	45%	55%	0%	0%	0%
Segmentation	45%	50%	0%	0%	0%
Supervised classification	45%	55%	0%	0%	0%
Verification of the automatic procedure	20%	70%	10%	0%	0%
Statistical analyses	29%	57%	5%	5%	5%

Are you going to use pktools in the future? (14 participants replied to this questionnaire)			
Yes 90%		No 10%	
Have you improved in working with pktools? (16 participants replay to this question)			
Yes 100%		No 0%	
Which part of the pktools would you have liked to see in more in detail? (16 participants replied to this questionnaire)			
Image filtering 50%	Image compositing, merging, subsetting 67%	Image extraction (vector to raster overlay) 73%	Image classification 53%
Do you see potential in the use of pktools? (18 participants replied to this questionnaire)			
Yes 89%	No 11%		
Is the help info understandable? (18 participants replied to this questionnaire)			
Yes but need to be improved 94%	Not understandable 6%		

Table 4: Overall courses evaluation considering the structural format and the trainers performances

Was the course adapt to your skills and knowledge ? (120 participants replied to this questionnaire)			
Well adapted 78%	Advanced and too difficult 18%	Basic and boring 4%	
Which is my prefered learning format ? (54 participants replay to this question)			
Exercise 40%	Theory 23%	Tutorials 23%	Projects 15%
The trainer were clear and prepared? (84 participants replay to this question)			
Yes, very well 60%	Yes, enough prepared 39%	Not enough 1%	
The trainer's supervision during some coding exercises was satisfactory? (82 participants replay to this question)			
Very much 54%	Yes satisfactory 44%	Not enough 2%	

5   Conclusion

This paper presented some experiences providing geo-spatial data processing educational courses to students, professionals, and technicians from a number of different countries using open source image processing tools, a standardized Linux OS environment with software pre-installed, and additional learning materials and tutorials available via web applications. Results confirmed several benefits of this approach including improving students' motivation and engagement, encouraging experimentation and collaboration, and utilizing the allocated time for classes more effectively. Students indicated that they are interested in continuing to use the open source tools. In trainings with a duration of one week the most commonly reported problem in the organization of the course was the lack of sufficient time to fully grasp the concepts of the new computing environment. Many students expressed the need for several days dedicated solely to becoming comfortable with the command-line nature of a completely new OS and interface.

Regarding day-to-day usage of open source tools and software, there is a tremendous effort from the developers to maintain and improve the functionality of the various systems, but our experience shows that usage on a much broader scale could be enhanced through increased formal course offerings as part of all geospatial science curricula. Today, most students in developed countries have enough computing power in their home computers to easily reproduce class experiments and exercises. However, in the developing countries the situation is often challenging due to general lack of computing and internet infrastructure which necessitates that class lessons are processed on old, typically low-performance hardware. Therefore, the selection of an appropriate approach (VM/Live OS) depending on the available computing resources, is a key for satisfactory user experience and a successful course.

Finally, we are confident that with the continuing advances in technology, increased access to internet and the overall improvement in available computing

www.ogrs-community.org

resources, interest in open-source tools and computer programming will subsequently increase and, as such, will be a very important tool in the learning about and processing geospatial data.

References

Brito, M. A. & de Sá-Soares, F. (2014), 'Assessment frequency in introductory computer programming disciplines', *Computers in Human Behavior* 30 (2014), 623–628.

Camara, G. & Fonseca, F. (2007), Information Policies and Open Source Software in Developing Countries. *Journal of the American Society for Information Science and Technology*, 58(1), 121–132.

Cook, W.R. (2008), 'High-level problems in teaching undergraduate programming languages', SIGPLAN Notices, 43, 55–58.

Donovan, S. (2008), Big data: teaching must evolve to keep up with advances. Nature, 455(7212), 461-461.

Giraffa, L.M., Moraes, M.C. & Uden, L. (2014), Teaching Object-Oriented Programming in First-Year Undergraduate Courses Supported By Virtual Classrooms. In The 2nd International Workshop on Learning Technology for Education in Cloud (pp. 15-26). Springer Netherlands.

Govorov, M. & Gienko, G. (2013), 'GIS learning objects: an approach to content aggregation', *International Research in Geographical and Environmental Education*. Vol. 22(2):155-171.

Nieh, J. & Vaill, C. (2005), Experiences teaching operating systems using virtual platforms and linux. In ACM SIGCSE Bulletin (Vol. 37, No. 1, pp. 520-524). ACM.

Kempeneers, P., Sedano, F., Seebach, L., Strobl, P. & San-Miguel-Ayanz, J. (2011), 'Data fusion of different spatial resolution remote sensing images applied to forest type mapping', *IEEE Transactions on Geoscience and Remote Sensing*, 49(12): 4977-4986.

Pekkarinen, A., Reithmaier, L. & Strobl, P. (2009), 'Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data', *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(2):171-183.

Roy, D.P., Ju, J., Mbow, C., Frost, P. & Loveland, T. (2010), Accessing free Landsat data via the Internet: Africa's challenge. *Remote Sensing Letters*, *1*(2), 111-117.

Spohrer, J. C., & Soloway, E. (1986). Novice mistakes: Are the folk wisdoms correct? Communications of the ACM, 29, 624–632.

Tan, P.H., Ting, C.Y. & Ling, S.W. (2009), Learning difficulties in programming courses: undergraduates' perspective and perception. In Computer Technology and Development, 2009. ICCTD'09. International Conference on (Vol. 1, pp. 42-46). IEEE.

Teague, D. & Roe, P. (2008), Collaborative learning: Towards a solution for novice programmers. Proceedings of the tenth conference on Australasian computing education (Vol. 78). Wollongong, NSW, Australia: Australian Computer Society, Inc.

Open source software for tourism promotion: Lake Poli School (LaPS) education project

Michela Arnaboldi[1], Maria Antonia Brovelli[2], and Marco Minghini[2]

[1]Politecnico di Milano, Dipartimento di Ingegneria Gestionale, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
[2]Politecnico di Milano, Dipartimento di Ingegneria Civile e Ambientale (DICA), Laboratorio di Geomatica, Como Campus, Via Valleggio 11, 22100 Como, Italy

Abstract: Management of geospatial data has become crucial in a broad range of new and non traditional disciplines. Here we investigate the application of open source geospatial software in the contexts of tourism and cultural heritage, and education. In our case master level students of environmental science, management science, and computer engineering work in groups to develop real projects proposed by real organizations with the assistance of junior and senior academics. The case study presented here tackled the issue of promoting tourism (particularly slow tourism) in the wonderful nature region located at the border between Italy and Switzerland. The project was carried out in collaboration with stakeholders, and it included a comprehensive managerial analysis and development of practical tools. After data collection, two web-based systems, one for desktop and the other for mobile devices, were developed using open source software allowing browsing and querying a rich set of tourism information. Both platforms were successfully tested with stakeholders. The study demonstrates the applicability of open source software not only for tourism promotion but also for educational purposes.

Keywords: Cultural heritage, education, mobile, open source, Web mapping

1 Introduction

1.1 Lake Poli School (LaPS)

A frequent weakness point of current research as well as business projects is the tendency to separate, instead of integrating, the different skills and competencies required for an overall understanding of the issues addressed. Reasoning in a "working packages" way often makes managers lose sight of the problem complexity, and accordingly find out inadequate or only partially-satisfactory solutions. Common examples occur in the field of cultural heritage and more generally in the public sector, where "perfect" applications and software are developed without properly designing/implementing the managerial infrastructure needed to operate and maintain them.

Against this background, an interdisciplinary educational project is carried out each year at Politecnico di Milano (Como Campus) which is called Lake Poli School (LaPS). The name recalls the city (Como is famous for its "Lake") and the university ("Poli" is an abbreviation for Politecnico di Milano) where the initiative is performed; while "School" highlights its academic nature. The idea at the origin of LaPS was linked to the aforementioned difficulty that both practitioners and academics have in overpassing the boundaries of their disciplines and languages. This limits the potentiality to innovate, creates tensions when cooperation is required among different groups, and can undermine long-term results. Starting from this seminal idea, LaPS is a multi-disciplinary initiative where MSc students from Management, Environmental, and Science Computing Engineering work in group to develop real projects proposed by local companies, and are supported by junior and senior academics. The course (5 ECTS) develops over 6 months, where "lessons" are substituted by "meetings". There are weekly meetings with three teachers (senior academics) from the three disciplines involved. Biweekly meetings at the university with the external companies are also scheduled. A key role in the initiative is played by the junior researchers assigned to each group, who become project managers as well as the main reference for both the group and the external company. There are no traditional lessons taught, but seminars or lessons "on demand" when gaps in knowledge emerge during the project development. The main role of senior researchers is to provide more experience advice to tutors and students, and to solve problems that can rise.

1.2 The case study

One of the proposed case studies in the 2013-14 LaPS edition was related to the valorisation of cultural heritage. Central to the project was the concept of slow tourism, i.e. the range of tourism activities focused on environmental friendliness and rediscovery of local traditions and cultural knowledge (Gardner 2009, Dickinson et al. 2011). Slow tourism is a special case of sustainable tourism, which, according to the definition of sustainable development (World Commission on
www.ogrs-community.org

Sustainable Development 1987), outlines a tourism addressing the needs of the visitors, the industry, the environment and the host communities by taking full account of its current and future economic, social and environmental impacts (UN World Tourism Organization 2005). As the amount of spatially indexed available information (as well as the technologies to acquire them) have known a tremendous growth even in the field of tourism, GIS tools represent crucial elements to be exploited for a successful promotion of tourism-related activities including slow tourism. Nevertheless, it often happens that the use of GIS technology is either neglected or not accompanied by an adequate tackling of the problem also from the managerial point of view.

The LaPS project was focused on the promotion of slow tourism in the area of Via Regina, i.e. the road which overlooks the west coast of Lake Como and highlights a system of soft mobility links across a mountainous region at the border between Italy and Switzerland. Via Regina has been a fundamental trade and pilgrim route since the old Roman times, fostering cultural and commercial exchanges between Italy and Switzerland and throughout Europe as well. Despite outlining a variety of beautiful paths surrounded by an astonishing naturalistic landscape, this transalpine system has never been properly grasped from the tourism point of view. For this reason an INTERREG project between Italy and Switzerland, named "The Paths of Via Regina", has recently started (Brovelli et al. 2013) in which Politecnico di Milano is the project leader. Gathering experts from several disciplines (landscape planners and designers, geomatics engineers, and experts in cultural heritage), the project seeks to valorize the cross-border area covered by the paths through a rediscovery of the common European identity into the artistic and cultural heritage. Main beneficiaries of the project are the local communities, that can enhance and exploit the tourism potential of their territories, and the travelers and tourists themselves (both local and foreign) who can safely undertake a trip to appreciate the nature, and together exploit the services available along the paths. The LaPS case study derived from the whole INTERREG project but addressed the specific purpose of advancing slow tourism activities (e.g. hiking, cycling, and walking) in the cross-border area formed by Cernobbio municipality (Italy) and the extreme south part of Ticino Canton (Switzerland), covering together just a part of the INTERREG project region.

2   Methodology

According to the LaPS organization, the project team was composed of students from the three MSc Schools taught at the Como Campus of Politecnico di Milano, in particular three students from Management Engineering, one student from Environmental Engineering, and one student from Science Computing Engineering. The project aimed first of all to create a strategic plan and perform a market analysis, in order to control its development from the managerial perspective. Exploiting secondary data provided by both cultural organizations and

administrations (specially Cernobbio and Como municipalities) as well as primary data collected from direct interviews, the characteristics of the typical tourist flow in the area (i.e. provenance, period, length of stay, and age of tourists) were determined. The segmentation and targeting steps were then carried out to identify the most significant customer groups according to the project objective, which turned out to be tourists from Italy and Europe. A competitive analysis was performed afterwards to outline advantages and disadvantages of similar slow tourism projects according to four discriminating factors, i.e. tourism facilities, information access and quality, use of Web technology (including GIS), and focus on local traditions. Thanks to a positioning analysis, the combination of the last two factors was recognized as a key for the project to be perceived as a high-performance offer compared to the existing competitors.

Based on all these analysis, and in strict cooperation with the stakeholders involved (mainly administrations and cultural associations), some concrete products were created in order to fulfill the project objective. Technological innovation was first achieved by developing ad hoc GIS Web platforms to guide tourists through the discovery of the cultural richness and slow tourism potential of the area. To build these Web platforms a preliminary step of data collection and pre-processing was essential. The official vector cartography of the involved administrative entities (Cernobbio municipality and Ticino Canton) was already available at Politecnico di Milano as it was provided in the frame of the main INTERREG project. Other geospatial data were obtained from the cultural associations involved in the INTERREG and the administrations themselves, which provided several layers representing hiking, trekking, and mountain biking paths, historical and cultural information, transportation and public services. Several geospatial data (mainly related to lodging and eating services) were then manually obtained from an extensive research and/or a field GPS survey. Other useful data (including base maps and aerial orthophotos) are served through standard Web geo-services from institutional administrations (e.g. the Italian Ministry of Environment), and a Digital Elevation Model (DEM) of the area with ground resolution of 20 meters is also available. Many of these geospatial data required pre-processing operations to be optimally arranged (in terms of both geometry and attributes) for Web publication. This was mainly achieved using QGIS, on which a short course was given during one of the LaPS meetings.

Due to its high performance (McKenna 2011) and ease of use compared to other geospatial Web servers GeoServer was chosen to serve data as WMS/WFS layers. Two Web platforms were then developed allowing users to visualize and query the published data. The first was thought for an access from traditional computers and was developed using OpenLayers, Ext JS, and GeoExt (see Figure 1). Besides a map panel displaying the layers, it provides a layer menu, the traditional navigation commands, a scalebar, and an indication of the current map scale and cursor position. Users can select one among various base maps (including Google layers

and OpenStreetMap), and they can query the layers published by GeoServer in order to access their descriptions and pictures inside a popup.
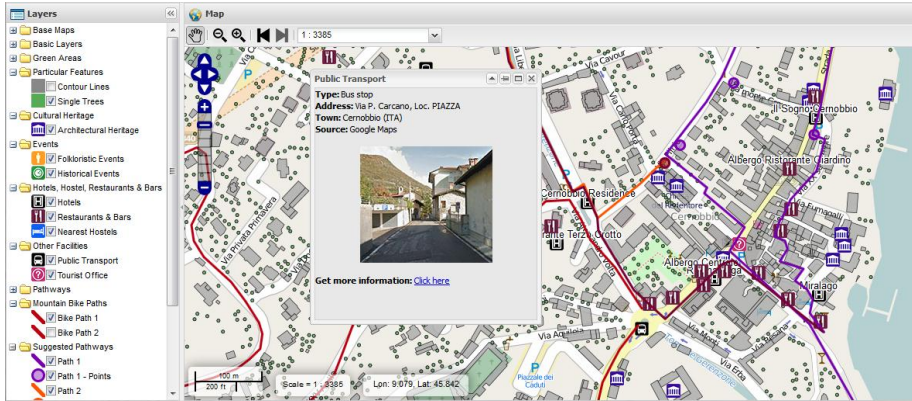
Figure 1: Web platform for traditional computers

The second Web platform was instead thought for being accessed from small-screen mobile devices (typically smart phones) while walking on the mountain paths. Coupled with the jQuery Mobile framework, OpenLayers was again used to build a touch-screen enabled application with an essential graphics (see Figure 2). Layers are this time manageable from a separate page accessible through the *Select Layer* button. Exploiting one of the mobile device geolocation services (e.g. the GPS), the *Your Position* button adds the user position as a new layer on the map. Query results are this time shown in a separate page (see Figure 2), thus resulting to be much more suitable for the small screen of mobile devices.

The managerial strategy finally included a promotional plan and a performance evaluation, which are key steps for the overall success of the project. Promotion was carried out through both online channels and physical channels. The former consisted of a dedicated website as well as social network interactions, whose purpose was to advertise slow tourism and attract new people to the project area. Physical channels addressed instead the tourists who were already present in the area. Examples were brochures and posters about slow tourism around Via Regina which were spread in restaurants, hotels, tourist offices, local shops, and travel agencies. A set of Key Performance Indicators (KPIs) was finally identified to quantitatively control the project development over time and measure its overall impact.

Figure 2: Web platform for small-screen mobile devices

3 Conclusions

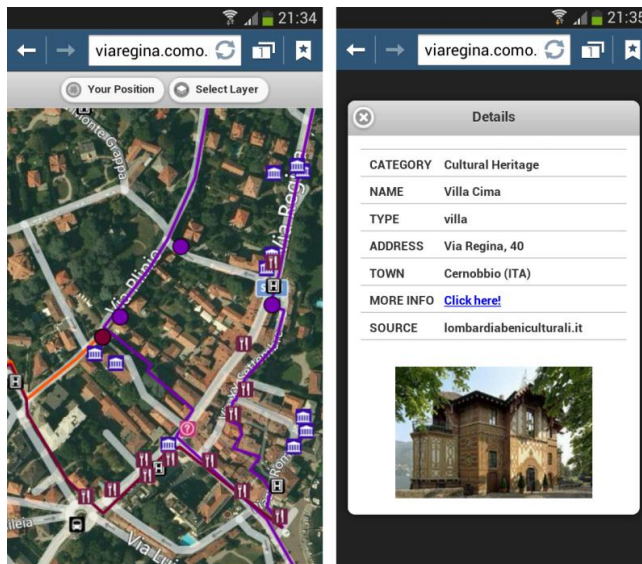Management of geospatial data is becoming increasingly important in a wide range of new and unconventional fields. One of them concerns disciplines such as cultural heritage and tourism promotion, in which projects tend often to discard either the use of GIS technology or the coupling of the related technical achievements with a proper managerial infrastructure that can frame and monitor the development. This study investigated the use of open source software in the highly multidisciplinary context of the Lake Poli School (LaPS) educational project. In the frame of a real research project, the student team addressed the issue of slow tourism promotion by tackling it from many complementary perspectives, developing tools that could be largely reused for similar case studies. The role of Management Engineering students was crucial in building the business model, which answered questions such as: who is going to maintain the system? Which tools can be designed to pursue a self-financing cultural network? GIS experts, on their side, understood the importance of considering these elements and tried to identify pilot areas for testing the INTERREG project feasibility on a small scale. The developed Web platforms provided support for a fruitful use of open source software in the field of tourism, as full customization of the products was possible according to the stakeholders' needs. The study also showed the suitability of open source software for educational purposes, as students were instructed on how to use the software and guided in the development of the applications. All in all, the work constituted a shining example of the interactions needed in both the

www.ogrs-community.org

university and business worlds to overpass the boundaries of single disciplines, in order to better address and solve the complex issues of current society.

Acknowledgments

References

Brovelli, M. A., Hogan, P., Minghini, M., & Zamboni, G. (2013), 'The power of Virtual Globes for valorising cultural heritage and enabling sustainable tourism: NASA World Wind Applications', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XL-4**/W2, 115-120.

Dickinson, J. E., Lumsdon, L. M., & Robbins, D. (2011), 'Slow travel: Issues for tourism and climate change', *Journal of Sustainable Tourism* **19**(3), 281–300.

Gardner, N. (2009), 'A manifesto for slow travel', *Hidden Europe* **25**, 10–14.

McKenna, J. (2011), 'WMS Performance Shootout 2011', FOSS4G 2011, **URL**: *http://www.slideshare.net/gatewaygeomatics.com/wms-performance-shootout-2011* [accessed: 2014-04-29].

UN World Tourism Organization (2005), 'Sustainable Development of Tourism', **URL**: *http://sdt.unwto.org/en/content/about-us-5* [accessed: 2014-04-26].

World Commission on Sustainable Development (1987), 'Our common future – Report of the World Commission on Environment and Development', Annex to the UN General Assembly Document A/42/427, Development and International Economic Co-operation: Environment, 4 August 1987.

Educating the next generation of open source GIS workers

Phillip Davis

Department of Computer Science, Del Mar College, Corpus Christi, Texas, USA

Abstract: The global geospatial technology industry, in a recent study by UK-based Oxera commissioned by Google in January 2013, has been estimated at $150 USD billion to $270 USD billion per year ($110 billion euro to $199 billion euro). In a similar US-focused study, also commissioned by Google in 2013, the Boston Consulting Group (BCG) found the geospatial services industry employs approximately 500,000 people and generates around $75 (USD) billion in annual revenue ($55 billion euro). By any measure, the geospatial industry is large one, both in the US and globally. With many of the current generation of the world's geospatial workers nearing retirement age in the next decade, it has become imperative to increase the number of well-qualified graduates knowledgeable in the latest geospatial technology to replace retiring workers and meet the demand for even more workers in this expanding industry.

Keywords: FOSS4G, education, QGIS, GTCM, training

1 Introduction

The global geospatial technology industry, in a recent study by UK-based Oxera commissioned by Google in January 2013, has been estimated at $150 USD billion to $270 USD billion per year ($110 billion euro to $199 billion euro). In a similar US-focused study, also commissioned by Google in 2013, the Boston Consulting Group (BCG) found the geospatial services industry employs approximately 500,000 people and generates around $75 (USD) billion in annual revenue ($55 billion euro). By any measure, the geospatial industry is large one, both in the US and globally. With many of the current generation of the world's geospatial workers nearing retirement age in the next decade, it has become imperative to increase the number of well-qualified graduates from higher education programs, knowledgeable in the

latest geospatial technology, to replace retiring workers and to meet the demand for even more workers in this expanding industry.

2 Barriers to open source software adoption in higher education

In the US, software from a single vendor is used almost exclusively by 90% of the 1400 colleges and universities offering an academic GIS program. By focusing heavily on the US higher education market for more than 20 years, this vendor can legitimately claim their product is used to train 9 out of every 10 graduates in the US. Although their nonacademic business share of the global market was estimated at 40% and approximately 70% in the US in 2012, they dominate the US higher education section by a disproportionally large margin. Open source geospatial software, by comparison, is found only infrequently in US academic programs.

By providing a robust and well marketed GIS education program, this vendor dominates the academic GIS market by a large margin. In the latest national survey of US academic GIS departments (GeoTech Center; annual report, 2012), only 5% of colleges and universities reported offering any form of open source geospatial software. This same survey revealed that more than half of the faculty responding indicated an interest in using open source in their classrooms and labs. While only 5 out of 100 US educators reported using open source software, more than 50 out of 100 are interested in using open source, but there are barriers to its adoption that must first be overcome.

The vendor's well-funded higher education division is able to offer a complete, ready-to-use curriculum and support products to academics at competitive pricing. This support includes: a) textbooks and lab manuals for all levels of learners, b) a virtual campus of online courses, c) professional development for educators through workshops, and d) robust community of practice through regional user's group and conferences. Compare this to the open source educational resources, where there is a lack of ready-to-use curriculum, limited opportunities for professional development, limited number of textbooks and online courses, and a small, but growing community of practice among educators using open source software. The goal of our research is to increase the quantity and quality of open source software curriculum resources available to higher education faculty in order to boost its broader adoption in colleges and universities across the United States by removing the curriculum resource barrier to open source software.

3 Developing a national standard for geospatial technology workers

In 2010 the US Department of Labor's widely recognized Geospatial Technology Competency Model (GTCM) was completed and published. This model
www.ogrs-community.org

(http://www.careeronestop.org/competencymodel/pyramid.aspx?geo=Y) provides a comprehensive list of the knowledge, skills, and abilities (KSA) required of workers in the geospatial technology industry. The model is represented as a pyramid, with the most fundamental skills at the base and building upward into more specialized knowledge areas. Figure 1 depicts this tiered model, with the most fundamental KSAs at the bottom.
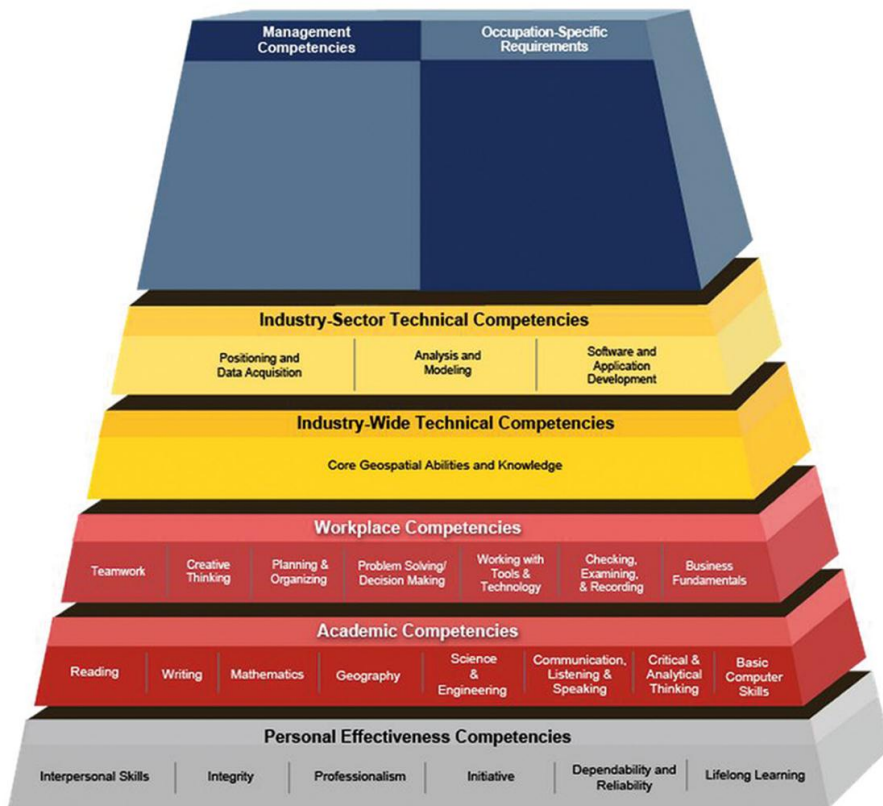
Figure 1: The US Department of Labor Geospatial Technology Competency Model (GTCM)

This model has been used by hundreds of educators in the US, Europe and Asia since 2010 to align GIS courses and curriculum with industry-identified KSAs. The GTCM was used between 2011 and 2013 to direct a team of 40 US geospatial educators to build a series of GIS course outlines that meet the KSAs defined by the model. Figure 2 depicts part of a sample of the KSAs identified by a detailed job analysis of the GIS Technician position. This job analysis is determined through an intensive two day workshop of actual GIS workers led by a DACUM facilitator. The facilitator is a professional trained in group session workshop management.

These "DACUM" workshops are used by professionals to determine what knowledge and skills are needed precisely by workers in that particular industry. DACUM is an acronym of the terms "Developing A Curriculum" and has been used since the 1960s, where it was developed by the Ohio State University Center on Education and Training for Employment. A decades-long series of geospatial technology DACUMS (1997-2013) was consolidated using the meta-analytical methodology and published by the GeoTech Center in 2010 in the URISA Special Education Journal. This consolidated DACUM document was used in establishing the curriculum required by modern GIS programs of study at colleges and universities in the US.

DACUM Research Chart for GIS Technician

Duties		Tasks			
A Create/Maintain Geographic Data Sets	A-1 Research Data Sources	A-2 Acquire Data	A-3 Design Database Structure	A-4 Create Database Tables	A-5 Perform Data Entry (e.g. digitize geographic features, college GPSdata, collect field data, create Auto-Cad
B Fulfill Map/Data Requests	B-1 Gather User Requirements	B-2 Research Data Sources	B-3 Design Maps	B-4 Design Cartographic Elements	B-5 Create Maps
C Conduct Data Analysis	C-1 Respond to Geographic Queries	C-2 Run Database Queries and Calculations	C-3 Interpret Imagery	C-4 Automate Manual Processes	
D Manage Geographic Data Sets	D-1 Manage Data Files	D-2 Design Database Structure	D-3 Create Database Tables	D-4 Create Database Queries	D-5 Create Database Reports
E Create Technical Documentation	E-1 Maintain Data Dictionaries	E-2 Create and Update Meta Data	E-3 Document Project Results		
F Perform Project Management	F-1 Gather User Requirement	F-2 Develop Project Plan	F-3 Estimate Task Schedule	F-4 Coordinate with Project Team	F-5 Coordinate Projects with Stakeholders
G Manage Equipment and Supplies	G-1 Maintain Plotters and Printers	G-2 Order Supplies	G-3 Schedule Equipment	G-4 Maintain Hardware and Software	
H Participate in Training and Outreach	H-1 Create Presentation Materials	H-2 Educate and Promote GIS Capabilities	H-3 Give Informational Presentations	H-4 Organize Staff/Department Training	H-5 Train Other Staff
I Perform Related Technical Duties	I-1 Manage Web Content	I-2 Prioritize Work Load	I-3 Respond to Public Records Request	I-4 Provide Technical Support	I-5 Provide Technology Recommendations

Figure 2: Sample DACUM of the Knowledge, Skills, and Abilities for GIS Technician (Source: http://www.instruction.greenriver.edu/gis/DACUM.pdf)

4   Building a national curriculum framework

Beginning in 2012 the US Department of Labor National Information, Security & Geospatial Technology Consortium (NISGTC) created a series of five GIS courses, aligned with the GTCM, based on the Meta-DACUM report. The result is a complete set of five GIS courses, including lecture and laboratory curriculum materials. These courses include: a) GST 101—Introduction to GIS, b) GST 102—Spatial Analysis, c) GST 103—Data Acquisition and Management, d) GST 104—Cartography, and e) GST 105—Remote Sensing. These courses meet the knowledge, skills, and abilities (KSA) as defined by the Geospatial Technology Competency Model (GTCM), depicted in Figure 1. The lecture theory portion of these five courses is generic in nature and applicable to any software implementation. The lab component for the current courses is built using Esri ArcGIS 10.1 proprietary software, as required by our contract award with the US federal funding agency, the US Department of Labor. The courses are complete now and ready for distribution under the Creative Commons 3.0 license allowing for the free use and redistribution with attribution from the Dept. of Labor's website (http://nterlearning.org).
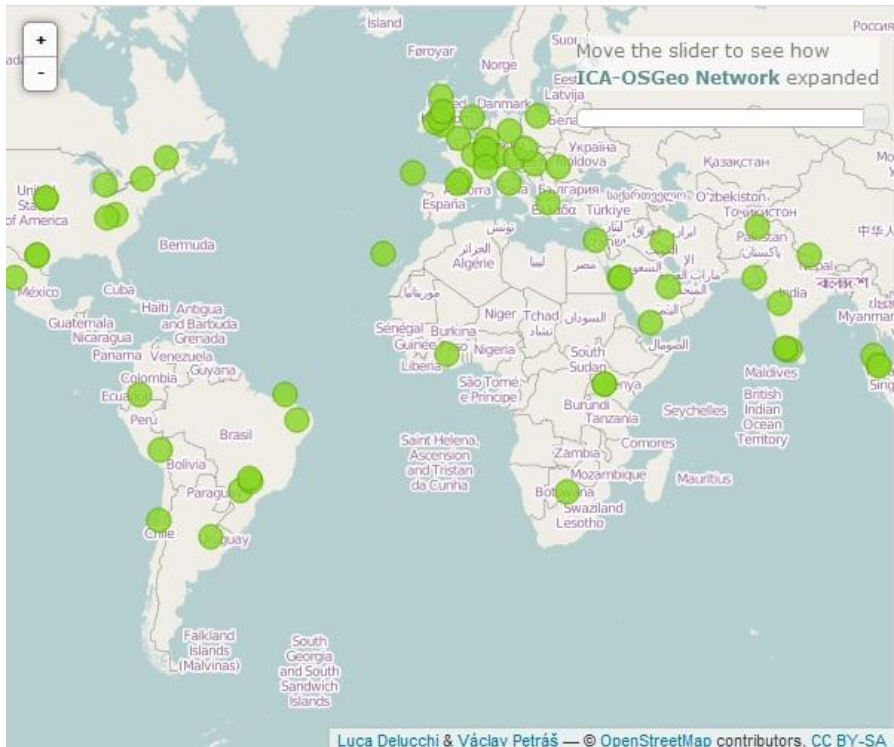

Figure 3: "Geo for All" Open Source Global Lab Network

OGRS 2014

5 Open source software curriculum resources

In 2014 a new series of complimentary QGIS laboratory has been published for the five GTCM GIS courses. With this complete set of courses, labs, and support material, we are now prepared to begin building the open source geospatial educational community of practice with educators at colleges and universities worldwide. When these QGIS laboratory materials are completed in August 2014, we will be prepared to launch a new initiative to increase the adoption of open source geospatial software in colleges and universities in the US and around the globe. It is our goal to both compliment the proprietary software in existing GIS programs, as well as assist those colleges and universities desiring to start new GIS academic programs based on the open source software community. By leveraging the rapidly expanding ICA-OSGeo Open Laboratory Network "Geo for All", we will offer our curriculum free of charge under the Creative Commons BY 3.0 license. It is expected that the curriculum will receive a rapid uptake among academic educators, commercial trainers, and self-directed individuals desiring to learn about open source geospatial software.

Acknowledgments

References

Johnson, J.   (2010), 'What GIS Technicians Do: A Synthesis of DACUM Job Analyses.' Volume 22 • No. 2 • 2010. Journal of the Urban and Regional Information Systems Association, 31-40.

Schutzberg, A. (2013), 'Ten Things Esri is Doing to "Move the Needle" on GIS Education', May 02, 2013 Directions Magazine

Study, (2013), 'What is the economic impact of Geoservices?', Oxera, Inc., **URL**: *http://www.oxera.com/Latest-Thinking/News/January-2013/Oxera-quantifies-the-benefits-of-Geo-services-to-g.aspx*

Study, (2012) GeoTech Center; annual report, **URL**: *http://www.geotechcenter.org/uploads/2/4/8/8/24886299/2010_national_faculty_survey_results.pdf*

www.ogrs-community.org

The impact of teaching using open source geospatial software

Jens Ingensand and Bertrand Cannelle

School of Business and Engineering Vaud (HEIG-VD), University of Applied Sciences Western Switzerland (HES-SO), CH-1401 Yverdon-les-Bains, Switzerland

Abstract: This paper explores the impact of using open source geospatial software for education in general. On one hand, educational institutes are a driving force behind open source software development, on the other hand students, once they have completed their education, are likely to use the software that they know from their studies in their future jobs. Several researchers have proven that many courses that use geospatial software can be entirely taught using open source software. In this paper we present the results of a questionnaire that we sent to several educational institutes throughout the world regarding their usage of open source software, their relation to propriety software companies and also what triggered the change from proprietary software towards open source software.

Keywords: Open source geospatial software, user study, teaching

1   Introduction

In recent years the use of open source geospatial software in public administrations and private companies has been increasing. For instance in France the IGN (French National Mapping Agency) uses open source geospatial tools such as PostgreSQL/PostGIS for storing and distributing their data (Fuchs 2006); in Switzerland, several federal and regional agencies use open source API's for the development of their online platforms such as map.geo.admin.ch (Moullet 2013). Also companies, as for instance the Swiss Post, started using open source geospatial software.

However the usage of open source software in companies or administrations is absolutely not equal to no cost as Shaikh and Cornford (2011) pointed out. In order to be able to use open source software, an organization needs experts who know

what they are doing and where to look for support. Furthermore the organization needs to assure a continuity in their services. If an expert leaves the organization, another expert who is familiar with similar open source software knowledge needs to be recruited.

Another important point is that open source software is known to be less "user-friendly" than proprietary software (Iivari et al. 2008, Nichols & Twidale 2003) and often, in order to create new systems or to implement data handling processes, developers and administrators need to assemble several different products that use different input or output formats, or that are utilized through very differently looking interfaces. Software coming from companies selling proprietary software on the other hand has interaction manners, terms, and interfaces and so forth that are standardized for all their products.

2 The impact of using open source software in teaching

We suggest that the fact that open source software is used in courses and projects at educational institutions plays a major role in the increasing usage of open source software in general. On one hand the initiators of open source software projects are often to be found at educational institutions because the need for a specific tool made them develop the software themselves - on the other hand open source software offers the possibility to inspect and to modify the source-code and thereby to understand how for instance an algorithm has been implemented.

Furthermore using open source software in education also allows to shape tools for special needs, for instance within student projects or research projects. Open source software also helps introducing students to programming. In the case of a research project it is easier to continue the development process and allow all partners of the project to use the software and to maintain and upgrade the source code.

However within the limits of a study program it is almost impossible to present all existing software and systems in depth to the students and lecturers need to make choices. Therefore we can ask the question - why teach one or several proprietary systems rather than open source systems?

Several researchers; e.g. Steiniger & Hunter (2013) or Ellul (2012) have shown that it is fully possible to replace proprietary geospatial software products such as GIS, geospatial databases, remote sensing software and so forth with open source alternatives in courses taught at their institutes.

So what reasons prevent educational institutes from exclusively using open source software? Rey (2012) for instance has described concerns of students who wanted to learn the proprietary software tools that they needed for their CV in order to apply for jobs. Moreover colleagues were worried about the relationship to

proprietary software companies, since they also sponsor research; organize meetings, seminars and so forth.

However from the perspective of proprietary software industry, the competition has changed. 20 years ago software development companies were facing competition from other proprietary software companies. Today they are also facing open source software. Software development companies of course are aware of the fact that people, once they have learned a specific tool, tend to prefer the same product at their new jobs and are thus influencing the usage of geospatial software. Educational institutes thereby have a certain power in influencing the usage of open source software: they are often involved in the programming of the software and more importantly shape future users.

3   User study

In order to analyze the impact of teaching open source software, we decided to send questionnaires to colleagues working at other research institutes. In this questionnaire we wanted to know:

- if open source software has been chosen for teaching and why
- if proprietary software had been used before: what has triggered the change and what was the reaction of the students
- if the proprietary software industry has undertaken steps in order to promote their products for teaching (or/and to prevent institutions from using open source software)
- if there has been feedback from former students regarding the fact that open source software has been taught / has not been taught. Did they have problems adapting to a different kind of software.

We sent the questionnaire to 20 lecturers in Switzerland and to 50 lecturers teaching at institutes throughout the world.

4   Results

We received 15 answers from Switzerland, Germany, France, UK, Finland and Canada. Four out of these 15 answered that no or almost no open source software is used.

Major reasons for both the usage of open source software and the switch from proprietary software were the absence of license costs and the possibility to modify the code. Three times we received the answer that lecturers already had acquired the experience during their research and were therefore also using it in their courses. Other reasons for using open source software were performance, scalability, flexibility, major support for interoperability, good online documentation and easiness to install (no complicated licensing software) and also

the fact that for some specific subjects there is no proprietary alternative to open source software.

We also received the answer that the lecturers wanted to open the student's minds to alternatives to proprietary software. One lecturer answered that their faculty also encourages the usage of open source software. Three of the four people who answered that no open source software was used were from the same university.

The answers received did not confirm Rey's (2013) observation that students wanted to learn to use specific proprietary software and thus had a negative attitude to learning open source software. We did not receive any answer about negative student feedback on the usage of open source software except one person mentioning that students were annoyed when the software wasn't stable. Overall students were either indifferent or positive.

Concerning the question if proprietary software industry has undertaken steps in order to increase the usage of their products we received very different answers. Three out of the 10 people who use open source software for teaching answered that proprietary software industry neither puts them under pressure nor are doing an aggressive promotion of their products. Seven people replied they do feel a pressure from the software industry. Methods used to promote software products included aggressive promotion campaigns, discounts, access to resources, agreements with the institution or even the ministry of education, the requirement to book a vendor once a year or the requirement to report software usage.

There was little feedback of former students who had to adapt to different software than the software they learned to use in courses. One answer was that some students might have had problems due to a different GUI; three people answered that students did not have any problems; also due to the fact that conceptual know-how and skills are more important than the knowledge of how to use specific tools.

5   Conclusions

Our survey showed that open source software is used in most of the educational institutes who completed the survey. One interesting finding is that three out of the four people who answered that they were not using open source software at all were from the same research institute. At the same time one person answered that using open source software was encouraged by the faculty. These facts suggest that the influence of the department, the educational institute or even the ministry of education has a significant influence on the choice of software for teaching. Moreover, as Rey (2012) describes, relationships with proprietary software industry are sometimes cultivated over a long period of time and thus not so easy to change or to break.

The change from proprietary software towards open source software appears to be frequently triggered by people who used open source software for their
www.ogrs-community.org

research projects. Once they had learned to use a particular software, they felt that they also wanted to use it for teaching.

The measures proprietary software industry has taken in order to increase the usage of their software in courses appears to be varying; some people answered that local vendors did not care very much, while others reported aggressive campaigns and contracts with their institutes or the ministry of education in order to force the usage of their projects.

None of the people who answered the survey were using open source software at 100%. Most people answered the usage was around 50%.

We argue that the main role of an educational institution is to prepare students for the life after their studies. Most importantly the goal is to make students understand the use of specific techniques within geodesy, photogrammetry, GIS, cartography etc- and to use specific software in order to apply these techniques. However, both open-source and proprietary software is changing constantly. Open source software changes sometimes even every week and it is not always as stable as proprietary software which changes roughly every other year. Lecturers and professors must be aware of this evolution and therefore need to constantly adapt their courses and exercises.

We argue that the choice of software for teaching geospatial concepts has an impact on the usage of geospatial software in general. Choosing open source software has several advantages, but it also has consequences. Professors and lecturers need to be aware of these consequences and need to make students aware of the choices they have made. Our own choices reflect the demand of our student's future employers, but also the necessity to make students understand the theoretical concepts that are taught.

References

Ellul, C. (2012), Can Free (and Open Source) Software and Data be Used to Underpin a Self-Paced Tutorial on Spatial Databases? In: Transactions in GIS, 16(4): 435-454.

Fuchs, F. (2006), Institut Géographique National (IGN) PostGIS Case study, URL: http://postgis.refractions.net/documentation/casestudies/ign/ [accessed 2014-02-03].

Iivari, N., Hedberg, H. & Kirves, T. (2008), Usability in Company Open Source Software Context - Initial Findings from an Empirical Case Study. OSS 275, p 359-365 Springer.

Moullet, C. (2013), User centred design for map.geo.admin.ch. Presentation held at FOSS4G2013 conference 2013. Nottingham, England, 17th-21st September 2013.

Nichols, D. & Twidale, M. (2003), The Usability of Open Source Software. First Monday 8(1).

Rey, S. (2012), Open source spatial analysis: lessons for research and education from PySAL. Proceedings of the OGRS 2012 conference, Yverdon les Bains, Switzerland, 24th-26th October 2012.

Shaikh, M. & Cornford, T. (2011), Total cost of ownership of open source software: a report for the UK Cabinet Office supported by OpenForum Europe. UK Cabinet Office, London.

Steiniger, S. & Hunter, A.J.S. (2013), Teaching GIScience with Free and Open Source Software? – A first Assessment. GIScience 2010: 6th Int'l conference Zurich, September 14-17.

E-Foto and MicMac: synergetic benefits of integrating open-source digital photogrammetry software

João Araujo Ribeiro[1], Marc Pierrot-Deseilligny[2], Jorge Luís N. S. Brito[3], Orlando Bernardo Filho[1], and Guilherme Lúcio Abelha Mota[4]

[1]Department of Systems and Computer Engineering, Rio de Janeiro State University, Rio de Janeiro, Brazil
[2]École Nationale des Sciences Géographiques, IGN/LOEMI, Université Paris Est, France
[3]Department of Cartographic Engineering, Rio de Janeiro State University, Rio de Janeiro, Brazil
[4]Department of Informatics and Computer Science, Rio de Janeiro State University, Rio de Janeiro, Brazil

Abstract: E-Foto is a project developed at Rio de Janeiro State University as an graphical educational software for digital photogrammetry. MicMac is a set of command line general purpose tools for photogrammetric or image processing. In 2013, the E-Foto research team decided to borrow some functions from MicMac software in order to improve E-Foto capabilities. This paper reports and analyses the work of integration of E-Foto and MicMac tools, as well as the difficulties to accomplish this task.

Keywords: Photogrammetry, Open Source, Software Development

1 Introduction

The main goal of the E-Foto project is to build a complete educational, digital photogrammetric workstation under the General Public License (GNU/GPL)(Brito et al. 2011, Mota et al. 2012). The E-Foto project started in 2002 at the *Instituto Militar de Engenharia* (IME, Brazil). In 2004 the research moved to *Rio de Janeiro State University* (UERJ, Brazil), where its development is ongoing.

The E-Foto modules were initially developed using a non-integrated approach. In 2010 the first integrated version of the E-Foto software was launched. At that time, the E-Foto developer's team also provided a set of didactic material and data, designed to permit a beginner to learn digital photogrammetry using a self-learning/teaching

approach. In August of 2012 the E-Foto project won a bronze medal in the 6th edition of "**The Computer-Assisted Teaching Contest**" (CATCON), organized by the *International Society of Photogrammmetry and Remote Sensing*, in Melbourne, Australia.

The MicMac software (Pierrot-Deseilligny 2014) is a general purpose tool for automatic matching of digital images developed since 2005 at *Institut National Géographique* (IGN, France) under the *CeCILL-B license*, which is an adaptation of the L-GPL license regulated by French law. The MicMac family of tools is composed of several programs that work together to accomplish photogrammetric tasks. The IGN team developed more than 20 tools, each for a specific photogrammetric or image processing task. For the sake of simplicity, in this paper we will use the word "MicMac" indistinctively in reference to all these tools.

In 2013, the E-Foto research team decided to borrow some functions from MicMac software. This allowed us to improve E-Foto capabilities, without having to program more complex code, since this was already developed on MicMac. As both projects implement photogrammetric functions and are developed under an open source philosophy, it was natural to join efforts to make a better program with less programming.

This paper reports and analyses this work of integration of E-Foto and MicMac tools, as well as the difficulties to accomplish this task.

2   Development philosophy

The main objective of E-Foto project has always been to develop a pedagogical tool for the self-teaching/learning of photogrammetry. In order to accomplish this didactic goal, we created a graphical user interface (see Figure 1) which is *as simple as possible*, and which can guide the student through the several steps of a complete photogrammetric project. Each step is taken by an E-Foto module: Project Manager, Interior Orientation, Exterior Orientation by Spatial Resection or Phototriangulation, Stereo Plotter, DEM Extraction, and Orthoretification. Since these concepts are broadly used in photogrammetry, E-Foto users could better understand and practice the theoretical concepts taught in formal classrooms. Figure 2 shows a simplified workflow of a photogrammetric project inside E-Foto. Dashed lines represent alternate paths in the execution of a project.

Even though it is conceived as an educational tool, we have achieved good results in real world photogrammetric projects, comparable to commercial workstations (Laranja et al. 2013, Lengruber et al. 2012). However, it must be pointed out that E-Foto has its limits. For instance, it only works well with airborne imagery taken by a calibrated photogrammetric camera, based upon a classical photogrammetric mapping project. Moreover, E-Foto currently still lacks an automatic tie point generation tool.

The project was developed using free environments such as GCC/G++ (C++) and Qt, a cross-platform application framework widely used for developing application software with a graphical user interface. Thanks to this, E-Foto has both Windows and Linux versions, for 32 and 64 bits operating systems. Nevertheless, we have achieved better results on Linux, thanks to its better memory management. Thus, all E-Foto development work is first done in Linux and then ported to the Windows operating system.

The focus of the MicMac project is to develop a powerful system, in which each photogrammetric task is accomplished by a different tool. Almost all tools have a

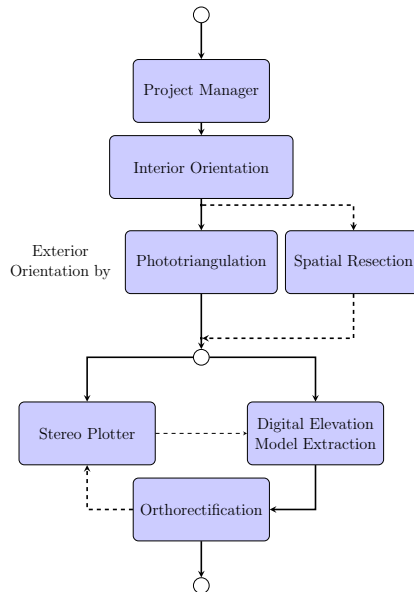Figure 1: E-Foto Phototriangulation Module.

Figure 2: Basic E-Foto Workflow.

Figure 3: Photo and Digital Elevation Model generated by MicMac tools.

command line interface with great use of regular expressions or parameters. These could be intimidating to a new user. For more complex tasks, MicMac uses an XML file with a specific syntax for parametrization. However, the complexity of use is also MicMac's strength (see Figure 3). The software is capable of complex photogrammetric tasks and it works with non-calibrated consumer cameras, thanks to its tools for camera callibration and implementation of the *Shift-Invariant Feature Transform* (SIFT) algorithm (Lowe 2004) to perform automatic tie points generation.

Like E-Foto, MicMac is also developed in a free environment (C++). However, unlike E-Foto, that implements all necessary functions, MicMac can use other free software like *DCRaw*, *Image Magick*, *Meshlab* to accomplish a specific task.

3 The integration

Both projects have independent research teams and development approaches. The first decision was to leave the code of MicMac unaltered. As such, the integration of functions between both programs would be made only by calling each individual MicMac tool and not by incorporating code from it. The only way to have a communication between E-Foto and MicMac is through input and output data generated by operating system calls. Both software, E-Foto and MicMac are developed in C++, which, a priori, facilitates the incorporation of code. However, in doing this, we would have problems with code maintenance. Since both projects have small teams, we must optimize all resources. By using only calls to programs execution, we are sure to benefit from of all future developments of MicMac with minimum effort, as long as MicMac does not alter its input and output data format too much. As E-Foto implements all its photogrammetric functions, MicMac is, on this sense, the first external program used by E-Foto in a photogrammetric task. Besides that, E-Foto already uses *xsltproc* as an external program to parse *epp* file and generate the project report.

All data generated by E-Foto software is saved in an XML file with an extension

format called *epp* (namely, *E-Foto Photogrammetric Project*). This format was designed by the project's development team and allows communication between each module, keeping coherence across all the steps of a photogrammetric project. The manipulation of this *epp* file and its XML syntax are transparent to the user, so it is dangerous if the user edits its content directly. The user has access to all the internal data through a report tool integrated in the E-Foto interface. By using this interface, the user can audit and analyse all data generated without compromising the *epp* file integrity. Moreover, the user does not need to know its internal syntax. The report output can be a text file, which can be easily imported by a word processor, or an HTML page, visible on a simple ordinary browser, with full parametrization of which fields are wanted.

MicMac also uses XML files, but with a different approach. XML files serve mainly to pass configuration parameters to each MicMac tool and to deliver some intermediary results. Unlike E-Foto, which works with only one XML file, MicMac can generate several XML files during its processing.

The next decision was to choose where MicMac functions could improve E-Foto. As already mentioned, currently E-Foto deals only with classical photogrammetric airborne imagery taken by calibrated cameras. If we want to work also with digital imagery taken by *Unmanned Aerial Vehicles* (UAV), it will be advantageous to know how to auto-calibrate cameras used in image acquisition and generate photogrammetric tie points automatically.

E-Foto also lacks an algorithm for structuring a *Triangular Irregular Network* (TIN) from a cloud of 3D points generated by its *Digital Elevation Model* (DEM) module. In that hypothesis, E-Foto would benefit from integration with MicMac by requesting the specific MicMac tool not only for accomplishing this task, but also for the exploitation of DEM applications, such as the automatic interpolation and drawing of contours and topographic profiles.

MicMac software works with a lot of temporary files and uses specific names of directory and file to output. However, the user can redefine these default file names from the command line. Therefore, the first task to integrate MicMac results is to read and parse XML/text files generated at the end of its execution.

We encountered some difficulties to accomplish our task. The most important of these is a conceptual one: the integration of MicMac generated data must be done without compromising the pedagogical principles of E-Foto. The technical problems are easier to solve, but we must be careful to manage the potential impacts on the E-Foto structure. For instance, to allow an UAV photogrammetric project, we must adapt the *epp* format to new kinds of information without compromising backward compatibility.

Despite these difficulties, we are certain that E-Foto will be more flexible and useful to the photogrammetric community as a result of its communication and integration with MicMac.

4 Conclusion

This paper describes the work in progress for integrating MicMac's photogrammetric functions into the E-foto software. This kind of experience is useful to the geospatial open source community because it shows how it is possible to join efforts and benefit from different research environments and experiences to improve distinct projects, even programs with particular approaches and development philosophies.

Both projects, E-Foto and MicMac, lack updated documentation of the code and of the input and output file format used. This is very common on open source projects, which have small teams to perform several tasks, from coding to documentation. Despite this problem, as both software use XML files, some characteristics are easily understandable by a simple reading of the output file. As a final idea we believe that a formal definition of XML files used by both projects will be very useful to future developments.

Thanks to the use of tools developed by MicMac team, we hope that by mid-2014, E-Foto will have its first version capable of working with UAV imagery.

5 Acknowledgments

References

Brito, J. L. N. S., Aguiar, R. A., Silveira, M. T., Filho, L. C. T. C., da Silva Badolato, I., Pupim, P. A. B., Reolon, P., Ribeiro, J. A., Ribeiro, J., Bernardo Filho, O. & Mota, G. L. A. (2011), E-FOTO: Development of an Open-Source Educational Digital Photogrammetric Workstation, Barcelona, Espanha, pp. 356–361.
URL: *http://www.thinkmind.org/index.php?view=article&articleid=icsea_2011_15_10_10194*

Laranja, R. M., da Silva Correa, N. C. & Brito, J. L. N. S. (2013), Mapeamento Fotogramétrico Digital: Um Estudo Comparativo da Bacia Hidrográfica do Rio Piabanha nos Ambientes E-Foto e LPS, *in* L. S. G. José Carlos Neves Epiphanio, ed., 'Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR', Instituto Nacional de Pesquisas Espaciais, Instituto Nacional de Pesquisas Espaciais (INPE) , São José dos Campos , SP, pp. 2134–2140.
URL: *http://www.dsr.inpe.br/sbsr2013/files/p0222.pdf*

Lengruber, N. V., Bahia, J. G., Barbosa, L. G. & Moffati, M. S. (2012), Avaliação da Qualidade das Medições Realizadas no Módulo de Estereoplotter do Software Livre E-FOTO, *in* 'Anais do IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação'.
URL: *http://www.ufpe.br/cgtg/SIMGEOIV/CD/artigos/SReFOTO/141_4.pdf*

Lowe, D. G. (2004), 'Distinctive image features from scale-invariant keypoints', *Int. J. Comput. Vision* **60**(2), 91–110.
URL: *http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94*

Mota, G. L. A., Nunes e Silva Brito, J. L., Ribeiro, J. A., Bernardo Filho, O., Silveira, M. T., Aguiar, R. A., Badolato, I. S., Costa, S. L. & Reolon, P. F. (2012), The E-Foto Project and the Research to Implement a GNU/GPL Open Source Educational Digital Photogrammetric Workstation, *in* E. Bocher & M. Neteler, eds, 'Geospatial Free and Open Source Software in the 21st Century', Lecture Notes in Geoinformation

and Cartography, Springer Berlin Heidelberg, pp. 89–106.
URL: *http://dx.doi.org/10.1007/978-3-642-10595-1_6*

Pierrot-Deseilligny, M. (2014), 'MicMac, Apero, Pastis and Other Beverages in a Nut-shell!'.
URL: *http://logiciels.ign.fr/IMG/pdf/docmicmac.pdf*

Interactive maps on peace and conflict: an open source WebGIS application for civil education

Alexander Strunck and Lars Wirkus

Data and GIS Section, Bonn International Center for Conversion (BICC),
Pfarrer-Byns-Str. 1, 53121 Bonn, Germany

Abstract: This article gives an overview of an application, which has been developed to provide up-to-date information on various peace- and conflict-related topics and processes with the aim of supporting the political education of the interested public. The cooperation between the Bonn International Center for Conversion (BICC), a German peace and conflict research institute and the Federal Agency for Civic Education (bpb) resulted in a web- and geographic information system-based information portal. Open Source software has been used to create global interactive maps for currently seven thematic modules. BICC regularly updates the information and a complete redesign will be implemented in 2014; including the development of an English version of the portal.

Keywords: WebGIS, civil education, war, peace, map

1 Introduction

War and violent conflicts are omnipresent—be it the war in the Middle East, violent conflicts in failed states, or increasing military expenditures and exports/imports of military goods. To understand certain conflicts or peace processes and their possible interrelations, to conduct a well-founded political discussion and to support or influence decision-making, one matter is of special importance: easily accessible and, above all, reliable data and information.

The Bonn International Center for Conversion (BICC) has been conducting applied research on different themes and topics of peace and conflict for 20 years. BICC's work is guided by its vision to contribute to a more peaceful world through "Conversion Studies" (BICC homepage 2014). Since its foundation, it has been an important part of BICC's work to make available its research outcomes to the

general public via publications, conferences, presentations and discussions. The information portal uses open source geospatial software to present information on peace and conflict to the public in a new and innovative way.

Peace and conflict itself, most of their dynamics and effects but also many related interfering socio-political processes have a strong spatial reference and are highly suitable for an interactive presentation. In the context of the given spatiality, maps particularly interactive maps, are an excellent way to provide information about afore mentioned theme for a variety of different users. Since the inception of the project in 2011 all addressed topics of the website *sicherheitspolitik.bpb.de* combine interactive maps, infographics, country portraits, a glossary as well as concise information and background articles. The content presented is mainly built upon BICC's expertise. Various researchers have been contributing data and articles to the so far realised modules. For the time being the portal is meant to serve as a knowledge transfer and educational tool for mainly German-speaking communities.

2   The information portal on peace and conflict

2.1   Why develop a peace and conflict information portal?

The objective of the portal is to broaden the knowledge of the interested public and deepen their understanding of complex political issues related to either peace or conflict. It not only plays a crucial role in supporting the educational mandate of the Federal Agency for Civic Education (bpb), founded in 1952. The portal also directly supports various target groups in obtaining the knowledge needed: it provides a collection of well designed, easy-to-use maps and info graphics. Yet, the main motivation is to initiate or stimulate an intense examination of themes and processes crucial for understanding the given context in the framework of public education.

In the following article we would like to share our idea and experiences in developing and implementing such a platform. We hope to inspire other research groups of other disciplines to follow in our footsteps and create comparable open source-based geo-information portals to strengthen the connection between scientific research and education practices.

2.2   Cooperation between bpb and BICC

The bpb and the data department of BICC started planning the geo-information portal on peace and conflict in 2009. "Through civic education, the bpb is assigned to promote the understanding of political issues, strengthen awareness for democracy and willingness to participate in political processes amongst all citizens" (Enactment of the bpb 2001). In recent years the multimedia editorial department of bpb have broadened their output by intensifying the use of modern

www.ogrs-community.org

media tools and visuals like info graphics as well as interactive solutions. The information portal is one element of this multimedia approach. The cooperation created a win-win situation for both institutions; the bpb received a new, modern, interactive geo-information portal with the latest peace and conflict data and information available which supports bpb's mission to inform the public, whereas BICC has gained access to the huge user group of bpb's website and has been enabled to present its latest research findings to a wider general public.

The first phase of the project was devoted to developing and testing a portal structure which best fits the given requirements. The design of the portal, the structural composition, the orientation of its content as well as development and update procedures were extensively tested. After four years of cooperation, in 2013, both institutions decided to intensify their cooperation to sustain, update and further develop the portal. The intended long-term cooperation includes a redesign of the whole portal (2014), the development of additional modules (nine modules were originally planned), improved user-friendliness, the development of new ways of accessing the information and last but not least an English version of the portal.

2.3 Why open source?

Free and Open Source software build up on simple principles: "(1) the freedom to run the software for any purpose (e.g., may it be education or business), (2) the freedom to study and adapt the software for own needs, (3) the freedom to redistribute the software, and (4) the freedom to improve the software and to release improvements to the public" (GNU Operating System 2014). Popular open source software "attract a strong following of users and developers" (Steiniger & Hunter 2012) and support free of charge through the community. These principles enable a long term development and modification of the project software infrastructure. In the majority of cases open source software also has a stronger commitment to industry standards, such as the Open Geospatial Consortium (OGC). From an economic standpoint, open source software allows you to invest money saved through not buying proprietary software either into the conceptual development of the planned product or into tweaking the used software.

BICC used PostgreSQL with the PostGIS extension as database, UMN MapServer as geo-server, OpenLayers as user interface and MapProxy as map proxy for the WebGIS infrastructure embedded in CMS Made Simple as content management system (CMS) and pChart for graphics. The use of the aforementioned software was bound to certain requirements: All components had to be stable, had to have a good documentation and preferably a huge community to ask questions. Moreover all open source components had to prove their compatibility and the potential to be used and manipulated easily.

Most components are open source and most of them have a long history in the open source community: PostgreSQL since 1995, PostGIS since 2001, MapServer since 1995, OpenLayers since 2006, MapProxy since 2010, CMS Made Simple since

2004 (http://www.cmsmadesimple.org/) and pChart since 2008 (http://www.pchart.net/). Due to their long history most of the components are often referred to in the literature. We use the most popular open source WebGIS software on the market. PostGIS is " (...) the most powerful open source spatial database engine" (Obe & Hsu 2011). "MapServer is perhaps the most popular of these tools because it contains a lot of functionality and can run under practically any web server without requiring installation" (Obe & Hsu 2011). OpenLayers "(...) is very powerful but also easy to use—you don't even need to be a programmer to make a great map with it" (Hazzard 2011).

The database PostgreSQL and PostGIS stores all statistical and geodata that is needed for the maps, tables and graphics.

From the spatial data in the database MapServer creates one or more picture files with the map. For every map, the appearance of a layer on a map is set in different map-files. The application works on the OGC standards Web Mapping Services (WMS), Web Feature Services (WFS), or Web Coverage Services (WCS).

OpenLayers enables the users to interact with the map on the website. The user can zoom in/out, pan the map, and create an individual view of the map.

MapProxy accelerates the speed of the interactive maps. Most of the picture files that are generated by MapServer get used on a regular basis. These pictures do not need to be created for every map request. MapProxy calculates these pictures in advance and stores them systematically. After the calculation the picture files replace the map creation from MapServer.

CMSmadesimple is a PHP written CMS that provides a framework for the geodata infrastructure and a user interface to administer, edit and store all static information in a CMS database (MySQL).

As concerns the open source and open data movement, the application is, licensed under Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0 http://creativecommons.org/licenses/by-nc-nd/3.0/). The first version of the license was published in 2002 and the version 3.0 has been available since 2007. All material can be copied and redistributed, when an appropriate credit is provided, a non-commercial use is ensured and the material is not manipulated. The spread of such information on peace and conflict is supported by such a creative commons license.

3   Components and functions of the information portal

3.1   Implementation procedures

The development of an ergonomic and user-friendly web portal on peace and conflict started in 2010 with the first thematic module *military capacities*. Every module discusses a specific theme in a coherent manner. The first ideas for the website included nine thematic modules, which would be developed successively and cover the most important aspects of peace and conflict. With this approach the

www.ogrs-community.org

website could be used immediately after the completion of the first module. Each new module that was incorporated into the portal added new content and perspectives to the topic of peace and conflict in the portal which as it became more complete, attracted new user groups.

After bpb and BICC had come to an agreement with regard to the portal's content, structure and the concept of the mapping component, BICC started developing the first version of the information portal. The project team worked simultaneously on the data inquiry, preparation and manipulation, development and programming of the geo-software infrastructure, the creation of graphics and maps and the writing of accompanying information and background articles. The team joined all parts in the prepared CMS for the first module. In this phase the development of a stable geodata infrastructure and its adaption to the needs of the portal was most important. Here, the development of the layer-switcher in particular, which is used to select/de-select layers took time. Once all major functionalities had been implemented with the first module, less working hours were required for the following ones; data research and processing, map production and text writing remained the only time consuming processes.

3.2   Components and functions

The layer-switcher (Figure 1, part 3 below) was programmed to meet the coordinated concept. A maximum of three layers (polygon, hatch and point layer) can be selected simultaneously. For every selected layer (Figure 1, part 7 below) additional corresponding information (articles and graphics Figure 1, part 10) if available is provided below the map. Only predefined meaningful layer combinations can be selected, the combinations can involve one or more modules. Whenever a layer is selected the user can only select the predefined compatible layers; all other layers are greyed out (Figure 1, part 8 below). The predefinition of the layer combinations is solely based on content and the expertise and knowledge of the involved peace and conflict researchers at BICC. From a technical perspective this has been quite challenging as with every additional module the number of potential combinations increase. Every time new layers or modules are added all combinations need to be checked and possibly adapted. New combinations might lead to changes in the visualisation characteristics of certain layers—whether it has to be displayed as polygon, hatch or point layer. For all these reasons it was only possible to develop the layer-switcher in an open source environment.

The portal offers a unique collection of peace and conflict data in the form of simple data, tables, maps, and graphics. The project aims at using as much sound and open data sources as possible. Several datasets have been searched and collected from different well-known official sources such as the World Bank, the International Monetary Fund, peace and conflict research institutions, and journals. The same applies to geodata which mainly originate from sources like Natural Earth, CShapes and the FAO. The statistical data have been manipulated, arranged in categories and classified according to certain features and types of information.

They were stored systematically in the project database. Spatial data require a similar process, but the preparation needs other tools such as the desktop GIS software QGIS. The spatial data are also stored in the PostgreSQL database. The module data are updated on a regular basis and the spatial datasets with the current and historic country borders are updated when required. The background information includes a full metadata set with source, description, unit, indications and, if possible, a table with raw data.
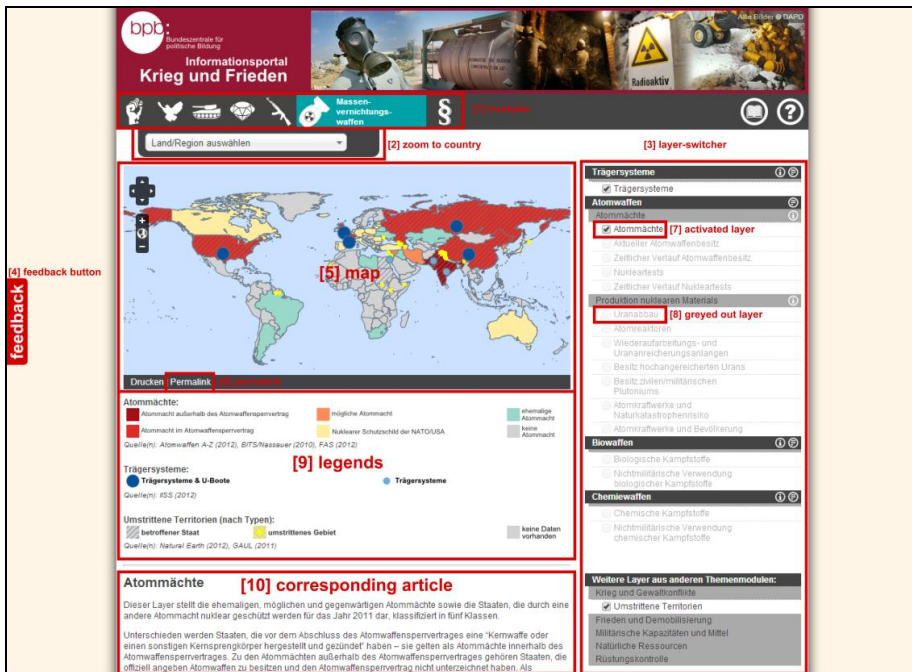
Figure 1: Screenshot—Thematic module weapons of mass destruction

3.3 Interaction with the user

At first the users could only access the information via the thematic modules. The user has to choose a module on the landing page. The interactive map (Figure 1, part 5) with the layer-switcher as the control unit on the right hand side (Figure 1, part 3) is the centerpiece of the following module page. Once one or more layer(s) are chosen, corresponding legends (Figure 1, part 9) and explanatory notes, graphics and teaser to related info and background articles are immediately shown underneath the map (Figure 1, part 10).

A click on the layer names in the layer tree offers the user links to short explanations on what is being shown in the map, how it should be read and which sources have been used. The layer tree also offers short definitions of what is being addressed by the respective layers under each subheading of a thematic module.

www.ogrs-community.org

This information can be accessed through a click on the info icon right next to the subheading.

Additional functions, besides zoom and pan, are a permalink (Figure 1, part 6), a zoom to country drop down menu (Figure 1, part 2), and an information pop-up window when you click on a country on the map. Many points of various point layers might deliberately overlay others in the standard view. The large size of the points was chosen to motivate the users to zoom in and to interact with the map. A single or a combination of layers with a specific view can be sent to another user via a permalink; this feature makes sure that discussions, for example in the classroom, are based on the same map.

In 2012 BICC developed another option to access the information stored in the project database. Besides the content-driven access through the selection of a thematic module and layers, a country-related access was put in place. By selecting a country from an alphabetical list or from a "country portrait link" given in the information pop-ups, all data used for the layers in one or in all modules are presented in a table-style country profile (Figure 2).

On the left of the information portal there is a feedback button (Figure 1, part 4). All users are invited to give direct feedback to BICC with general questions, technical problems and/or suggestions and criticism. Since 2011 some smaller errors have been reported about texts, legends or maps while feedback has triggered some content discussions. Some positive feedback has proven the portal's manifold implementation possibilities for teaching purposes in classrooms.

An external evaluation commissioned by the bpb and executed by a German University brought about a substantial revision of the website. The evaluation investigated the strength and weaknesses of the quality of information, further development, user-friendliness, accessibility, media competence, transparency, information design and the integration of other already existing products of bpb into the information portal. Suggestions and ideas that could be easily realized were directly implemented in a revised version of the website. Some suggestions which demanded major technological or structural modifications found their way into the concept of the redesign of the information portal (e.g. better readability of the texts, a bigger map and so forth) which started in late 2013.

3.4   Content addressed by the information portal

Conflicts are a normal part of our everyday lives. There are various nuances of a conflict and war between groups or nations is an extreme manifestation of conflict. The number of influences on the nuances of conflict is uncountable, but the information portal gives a good overview on the most important global indicators. Our first concept included nine different thematic modules, which are loosely based on the chapters of the German peace research textbook "Handbuch Frieden" by Gießmann and Rinke 2011. The portal fills the textbook-based modules with digital interactive data. Until now seven modules (Figure 1, part 1) and their sub-chapters have been realised: *War and Conflict (Violence, conflicts and war; Causes of*

violent conflicts), Peace and Demobilization (Peace and peace missions; Disarmament, Demobilization and Reintegration), Military Capacities (Military resources; Militarization; Arms trade), Resources and Conflict (Mineral resources and distribution of mineral resources; Resource conflicts; Resource control regime), Conventional Weapons (Conventional weapons; Small and light weapons; Landmines and cluster munition), Weapons of Mass Destruction (Launcher systems; Nuclear weapons; Biological weapons; Chemical weapons), Arms Control (History; Conventional weapons; NBC weapons; Space weapons). The last module *arms control* went online in December 2013. The two missing modules, *Good governance* and *Environmental security*, will be developed in 2014 and 2015.

At the moment the website is composed of 112 maps, 63 info graphics, 19 data tables, 71 focus articles, 22 background articles, and seven module glossaries explaining more than 210 terms.

All content is edited in plain and intelligible language to meet the different skills of the various target groups (interested members of the public, researchers, teachers and learners, policymakers, and representatives of the media).
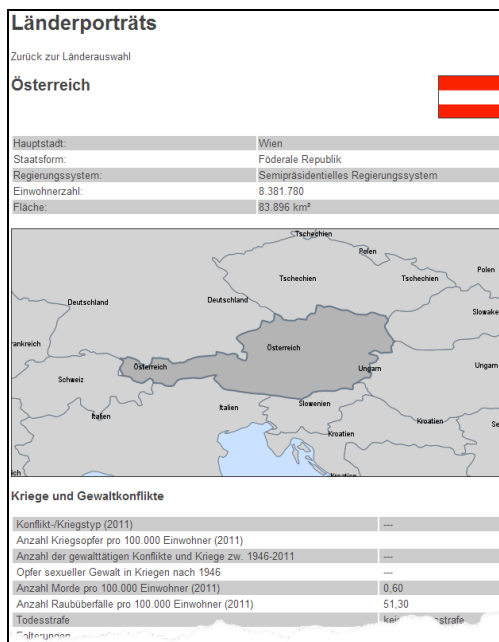

Figure 2: Screenshot - Country portrait of Austria

www.ogrs-community.org

4 Outlook

At the moment BICC is working on a complete redesign of the website. The objective of the redesign is to match the design of the portal with the main homepage of the bpb which includes a new color design, new allocations and structure of website components, a changed text presentation with more pictures and an improved on-screen viewing mode. In this context a redesign of the mapping component has also been implemented; a larger map covering the full width of the webpage not only offers a greater view but also improves interaction with the user. Finally, the backend of the information portal will be revised and changed to create a faster, more stable website.

It has also been discussed to change the country portraits (Figure 2). At the moment the information and data are only presented for one country at a time. It is planned to extend the table view to up to two more countries. Like this the user will be able to compare two or three countries in this context, which will be an advantage if, for instance, the user is interested in a specific conflict in which several countries are involved.

The IT world with hardware and software changes continuously. Therefore a constant monitoring of enhancements of the used software and testing of new products is necessary. This may lead to a revision of the existing system and to a change of certain components.

Another decision to be taken in the near future refers to the licencing model used so far—license CC BY-NC-ND 3.0 mentioned in Section 1.3. In November 2013 a new version of the license 4.0 was released. "The new licenses are more user-friendly and more internationally robust than ever before" (What's New in 4.0 2014). BICC and bpb need to review the new license and take a decision which one to use.

The biggest change in the coming months will be to set up an English version of the information portal. Both partners, BICC and bpb, are aiming at increasing the user community of the portal. The information portal that had initially been developed for the German educational market only, has proven its usefulness. Stimulated by international requests, the decision has been taken to develop an English version.

Acknowledgments

References

Bonn International Center for Conversion (2014) Homepage.
  URL: *http://www.bicc.de/* [Accessed March 21, 2014].

CMS Made Simple (2014) Homepage.
  URL: *http://www.cmsmadesimple.org/* [Accessed March 31, 2014].

Creative Common (2014) 'What's New in 4.0'.
  URL: *http://creativecommons.org/version4* [Accessed March 21, 2014].

Federal Agency for Civic Education (2001) 'Erlass über die Bundeszentrale für
politische Bildung (BpB)'.
  URL: *http://www.bpb.de/die-bpb/51244/der-bpb-erlass* [Accessed March 21, 2014].

Gießmann, H.J. & Rinke, B. (2011) *Handbuch Frieden*, VS Verlag.

GNU Operating System (2014) 'The Free Software Definition'.
  URL: *http://www.gnu.org/philosophy/free-sw.en.html* [Accessed March 21, 2014].

Hazzard E. (2011) 'OpenLayers 2.10 – Beginner's Guide', Packt Publishing Ltd.,
Birmingham, Great Britain.

MapProxy (2014) Homepage.
  URL: *http://mapproxy.de/* [Accessed March 31, 2014].

MapServer (2014) Homepage.
  URL: *http://mapserver.org/* [Accessed March 31, 2014].

Obe, R. O. & Hsu, L. S. (2011), 'PostGIS in Action', Manning, Stamford, Great
Britain.

OpenLayers (2014) Homepage.
  URL: *http://openlayers.org* [Accessed March 31, 2014].

pChart (2014) Homepage.
  URL: *http://www.pchart.net/* [Accessed March 31, 2014].

PostGIS (2014) Homepage.
  URL: *http://postgis.net/* [Accessed March 31, 2014].

PostgreSQL (2014) Homepage.
  URL: *http://www.postgresql.org/* [Accessed March 31, 2014].

QGIS (2014) Homepage.
  URL: *http://www.qgis.org/* [Accessed March 31, 2014].

Steiniger, S. & Hunter, A.J.S. (2013) 'The 2012 Free and Open Source GIS Software
Map – A Guide to facilitate Research, Development and Adoption', *Computers,
Environment and Urban Systems* **39**, 136-150.

Webel, C. & Galtung, J. (2007), *Handbook of Peace and Conflict Studies*, Routledge
Tylor & Francis Group.

www.ogrs-community.org

Experiences with digital aerial photogrammetry software 'e-foto' in an educational setting in Bahir Dar, Ethiopia

Martin Vermeer[1] and Getachew Tesfaye Ayehu[2]

[1]Department of Real Estate, Planning and Geoinformatics, Aalto University, Espoo, Finland
[2]Institute of Land Administration, Bahir Dar University, Bahir Dar, Ethiopia

---

Abstract: The digital photogrammetry software e-foto was developed at the State University of Rio de Janeiro for teaching aerial photogrammetry using standard personal computers. At the University of Bahir Dar, Ethiopia, in 2013 an intensive course took place completed by 15 students. In the course, the complete tool chain was exercised using Intergraph's ERDAS Imagine LPS software, on imagery from Finland and Brazil. However, the less complex e-foto software was also successfully used and found easier to master. E-foto works best on Linux. We were inspired by our teaching experience to build a tailored version of 64-bit Ubuntu booting from a memory stick, to easily allow the educational use – including self-study – of e-foto in this fashion.

Keywords: Digital aerial photogrammetry, e-foto, Open Source, education, Ethiopia

---

1   Why digital photogrammetry?

Using aerial photogrammetry for constructing orthophoto maps is mission critical for developing a cadastral system in many developing economies including Ethiopia (Hailu & Harris 2013). It is the only technology offering the geodetic mapping precision required especially in urban and peri-urban areas. The Ethiopian ministry of Urban Development, Housing and Construction (MoUDHC) is assisting regional states and urban centers in carrying out the adjudication and registration of landholdings. As a result MoUDHC is preparing orthophoto base maps and offers tailor-made training in digital photogrammetry to experts in close collaboration with BDU ILA and the Ethiopian Mapping Agency (EMA).

Figure 1: The computer classroom in Bahir Dar.

2 Description of e-foto

The digital photogrammetry software e-foto (Brito et al. 2011) was developed at the State University of Rio de Janeiro for teaching aerial photogrammetry using personal computers, without the expence and licence hassle of commercial photogrammetric software packages. It offers a largely complete tool chain from camera calibration and interior orientation over exterior orientation to aerotriangulation and feature extraction, and a less complete one for digital terrain model construction and orthorectification aimed at the construction of orthophoto maps. The package is already well sufficient for educational use.

3 Training data used

The aerial images used in the training were three images from Rio de Janeiro provided on the e-foto web site, and five very high quality colour images digitized at high resolution (22,000 x 22,000 pixels) from the Ringroad II project around Helsinki, Finland. Camera calibration documents and ground control point descriptions for both image sets were available. The GCPs and their descriptions were quite different for the two cases, a useful learning experience on professional practices in different countries.

4 Bahir Dar master's course experiences

At Bahir Dar University an intensive course was organized in October 2013, which 15 master's students from the Institute of Land Administration completed, see Figure 1. In the course, the complete digital photogrammetry tool chain was exercised based on the ERDAS Imagine LPS ('Leica Photogrammetry Suite') software from Intergraph. However, the e-foto software was also used as an introductory learning experience.

In the half year preceding the intensive course, the course instructor (MV) trained the use of both softwares, completing the processing chain in both. He even submitted one bug report with code patch back to Rio. It would however be rash to call him an expert, especially on the more complex ERDAS/LPS.

Functionalities used by the students were Interior and Exterior Orientation ('Spatial Resection'), aerotriangulation, and Feature Extraction. Generally this appeared to be a good learning experience. ERDAS is more versatile, which benefits the experienced user but gets in the way for a novice.

These students had already a good background in Geographic Information System (GIS) software and operations, but photogrammetry is different. The very precise measurement in the images of fiducial marks and ground control points, followed by a least-squares adjustment that mercilessly exposes any operator error, was a new experience for them. As was the extreme precision, in three dimensions, that the technology uniquely affords.

In this two-weeks full-time course the students completed a full processing chain in ERDAS/LPS, from the five Finnish images given to the end product, an orthorectified photomosaic of the area, ready for mapmaking use in any GIS product. They also completed a partial processing chain in e-foto for this imagery, doing Interior and Exterior Orientation and aerotriangulation. The Feature Extraction, Digital Elevation Model generation and orthorectification modules were tried out individually on the Rio de Janeiro imagery.

5   Experience details and comparisons

The reason for the above limitation in use of e-foto is that the relevant functionalities were not mature enough for practical use. These modules, Feature Extraction, DEM Extraction and Orthorectification, have working core functionality, but do not provide a way to export/import data files in compatible formats, making it impossible to build a processing chain. Also they are less integrated within e-foto, not offering the same processing-chain management (*i.e.,* keeping track of which operations were completed on which images) that ERDAS provides also for these functionalities.

We did not receive any extensive comparative comments from students. One sentiment expressed by students was that it was useful to have been made aware of the existence of e-foto. Students appreciated e-foto's ease of installation, learning and use by beginners, while also understanding its more limited capabilities.

We noticed one problem with ERDAS/LPS happening several times: camera calibration data entered somehow did not make its way to later stages of processing, undoubtedly due to operator error. In e-foto this cannot really happen as there is only one camera per project and it applies to all images.

It may be noted that ERDAS and LPS started out as two separate pieces of software by different companies; even after integration the "seamline" remains visible. E-foto has a clearer design, but its photogrammetric roots show in the photogrammetric measurement and adjustment modules being mature and production-ready, while the DEM extraction and orthorectification modules again have a "demo feel" over them. Feature Extraction, e.g., exports a flat text file, where geoinformatics folks might expect shape files. All these will require work especially on the data I/O interfaces to become production ready.

Of course ERDAS/LPS is more sophisticated. E.g., it can automatically generate tie points, and in Interior Orientation it predicts the location in the image of the next fiducial point after a couple have been measured, saving operator time. This does however require the operator to correctly identify the fiducial point numbering scheme used, causing some puzzlement with our students.

OGRS 2014

6 The future of e-foto, and our future plans

E-foto has already found its place in the teaching of digital aerial photogrammetry at low cost and free from licencing restrictions. Weaknesses, in addition to those mentioned, include paucity of high-quality instruction materials in English. This prompted the writing of our own lecture notes for the course, and planning of a textbook.

Open Source software development typically addresses the 'itches' of developers, which so far have been traditional aerial photogrammetrists from Brazil. It would take little development effort to also make the later modules in the processing chain, leading up to orthophoto mosaicing, mature enough for production use. Most of that work would involve better integration within the e-foto framework and with the rest of the map-making world.

We are experimenting with booting a tailored version of 64-bits Ubuntu from a USB stick ('memory stick') to allow educational use, including self-study, of e-foto on any 64-bits work station or laptop that allows such booting. These sticks nowadays come at large capacities – 8 GB and over – at reasonable prices. Software exists for Linux – *remastersys* and *unetbootin* – to create such a booting stick with a 'persistent' area for user data. There we store the aerial imagery, calibration documents, ground control point data files, *etc.*

One must use a USB stick with a fast enough data rate. We found that all e-foto modules and their imagery loaded within some tens of seconds even with a 'cheapo' stick, and after that, all operations are in memory and fast. The completely soundless loading of gigabytes of stuff is eerie to experience.

A digital aerial photogrammetry textbook is in early stages of preparation. We have permission to use the Finnish Ringroad II imagery from the Municipality of Espoo for educational purposes without charge. Include the e-foto bootable USB stick and red-cyan anyglyphic glasses with it, and the laptop-owning reader could start self-study immediately!

Acknowledgment

References

Brito, J. L. N. S., Aguiar, R. A., Silveira, M. T., Coelho Filho, L. C. T., Badolato, I. d. S., Pupim, P. A. B., Reolon, P. F., Ribeiro, J. A., Ribeiro, J., Bernardo Filho, O. & Mota, G. L. A. (2011), E-FOTO: Development of an Open-Source Educational Digital Photogrammetric Workstation, Barcelona, Spain, pp. 356–361.
URL: *http://tinyurl.com/brito-et-al*

Hailu, Z. & Harris, D. (2013), How to Map 50,000,000 rural land parcels in Ethiopia, *in* 'Annual World Bank Conference on Land and Poverty', Washington DC.
URL: *http://tinyurl.com/hailu-harris*

Innovative early warning and risk management system for flooding risk

Massimiliano Cannata and Milan Antonovic

Institute of Earth Sciences, SUPSI, Campus Trevano, 6952 Canobbio, Switzerland

Abstract: Several studies on predict an increase in high-intensity precipitations in Europe, which is likely increase the frequency and magnitude of flood events. Early warning systems (EWS) are used for predicting, communicating, and responding to floods. In this paper we propose a system architecture for a new type of EWS, and describe its implementation in a real case application in lake Verbano flood risk management. An existing EWS was renewed to meet new requirements. The new system design was based on the Service Oriented Architecture pattern. Open source software was used throughout the system. Open source approach was found desirable for several reasons: cost reduction, independence from proprietary policies, the right to access and modify the source code, and because of the good support provided by the open source software community.

Keywords: Early warning, open standards, sensor observation service

1 Introduction

Several studies on climate changes have predicted an increase in high-intensity precipitations over Europe (Lendernik & Meijgarrd 2008). This trend has been actually already observed in the Swiss Alps, particularly in fall and winter (Schmidli & Frei 2005). As a result, these changes are likely to impact the security of population located nearby flood prone areas, which may expect an increase in frequency and magnitude of events.

The risk management is therefore a key factor in guaranteeing a sustainable development and a secure society. The risk, which is the product of hazard, exposition and vulnerability, may be "controlled" through -the application of:

- planning strategies that reduce the exposition or vulnerability by regulations and constraints;
- structural interventions that physically remove the causes of the hazard or protect the exposed elements;
- early warning monitoring systems that reduce the exposition by timely intervention actions and emergency plan actuation.

Early Warning Systems (EWS) can be broken down in four components that are the monitoring, the predicting, the communicating and the responding. According to Molinari et al. (2013) latest European researches have highlighted a new type of EWS that is going to become the practice in the near future and that may be referred to as Innovative Early Warning System (IEWS). Such a kind of systems has the capability of:

- using wireless sensor networks characterized by dense, interconnected and low cost sensors for monitoring;
- producing dynamic prediction of the impending hazard based on real-time integration and modelling of sensor measures, contextual data and phenomenological properties for predicting;
- dispatching alerts using also non-standard communication channels like social networks and smart devices applications for communicating;
- involving citizens in emergency actions throughout crowdsourcing practices like community mapping or translation of emergency communications for responding.

In this paper the authors propose a system architecture for IEWSs based on the above described features and its implementation in the real case application of the Verbano Lake flooding risk management.

2   The early warning and management system

After yet another flooding of the Verbano Lake in 2000, which not caused fatalities but generated huge economic losses, the Locarno and Vallemaggia Civil Protection Agency (Ticino, Switzerland) decided to create a tool capable of supporting the risk management through an EWS.

The Institute of Earth Sciences of the University of Applied Sciences and Arts of Southern Switzerland (IST-SUPSI) between 2000 – 2002 set up a Geographical Web based application named SIT-GAP (Italian acronym of Geographical Information System for the Management Alert and Planning of interventions) capable to fulfil the civil protection agency capability requirements. The major features individuated and implemented were:

- visualization of information (cadastral, population, exposed elements, critical infrastructures, etc.) on cartographic basic data at different scales;

www.ogrs-community.org

- availability of updated information (location, resources and staff) related to emergency leading agencies like Municipalities, Police, Firefighters and Red Cross;
- visualization of lake boundaries based on current lake level and hourly forecasting levels;
- dynamic identification and localization of exposed elements based on lake level,
- triggering of alarms and visualization of planned activities;
- modification or creation of new action plans during intervention phase;
- management of information on evacuated population;
- dispatching of warnings and impending hazard through public access of base informative Web interface with maps.

This system was implemented on Windows NT4 operating system, based on Oracle8i database and ArcIMS web services and Neapoljs web mapping framework (http://www.mapserver.ch/en/samples.htm). Unfortunately, the mix of these technologies today causes several issues: in fact, ArcIMS is no longer supported, the whole system is no more compatible with modern browsers and is not portable on modern devices; moreover, the data are not editable on-line but need manual manipulation and a following publication for their update.

For these reasons, SIT-GAP required a renewal to keep pace with up-to-date technologies and instruments, but also to meet the new concept of IEWS and to integrate the standards promoted by current legislations: SITGAP 2.0 was thus designed and engineered.

The new system design is based on Service Oriented Architecture (SOA) pattern (Rotem-Gal-Oz et al. 2012), where different functionalities are provided as a service to other applications. This architecture was selected because has the advantage to be independent of any vendor, product or technology.

The whole system was decided to follow the key-concepts of interoperability and open-source. The interoperability, strong encouraged at European level by INSPIRE Directive (Directive 2007/2/EC), has as objective the adoption of standard data, protocols and procedures in order to obtain the most correct, accurate and comprehensive information for decision-making support. In practice, this leads to the system capabilities to access institutional or third party services which provide updated data despite the used software (vendor or version) or the original data format.

The open-source approach is desirable for several reasons: in addition to reducing costs by eliminating license fees, it is independent from proprietary policies and thus ensures long term project sustainability. Moreover, the right to access and modify the source code allows for upgrades that guarantee the compatibility with latest technologies. Finally, the development process of open source software generally involves a community of developers that provides prompt bugs fix and rapid release.

OGRS 2014

As illustrated in Figure 1, SITGAP 2.0 is based on OGC Open Standards (WMS, WFS, SOS, WPS), specific ad hoc SITGAP RESTful web APIs and social network APIs.
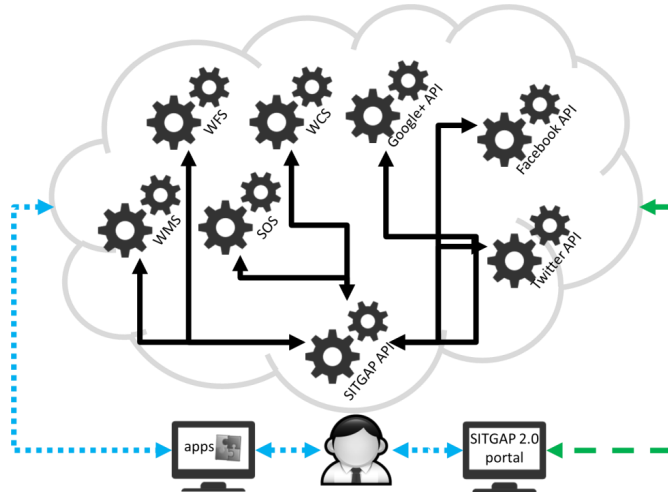

Figure 1: SITGAP 2.0 Service Oriented Architecture design and adopted open services

The SITGAP RESTful web API (Richardson & Ruby 2007) are developed in Python language (Van Rossum 2003) and rely on Tornado Web Server (www.tornadoweb.org) which is capable to handle a large number of clients at the same time. The requests and responses use the JSON open standard format. The data store is the PostgreSQL (Worsley & Drake 2002) with the PostGIS (Obe & Hsu 2011) extension.

The interface was developed using Sencha Ext JS and orchestrates the services to provide the desired business logic of the application. It integrates hydro-meteorological data concerning the areas surrounding the Verbano basin, it runs a conceptual hydrological model that combines the collected data and high-resolution meteorological forecasting to forecast lake level, and dispatches alerts to local stakeholder.

Hydro-meteorological data are collected and managed in real-time thanks to the usage of the SOS standard by means of two istSOS platform which serve more than 240 sensors and 30 million of observations coming from the monitoring networks of Canton Ticino, MetwoSwiss, Swiss Federal Environmental Office, Piemonte Region Environmental Agency and Lombardy Region Great Lake Consortium.

The hydrological model is a SCS-CN-based conceptual model that applies the Nash method (Maidment & Mays 1988). It is activated in case of meteorological alert only. Every model run, after a first phase of calibration based on measured observations collected from the istSOS services, it provides 72 hours lake level forecasts by using COSMO-7 meteorological model rainfall outputs by MeteoSwiss
www.ogrs-community.org

(http://www.meteoswiss.admin.ch/web/en/weather/models/cosmo.html) and elaborated thanks to a local WPS service.

SITGAP2 presents three operational modes: when no floods events are expected the normal mode provides information about emergency institutions and contact points; when floods are likely expected the alarm level gives to stakeholders information about the lake level forecasted, the possible flooded areas and the exposed elements. Finally, if severe floods are occurring, evacuation level supplies tools for managing operations such as registrations and tracking of evacuated persons.
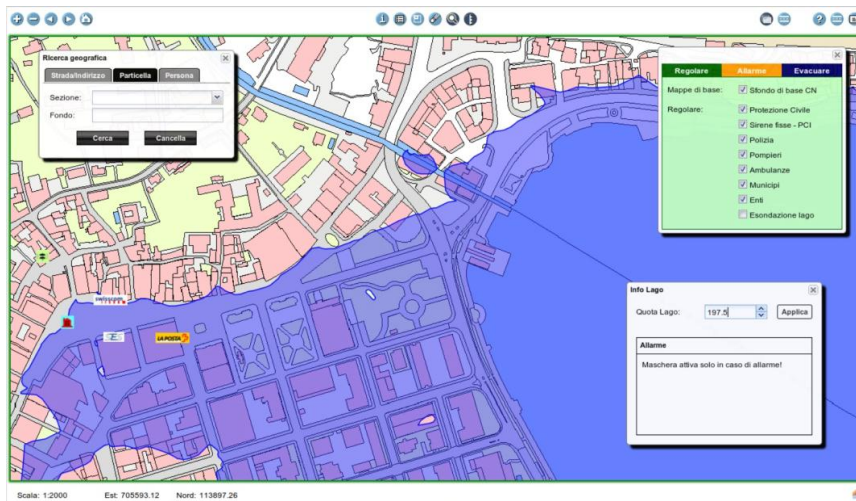
Figure 2: SITGAP 2.0 portal snapshot

Acknowledgments

References

Lendernik, G. & Meijgarrd, E. (2008), Increase in hourly precipitation extremes beyond expectations from temperature changes, Nature Geoscience, 1, doi: 10.1038/ngeo262

Molinari, M.E., Cannata, M. & Meisina, C. (2013), r.massmov: an open-source landslide model for dynamic early warning systems. Natural Hazards. Advance online publication. doi: 10.1007/s11069-013-0867-8

Obe, R. & Hsu, L. (2011), PostGIS in Action. Manning Publications, Stamford, 520 p.

Richardson, L. & Ruby, S. (2007), RESTful Web Services. O'Reilly Media, Sebastopol, 419 p.

Rotem-Gal-Oz, A., Bruno, E. & Dahan, U. (2012), SOA patterns. Manning.

Schmidli, J. & Frei, C. (2005), Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century. Int J Climatol 25(6):753–771

Van Rossum, G. (2003), The Python Language Reference Manual. Network Theory Ltd, Bristol, 144 p.

Maidment D.R. & Mays, L.W. (1988), Applied Hydrology. Tata McGraw-Hill Education,Worsley JC, 277 p.

Drake, J.D. (2002), Practical PostgreSQL: a hardened, robust, open source database. O'Reilly Media, Sebastopol, 640 p.

A toolkit for identifying and measuring physical streams features from high resolution DEMs and digital image

Corrado Cencetti, Pierluigi De Rosa, and Andrea Fredduzzi

Department of Physics and Geology, University of Perugia, Perugia 06100 Italy

Abstract: In fluvial dynamics remote sensing using airborne LiDAR (Light Detection and Ranging) may provide an alternative basis for identifying and measuring physical-geomorphological stream features that are traditionally recorded by field surveys. The ability to remote sense bankfull elevations was of particular interest in this study because bankfull mapping depends on topographic indicator. Bankfull is defined as the elevation at which flow spills onto the floodplain. In order to objectively define the bankfull elevation, we plotted the hydraulic depth as function of flow height for each cross section. The local maxima values indicate a sudden increase in flow width where water spills across floodplain. Such approach have been used to create a GIS tool that runs as plugin in QGIS software and provides as result a series of river indices (sinuosity index, braiding index, bankfull channel width, formative discharge) very helpful for fluvial geomorphologists and the polygon map of the bankfull limits. The algorithm was applied on a reachs of River Paglia an Umbrian rivers (in central Italy) affected by a strong disequilibrium.

Keywords: Fluvial dynamics, Bankfull Elevation, Airborne Lidar, hydraulic depth.

1   Introduction

In fluvial geomorphology there are some important physical features able to define the state of equilibrium of a river: such features can be obtained from field surveys, photointerpretation or both methods.

In fact, over the past centuries and especially in the last 50-60 years, the morphology and dynamics of most Italian rivers have suffered huge changes, mainly due to various anthropic interventions (eg construction of dams, removal of sediment from riverbeds, hydraulic works in riverbeds, changes in land use). This actions have
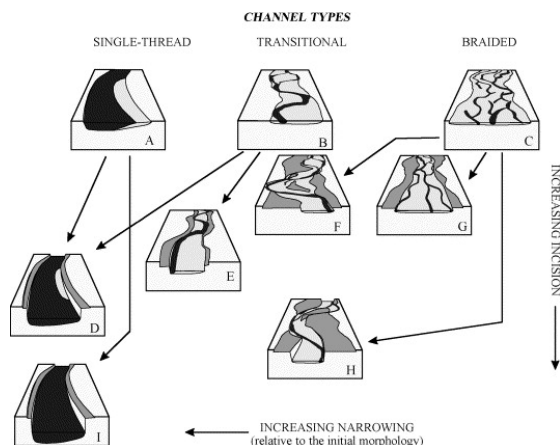
Figure 1: Classification scheme of channel adjustments for Italian rivers. Starting from three initial morphologies (A, B and C), different channel adjustments take place due to variable degrees of incision and narrowing (from Surian & Rinaldi (2003))

changed the liquid and solid flow regime, as well as having influenced other aspects, such as lateral mobility. The magnitude of changes of the riverbeds was considerable and in many cases has led to a morphological transformation, ie a modification of the planar configuration (for example from a braided channels to a wandering channel or even single channel, see Figure 1).

The physical features normally examined are based on indices able to read and quantify such variations; a list, non exhaustive, of them should contain at least sinuosity index, braiding index, bankfull channel width. The identification of planar indices listed above, however, is even directly related to the evaluation of an "*representative hydraulic height*". Consider, for example, the fact that varying the hydraulic height (and thus the flow discharge) varies consequentially the width of the channel.

In the morphological-sedimentary study of a water course is useful to define a level (and a corresponding flow discharge) "representative" (and responsible) about the shape and size of river channel. River geomorphologists have actually been struggling with the issue to investigate what flows are responsible for forming the channel for natural rivers.

The conventional view on the nature of formative discharge in rivers is that, although flow varies widely, a much more limited range of flows actually does most of the work in shaping a river channel. The notion here is that, for much of the year the lowest flows, although on a daily basis occurring more frequently than larger flows (note Figure 2), are simply too low to effectively erode and shape the channel. These low discharges the river simply flows through a channel that has been shaped by higher flows. But the highest discharges, although they are the most capable of eroding the channel, they occur so infrequently that their morphological impact is small. At the very highest discharges the flow will exceed the capacity of the channel and be conducted outside the channel and across the adjacent flood plain, thus becoming completely ineffective at shaping the channel itself. It is reasoned, therefore, that there is some much smaller middling range of flows that do most of the work shaping the
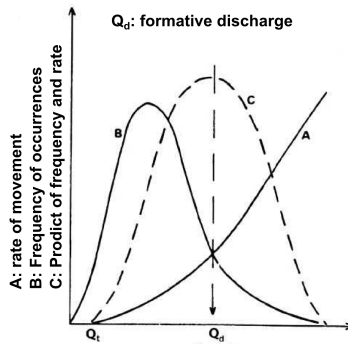
Figure 2: Notion of formative discharge as proposed by Wolman & Miller (1960)

river channel and that some summary value of these intermediate flows represents the formative discharge of the river Benson & Thomas (1966). Leopold et al. (1964) stated also that the dominant or formative discharge in an equilibrium channel can often be approximated by *bankfull discharge*: the discharge that completely fills the channel. The same authors say "in many rivers the bankfull discharge is one that has a recurrence interval of about 1.5 years".

Remote sensing using discrete return, near-infrared, airborne LiDAR (Light Detection and Ranging) and high-resolution digital imagery may provide an alternative basis for identifying and measuring physical-geomorphological stream features that are traditionally recorded by field surveys with a considerable monitoring effort.

The paper presents a toolkit, implemented as Quantum GIS plugin (QGIS) (QGis DT 2011), a desktop free and open source GIS that has rapidly gained popularity in the recent period, able to compute specific stream indicators and to identify some stream features directly from high resolution DEMs and aerial orthophoto. This software allows the user to derive the required informations minimizing, or even excluding, the field surveys, normally very expensive in terms of cost and time, which may be necessary to confirm the values of the indicators. The streams indicators here presented are: sinuosity index; braiding index; longitudinal stream profile; bankfull channel width; formative discharge. The plugin produces x-y scatterplots where each indicator is plotted against the distance from river head. Such diagrams provide many useful information to the fluvial geomorphologists as they investigate the river equilibrium, and identify problems related to the physical degradation.

2   Metodology

The main input data required for the river indices is represented by the stream centerline that can be obtained from high resolution digital imagery by opportune digitizing procedures or from a DTM using a GIS procedure for the extraction of drainage network (O'Callaghan & Mark 1984, Böhner et al. 2006) and eventually improve the result by a comparison to the digital aerial photo.

The toolkit, as first, calculate the sinuosity index defined as the ratio between the length measured along the stream centerline and the straight line distance (Brice 1964)
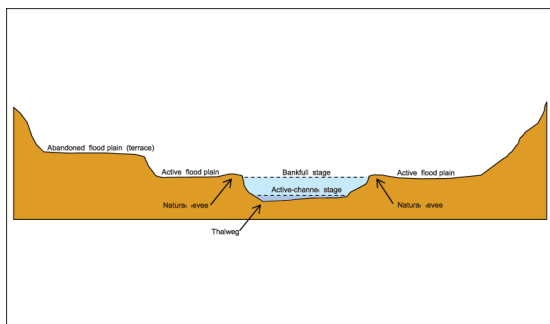
Figure 3: Schematic diagram showing geomorphic features of a stream channel (from Sherwood & Huitger 2005).

and trace the cross section lines orthogonally to the centerline with a step and a width defined by the user.

The next fundamental step in procedure is the definition, for each cross section, of the bankfull elevation end the elevation of first terrace (see Figure 3)

2.1 The bankfull elevation definition

The ability to compute bankfull elevations from remotely sensed data was of particular interest in this study because bankfull mapping depends on topographic indicator. Bankfull is defined as the elevation at which flow spills onto the floodplain (Leopold et al. 1964).

Bankfull locations are usually based on standard fields indicators (Dunne & Leopold 1978) given in order of reliance: break in bank slope corresponding with the active floodplain surface, high-flow markers (i.e., limit of bank scour, rock staining, sand/silt deposits, debris line), vegetation limits and bars tops.

In order to objectively define the bankfull elevation, we plotted the hydraulic depth (flow area divided by maximum width of a cross section) as function of flow height at each cross section. The local maxima values indicate a sudden increase in flow width where water spills across floodplain (i.e., bankfull - Faux et al. 2009). Such approach was suggested by McKean et al. (2005) as variant of the bankfull approach discussed by Williams (1978).

We recognized that the proposed method for identifying bankfull elevation would work best in floodplain rivers and would be limited in confined channels that lack floodplains. Moreover it worth mentioning that the detection of the bankfull elevation in field surveys, in case of confined channels, is not trivial.

The method of searching the local maxima of the hydraulic depth funcion is less definitive in channel with multiple terrace, where could be tricky decide which terrace and which plateaux in the hydraulic depth function correspond to the current bankfull elevation (Figure 4). Field observations can be used to guide the selection of the correct elevation corresponding to the active floodplain, but since the aim is to automatically derive from DTM LiDAR the bankfull elevation, and thus the bankfull limits, in case where multiple terraces are present we determined the elevation linked to the first terrace which may or may not be equivalent to the actual bankfull elevation.
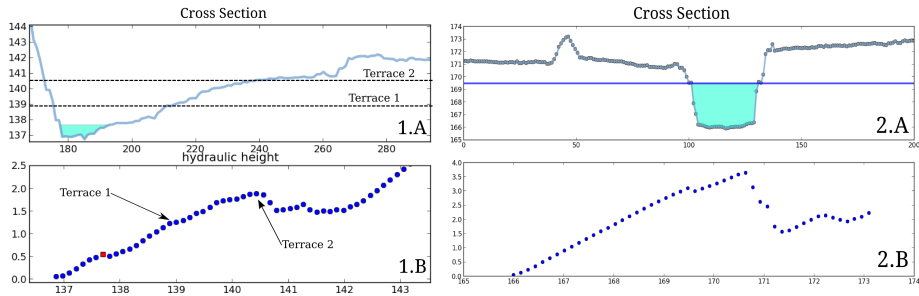
Figure 4: Cross section (1.A) shows two terrace and the relative example hydraulic depth (1.B) function. Dashed lines indicates terrace locations determined from (1.B). The red square in (1.B) indicates the actual bankfull elevation corresponding to the wetted area in (1.A) Cross section (2.A) shows a simpler case where not multiple terrace are present. The hydraulic depth function appears a have a more clear trend where the bankfull elevation extraction is more evident.

2.2 Calculation of indices

The tool implemented have been developed under Python environment (www.python.org) and it runs as plugin of Quantum GIS thanks to the possibility of extending the software potential by creating plugins in python language (Hugentobler et al. 2008).

A specific spline smoothing procedure have been applied in order to select the local maxima of the function hydraulic heights/water depths as could exists a big number of local maxima. The smoothing procedure should remove all local maxima coupled to a local "noise" effect. Particular attention was paid to the choice of the smoothing parameter *spar* that must be *small enough* to filter out only the local maxima due to small variations of the function. The *spar* parameter was selected through a process of 10 k-fold cross-validation. A *spar* parameters range is defined as values included between first standard deviation error related to the smoothing parameters having the minimum root mean square error. The best *spar* is the maximum value of all *spar* parameters included into the range limits above defined. The first local maxima found for the spline function with the best *spar* identify the bankfull water depth.

Once the bankfull depth is evaluated, for the cross section analyzed, the tool is able to calculate the number of channels, the bankfull channel width (corresponding to the channel with greatest wet area), and the discharge corresponding to the formative discharge. All these values are stored in the attribute table of the cross section lines. The analysis is repeated for all cross sections then the plugin produce the scatter plots where in y-axis are shown braiding index, the bankfull channel width, the formative discharge, the sinuosity index and in x-axis the distance from the stream source.

2.3 Case study: the River Paglia

The plugin has been tested on the Paglia River, a right tributary of the River Tiber (central Italy), a reach length of about 3 km. The Paglia River is a "wandering" channel characterized by an intensive fluvial mobility.

For the study area are available two high resolution DEMs: an airborne LiDAR

DEM (1 meter) referred to the 2008 and a high resolution DEM (2 meters) obtained by interpolation procedure referred to the 1964. The DEM resolution is an important discriminating as, of course, must be sufficiently small to reconstruct the cross-section river. For the case study the average channel width is always greater than 20 meters, therefore, it is suggested to use DEM with a ground resolution of at least 1/10 of the average channel width.

The river has suffered over the last 60 years, strong morphological modifications (Cencetti, De Rosa & Fredduzzi 2013*a*,*b*, Cencetti, De Rosa, Fredduzzi & Tacconi 2013) mainly due to the large removal of sediment from riverbeds, which has caused a shift from a braided morphology to the current wandering channel. The toolkit applied to the reach provide the fluvial morphological indices reported in section 2.2. The comparison between the mean bankfull channel width for all cross section lines is 160 m in 1954 and 51 m in 2008; the mean braided index is 1:32 in 1964 and 1:12 in 2008. Such values confirm how strong is the disequilibrium affecting the Paglia River. In fact the effects of this disequilibrium are evident in the study reach. In particular in field there are evidence of high banks frequently invested by local slides, channel armouring of riverbed sediments, substrate outcropping.

3   Conclusions

The study of channels morphological characteristics and of their degree of alteration is one of the basic cognitive activities for the proper management of fluvial system. High-resolution LiDAR data and derived DEMs present an opportunity for resource managers to expand the amount and quality of information available for river habitat monitoring. The availability of detailed topographic data is a fundamental requirement in this study.

The proposed toolkit represents a valid tool for assisting the fluvial geomorphologists the evaluate a series of river indices that normally are obtained through procedures of photointerpretation and field surveys with an important effort in time and money.

The application to the case study has provided a quantitative assessment of fluvial indices that can demonstrate the strong disequilibrium affecting the River Paglia.

Acknowledgment

References

Benson, M. & Thomas, D. (1966), 'A definition of dominant discharge', *Hydrological Sciences Journal* **11**(2), 76–80.

Böhner, J., McCloy, K. & Strobl, J. (2006), 'SAGA: Analysis and Modelling Applications.', *Göttinger Geographische Abhandlungen* **115**, 130p.

www.ogrs-community.org

Brice, J. C. (1964), *Channel patterns and terraces of the Loup Rivers in Nebraska*, US Government Printing Office.

Cencetti, C., De Rosa, P. & Fredduzzi, A. (2013*a*), 'Analysis of the evolution of a riverbed using CAESAR, a cellular automata model.', *Rendiconti online della Società Geologica Italiana* **24**, 49–51.

Cencetti, C., De Rosa, P. & Fredduzzi, A. (2013*b*), Evoluzione morfologico-sedimentaria dell'alveo del F. Paglia (bacino del F. Tevere) nella sua bassa valle., *in* 'Atti del Convegno 'Dialogo intorno al Paesaggio. Percezione, Interpretazione, Rappresentazione''.

Cencetti, C., De Rosa, P., Fredduzzi, A. & Tacconi, P. (2013), Studio sulla dinamica fluviale per la gestione morfo sedimentaria del sistema alveo-pianura fluviale del fiume Paglia, Technical report, Provincia di Terni.

Dunne, T. & Leopold, L. B. (1978), *Water in environmental planning*, Macmillan.

Faux, R., Buffington, J. M., Whitley, G., Lanigan, S. & Roper, B. (2009), 'Use of airborne near-infrared lidar for determining channel cross-section characteristics and monitoring aquatic habitat in pacific northwest rivers: a preliminary analysis', *Proceedings of the American Society for Photogrammetry and Remote Sensing. Pacific Northwest Aquatic Monitoring Partnership, Cook, Washington* pp. 43–60.

Hugentobler, M., Duster, H. & Sutton, T. (2008), 'Extending the functionality of qgis with python plugins'.

Leopold, L. B., Wolman, M. G. & Miller, J. P. (1964), *Fluvial processes in geomorphology*, Courier Dover Publications.

McKean, J., Isask, D. & Wright, W. (2005), 'Mapping channel morphology and stream habitat with a full wave-form green lidar', *EOS Transactions, American Geophysical Union* **86**(52). Fall Meeting, Suppl. Abstract H34B-05.

O'Callaghan, J. F. & Mark, D. M. (1984), 'The extraction of drainage networks from digital elevation data', *Computer Vision, Graphics, and Image Processing* **28**(3), 323 – 344.
URL: *http://www.sciencedirect.com/science/article/pii/S0734189X84800110*

QGis DT (2011), 'Quantum gis geographic information system', *Open Source Geospatial Foundation Project* .

Sherwood, J. & Huitger, C. (2005), Bankfull characteristics of Ohio streams and their relation to peak stream-flows., Scientific Investigations Reports 2005-5153, U.S. Geological Survey.

Surian, N. & Rinaldi, M. (2003), 'Morphological response to river engineering and management in alluvial channels in italy', *Geomorphology* **50**(4), 307 – 326.
URL: *http://www.sciencedirect.com/science/article/pii/S0169555X02002192*

Williams, G. P. (1978), 'Bank-full discharge of rivers', *Water resources research* **14**(6), 1141–1154.

Wolman, M. G. & Miller, J. P. (1960), 'Magnitude and frequency of forces in geomorphic processes', *The Journal of Geology* pp. 54–74.

Enhanced OGC:SWE – KIWI instrumentation nature-oriented observation suite

Mikołaj Dobski, Artur Jaworski, Norbert Meyer, Szymon Mueller, Tomasz Rajtar, Dominik Stokłosa, and Maciej Stroiński

Poznań Supercomputing and Networking Center, ul. Noskowskiego 10, 61-704 Poznań, Poland

Abstract: The Sensor Web Enablement suite of standards from Open Geospatial Consortium (OGC:SWE) is supported by several software and hardware solutions. The KIWI Team at the Poznań Supercomputing and Networking Center has over the last few years transitioned from developing its own custom tools to empowering both research and educational solutions with open-source OGC:SWE implementations. Starting with own proprietary sensor data acquisition system – communication protocols, sensor polling daemons and custom database based on Hibernate persistence layer – two notable success-story deployments were introduced: The Virtual Laboratory of Interactive Teaching and an ecological, phenology domain-oriented service within the PLGrid Plus project. Development of the latter triggered an attempt to reuse the formerly deployed software platform and entailed the need to generalize it. Assessment and evaluation of this task resulted in an alternative option – deployment of commonly accepted and verified OGC:SWE solution. By complying services with OGC:SWE standards we have introduced an open source, easy to deploy and use KIWI Instrumentation framework. Moreover by focusing on fauna and flora monitoring we have extended the OGC:SWE ontology with the concept of species. This added another filtering dimension for phenological queries performed on a OGC Sensor Observation Service.

Keywords: OGC Sensor Web Enablement, Phenological observations, Sensor data and blobs database, Remote Instrumentation, Workflows

1 Educational and research deployments success stories

The Virtual Laboratory of Interactive Teaching (Kaliszan et al. 2010, 2011, Stokłosa et al. 2012) – *WLIN* is an interdisciplinary e-learning platform. Up-to high school students can learn there elements of biology of plants and animals, phenology, meteorology, water conditions and environmental chemistry. Presented knowledge is not only

available from professional, prepared by teachers, lessons' materials but also from actual phenological observations, carried on-site. These consist of meteorological data acquired from weather station combined with water gauge deployed on local river and pictures of plants growing in the surrounding taken automatically on a scheduled basis.

Second deployment is an ecology domain oriented solution (PSNC 2013) within the PLGrid Plus project (Kitowski et al. 2012), focused on research and targeted at Polish Scientific Community. This project is based on the experience from virtual laboratories studies performed at PSNC (Meyer et al. 2004, Okon et al. 2006, Lawenda et al. 2004, Płóciennik et al. 2010). It consists of two novel services.

First is a fully automated fitophenological observations pipeline which archives continuous time-lapse pictures of plant species identified at the deployment site in Wielkopolska National Park. Such pictures, along with meteorological data measured on-site, are delivered to scientists. Conducted remote research allows them to accurately create full phenological spectra for those plants and observe long-term evolution changes over years via user-friendly portlets extending open-source Liferay portal platform (Akram et al. 2005).

Finally, a species preservation program has begun in Łopuchówko Forest Division, being part of the Polish State Forests National Forest Holding (Polish State Forests National Forest Holding 2013), which provides a 24/7 cold room monitoring platform. It ensures that environmental parameters (humidity, temperature and air composition) are maintained within required values in order to store the plants' seeds in perfect conditions. Monitored data are available for insight and analysis in the previously mentioned Web portal.

2 Integration with 52°North's SWE suite

KIWI revamped remote instrumentation platform was decided to be built upon a stack of acknowledged OGC:SWE (Botts et al. 2008) implementations, developed by the 52°North's Sensor Web community (Bröring et al. 2011).

2.1 OGC Sensor Observation Service facade

Acquired measurements and observations were to be stored in the Sensor Observation Service (Na et al. 2007, Bröring et al. 2010, McFerren et al. 2009). This software enabled our suite with reliable database storage and flexible data querying and filtering capabilities. Moreover this approach not only allowed us to pick from a variety of client implementations but also made our architecture coherent. We started by successfully migrating the WLIN observations database to a locally deployed 52°North SOS repository (Bröring et al. 2009). During this phase two important artifacts were created.

Firstly, a SOS complaint Java facade library wrapping the 52N:OX-Framework was created adding convenient interoperability API. It exposes SOS 1.0, 2.0 core profiles and partly SOS 1.0 transaction profile using OGC ontology standards to Java POJO model mapping, namely Observations and Measurements (Cox et al. 2011) and Sensor Model Language (Botts & Robin 2007).

Secondly, we embedded the KIWI SOS adapter within Spring Framework (Johnson et al. 2009) based Java servlet and deployed it next to the 52N-SOS service. Inverting

the client adapter library into a software as a service approach allowed us to isolate the client from the OGC standards' communication heavy XML bindings. With our extension the clients can now communicate with the SOS service using REST architecture and lightweight JSON data format. Additionally, we deliver SOS compliant Java client endpoints. The communication layer is done using open-source Apache CXF framework (Apache 2009).

2.2   OGC Sensor Planning Service facade

Access to deployed sensors is provided by 52°North's implementation of OGC Sensor Planning Service (Simonis 2007, Bröring et al. 2011) – *SPS*. Each type of sensor, ranging from simple temperature probe to a camera mounted on a pan-tilt turnplate, can be modeled and deployed within the SPS service. The framework allows scheduling and executing tasks on instruments, such as reading sensor's values or taking a picture using connected camera. Relevant SensorML documents describe sensors' capabilities, parameters and taskable actions. Customizable KIWI Instrument abstraction allows easy integration of new types of sensors within the SPS framework. We also provide special communication layer, KIWI SPS Adapter, hiding SPS XML bindings and providing Java API for managing sensor tasks.

Sensor Planning Service is tightly coupled with Sensor Observation Service – data gathered in result of certain SPS tasks, such as taking a picture by camera, is stored in SOS as an observation, enriched by execution context (camera settings, position).

3   52°North's Sensor Observation Service extensions

Special extensions to OGC SWE infrastructure have been provided in order to enrich the SOS standards with new operations and supplement 52°North's SOS profiles implementation.

3.1   ResultQuality management service

One major advantage of using 52°North's Sensor Observation Service is its compliance with optional Observations and Measurements section: result quality. This specification allows us to persist additional metadata related to particular observations as dictionaries of key-value pairs, retrievable from the database using standardized SOS-1.0 getObseravation operation.

KIWI Instrumentation Suite added value to this functionality is an extension built upon the 52°North's SOS service. It allows full CRUD operations on results' qualities – adding, modifying and removing. Moreover we implemented a querying API which filters results' qualities by keys' names, values or sets of key-value pairs. Finally, we are able to filter observations and retrieve only those meeting certain metadata criteria.

3.2   External blobs repository service and SOS missing operations implementation extensions

An external repository service supplements the SOS behaviour with handling observables which produce huge volumes of observation data. In our particular deployment these are digital pictures.

OGRS 2014

SOS O&M observations are automatically coupled with blob (binary large object) files provided by external services. Such binding allows propagating insertion and removal of data.

Our resultQuality CRUD extension, deployed next to the 52°North's SOS service provides two more notable features. First comes implementation of the missing SOS Transactional Profile deleteObservationById operation. This feature is fully integrated with the mentioned earlier External Repository extension. Together, combined modules, offer atomic deleteObservationById operation for blob observations. An observation is removed from SOS only once the external repository confirms binary data removal. Otherwise the operation fails execution with relevant error returned to the client.

The External Repository integration is however fully optional. DeleteObservationById operation only triggers an external repository services when an adequate entry is found in the SensorML description of procedure which generated the observation tasked for removal.

3.3   Numerical measurements aggregation – time series

Next we introduced another filtering dimension to the standardized SOS getObservations query – namely numerical measurements aggregation.

A client selects whether he wants raw measured numerical data read from sensors and stored in SOS or their on-demand computed aggregation. The available aggregation periods are as follows: day, decade, month, season, year and a phenology oriented "vegetative period" – April till October. For such temporal aggregation the client needs to choose from of the following statistical aggregation models: average (arithmetic mean), minimum, maximum or sum.

Adding those two filters plainly into the standardized query would effect in a Cartesian product. Therefore we had to introduce a modified one. The client provides a set of query filters – temporal filter, procedure, phenomenon and aggregation mode for each requested time series.

Our extension simplifies long-term meteorological analysis by filtering the dense and frequent (e.g. 10 minutes period) readings which may rather obscure the big picture. The user is presented with accurate and easy to read and interpret aggregated time series.

4   Nature–oriented novelties

Finally, KIWI Instrumentation Suite offers few unique tools which focus on nature oriented observations. Dedicated to technically support phenological research we introduced two novelties: the concepts of scene and species. By definition, phenological research, in order to discover phenological spectra for species in a particular habitat, requires two sources of information. These are standardized, measured weather conditions and repeatedly taken pictures of the same topic (meaning plant) over a long term (up to few years). To acquire such data we have developed a Pan-Tilt-Zoom, Digital Single Lens Reflex camera set. It is accessible via remote services with live view support which as a remote instrumentation software bundle is also available as open-source.

Researchers using Web portlets define and persist entities of species they are inter-

ested in observing. Species model consist of name, description and scenes in which they appear – name and pan-tilt coordinates. Moreover, scenes contain a list of sub-presets, extending their definition with lens focal length and focusing distance. Geolocation is resolved from camera SensorML document. Such dataset allows explicit plant identification as each persisted SOS observation might point to a particular scene using result quality entry.

These newly introduced entities are available for enumeration via an extension to standardized SOS adapter mentioned before. The researcher, using our deployed suite, is able to easily define workflows consisting of scheduled meteorological data acquisition along with pictures taken for defined set of scenes.

KIWI Phenological SOS extension introduces an additional observations filtering capability besides the standardized criteria defined in OGC:SOS by adding a list of chosen scenes into account.

5 Future work

Future work on KIWI Instrumentation Suite is focused on extending sensor side services, namely introducing Sensor Event Service (Jirka & Nuest 2010, Bröring et al. 2011) for standardized notifications mechanism. We also consider implementing OGC Sensor Instance Registry (Echterhoff & Everding 2008) for semi-automated sensor system configuration.

Finally we intend to migrate our suite to the recently introduced new, revamped implementation of Sensor Observation Service (52°North 2014) as well as to comply with the reference SPS 2.0.1 service (52°North 2013).

Acknowledgments

References

52°North (2013), 'SPS 2.0.1 official OGC reference implementation'. Accessed 2013/10/10.
URL: *http://52north.org/news/sps-2-0-1-official-ogc-reference-implementation*

52°North (2014), '52°North SOS 4.0.0 Final Release'. Accessed 2014/03/20.
URL: *http://blog.52north.org/2014/03/20/52north-sos-4-0-0-final-release/*

Akram, A., Chohan, D., Wang, X. D., Yang, X. & Allan, R. (2005), 'A service oriented architecture for portals using portlets', *UK e-Science AHM* .

Apache, C. (2009), 'An Open Source Service Framework', *See: http://cxf. apache. org* p. 111.

Botts, M., Percivall, G., Reed, C. & Davidson, J. (2008), OGC® sensor web enablement: Overview and high level architecture, *in* 'GeoSensor networks', Springer, pp. 175–190.

Botts, M. & Robin, A. (2007), 'OpenGIS sensor model language (SensorML) implementation specification', *OpenGIS Implementation Specification OGC* pp. 07–000.

Bröring, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., Liang, S. & Lemmens, R. (2011), 'New generation sensor web enablement', *Sensors* **11**(3), 2652–2699.

Bröring, A., Jürrens, E. H., Jirka, S. & Stasch, C. (2009), Development of sensor web applications with open source software, *in* 'First Open Source GIS UK Conference (OSGIS 2009)', Vol. 22.

Bröring, A., Stasch, C. & Echterhoff, J. (2010), 'OGC Interface Standard 10-037: SOS 2.0 Interface Standard', *Open Geospatial Consortium* .

Cox, S. et al. (2011), 'Observations and measurements-XML implementation', *Open Geospatial Consortium* .

Echterhoff, J. & Everding, T. (2008), 'OGC Discussion Paper 08-133: OpenGIS Sensor Event Service Interface Specification', *Open Geospatial Consortium* .

Jirka, S. & Nuest, D. (2010), 'OGC Discussion Paper 10-171: Sensor Instance Registry', *Open Geospatial Consortium: Wayland, MA, USA* .

Johnson, R., Hoeller, J., Arendsen, A. & Thomas, R. (2009), *Professional Java Development with the Spring Framework*, John Wiley & Sons.

Kaliszan, D., Koczorowski, F., Mazurek, C., Meyer, N., Procyk, M., Rajtar, T., Stoklosa, D. & Stroinski, M. (2010), New technologies in nature science teaching—Virtual Laboratory of Interactive Teaching, *in* 'eChallenges, 2010', IEEE, pp. 1–7.

Kaliszan, D., Koczorowski, F., Procyk, M., Stokłosa, D., Rajtar, T., Mazurek, C., Meyer, N. & Stroiński, M. (2011), 'VIRTUAL LABORATORY OF INTERACTIVE TEACHING–A LIVE LABORATORY FOR ENVIRONMENTAL SCIENCE EDUCATION', *ICERI2011 Proceedings* pp. 2214–2222.

Kitowski, J., Dutka, L., Mosurska, Z., Pajak, R., Sterzel, M. & Szepieniec, T. (2012), Plgrid plus: Toward domain-specific infrastructure for supporting international research collaboration, *in* 'Proceedings of the EGI Community Forum 2012/EMI Second Technical Conference (EGICF12-EMITC2). 26-30 March, 2012. Munich, Germany. Published online at http://pos. sissa. it/cgi-bin/reader/conf. cgi? confid= 162, id. 147', Vol. 1, p. 147.

Lawenda, M., Meyer, N., Rajtar, T., Okon, M., Stoklosa, D., Kaliszan, D., Kupczyk, M. & Stroinski, M. (2004), Workflow with Dynamic Measurement Scenarios in the Virtual Laboratory, *in* '6th CARNet Users Conference, ISBN', pp. 953–6802.

McFerren, G., Hohls, D., Fleming, G. & Sutton, T. (2009), Evaluating sensor observation service implementations, *in* 'Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009', Vol. 5, IEEE, pp. V–363.

Meyer, N., Lawenda, M., Okon, M., Rajtar, T., Stoklosa, D. & Stroinski, M. (2004), Effective integration of scientific instruments in the Grid., *in* 'TERENA Networking

Conference'.

Na, A., Priest, M., Niedzwiadek, H. & Davidson, J. (2007), 'OGC Implementation Specification 06-009r6: Sensor Observation Service'.

Okon, M., Kaliszan, D., Lawenda, M., Stoklosa, D., Rajtar, T., Meyer, N. & Stroinski, M. (2006), Virtual Laboratory as a remote and interactive access to the scientific instrumentation embedded in Grid environment, *in* 'e-Science and Grid Computing, 2006. e-Science'06. Second IEEE International Conference on', IEEE, pp. 124–124.

Płóciennik, M., Meyer, N., Bylec, K., Mueller, S., Lawenda, M., Davoli, F., Cheptsov, A., Liabotis, I., Pugliese, R., Watzl, J. et al. (2010), 'Integration of the Laboratory Instruments with e-Infrastructure', *Computational Methods in Science and Technology* pp. 125–133.

Polish State Forests National Forest Holding (2013), 'Łopuchówko Forest Division'. Accessed 2013.
URL: *http://www.lopuchowko.poznan.lasy.gov.pl/*

PSNC (2013), 'Phenological Domain Grid'. Accessed 2013.
URL: *https://fenologia.plgrid.pl/o-projekcie*

Simonis, I. (2007), 'OGC Implementation Specification 07-014r3: OpenGIS Sensor Planning Service', *Open Geospatial Consortium* .

Stokłosa, D., Kaliszan, D., Rajtar, T., Meyer, N., Koczorowski, F., Procyk, M., Mazurek, C. & Stroiński, M. (2012), New Technologies in Environmental Science: Phenology Observations with the Kiwi Remote Instrumentation Platform, *in* 'Remote Instrumentation for eScience and Related Aspects', Springer, pp. 117–127.

Web-based analysis of storm-inflicted forest damages

Jaakko Kähkönen, Pekka Latvala, and Lassi Lehto

Department of Geoinformatics and Cartography, Finnish Geodetic Institute,
PL 15, FI-02430 Masala, Finland

Abstract: The occurrence of storms has been increasing due to the change in climate. There is a need for Web-based tools to rapidly assess storm-inflicted damages in forests. Fast development of UAV-based imaging and laser scanning technologies has made quick acquisition of the needed data sets possible. The recently introduced 3D visualization capabilities in the Web browsers facilitate the development of a new category of online spatial tools. Together with the immediate SDI-provided access to content and to spatial analysis processes, browser-based applications can be seen as a promising solution for the analysis of forest damages. In this paper an exemplary application, entirely based on open-source software components, is presented.

Keywords: 3D visualization, WebGL, WPS, SDI, OGC

1 Introduction

Climate change has notably increased the probability of extreme weather phenomena. Heavy storms are an example of this development, specifically in the tropic. According to climate studies, the number of hurricanes in the Northern Atlantic has steadily increased since 1970s. At the same time it is expected that routes of heavy storms are gradually shifting towards Polar Regions. Even in Finland this phenomenon has been evident over the last few years. Several storms have been recorded, specifically during the recent unprecedentedly mild winter seasons. These storms have inflicted serious damage in Finnish forests - leaving thousands of households without electricity and leveling millions of cubic meters of timber.

There is an urgent need for tools to help forest owners to quickly do an inventory of the storm-inflicted damages inside their properties. Recent developments in Unmanned Aerial Vehicle (UAV) -based technologies enable fast

capture of laser scanned point clouds and photogrammetric imagery of limited target areas. These after-storm data sets can be compared with the currently existing, very accurate national laser-scanned surface models for first analysis of the situation. Full potential of the scenario can be achieved, when brought into the context of a modern Web Service-based Spatial Data Infrastructure (SDI) with abundant data content resources and a rich set of spatial processing services.

The fast development of network technologies generally, and the standardization of the open access interfaces for spatial services in particular, have made map applications one of the standard functionalities available on the Web browser platform. Introduction of 3D visualization capabilities into the browser environment has fundamentally changed the way spatial information can be presented to the user. In this paper an exemplary open source-based implementation of the discussed use case is presented. The main principles of the solution include online access to the spatial data sets needed, the use of Web Service-based processing facilities, and 3D visualization of the results in a Web browser-based application.

2 Web browser platform

The new Web browser-based map applications generally provide good support for open content access interfaces. This has motivated several content providers to make their data sets available via the popular Open Geospatial Consortium (OGC) - developed service interfaces, like raster map services (Web Map Service, WMS and Web Map Tile Service, WMTS) and geospatial data services (Web Feature Service, WFS and Web Coverage Service, WCS) (OGC 2014). Gradually also service interfaces for data processing and analysis have gained popularity. The main standard for this kind of services is the OGC's Web Processing Service (WPS). These interface standards currently form the basis of the service-based Spatial Data Infrastructures (SDIs). Increased availability of free and open data sources further emphasizes the importance of standardized access interfaces as a consistent mechanism for the provision of geospatial content.

Also the Web browser-based visualization of spatial content has recently been in a rapid development phase. The traditional 2D map rendering processes are gradually being complemented with richer, more descriptive and intuitive interactive displays. One particular topic under intensive development is the 3D visualization of geospatial data in Web browsers, both on desktop and in hand-held devices, like smart phones. A particular aim in this development is to support 3D visualization on top of the plain browser platform, without any user-installed, application-specific plug-ins. The latest results of these developments open entirely new possibilities for visualization tasks in geospatial applications (Kim et al. 2014). One can envision, for instance, an application for evaluating the effects of rising water levels in a real 3D model of a flood-prone area. Another scenario is to estimate the effects of a winter storm inside a forest owner's property borders.

www.ogrs-community.org

3 Web Graphics Library

One of the main drivers for the growing popularity of the Web browser-based map applications has been the development in the mid 2000s of the Google Maps application, consisting of a JavaScript-based user interface, a service API and a rich set of spatial content. At the same time the open source community was also activated and several dynamic map applications and libraries were developed. The most used of these are currently the OpenLayers and Leaflet libraries.

Web Graphics Library (WebGL) is one of the most important new technologies facilitating the 3D visualization of geospatial data on the browser platform (as an example, see Wüest 2012). WebGL is a JavaScript API for the rendering of 3D graphics. WebGL is natively supported in most of the modern Web browsers. With WebGL it is possible to display interactive 3D graphics, if the graphics board and the related drivers on the host computer provide support for it. However, programming in WebGL is quite complicated. Several higher-level libraries have been developed to help the application developer in this task (Khronos 2014). One of these is the Three.js library (Three.js 2014). Three.js is a JavaScript library that makes the use of WebGL easier by offering to the developer a set of high-level graphic objects and concepts that simplify the creation of 3D scenes. Three.js renders into the HTML5 canvas element using either Scalable Vector Graphics (SVG) or WebGL. There is also a plug-in available for the popular QGIS application to create Three.js exports (QGIS Plug-ins, 2014).

4 Case implementation

The application environment of the case implementation described in this paper consists of an OpenLayers-based client interface for defining the target area of the process, a MapServer-provided WCS interface for accessing the surface models, a GeoServer-based server application with a custom-built Java WPS process and a PostgreSQL/PostGIS database. Various external OGC-compliant content access interfaces are utilized for retrieving the data sets needed (for a similar approach, see Rocatello et al. 2013). The final results of the process are presented as a 3D model in the Web browser making use of WebGL and Three.js libraries. All of the software components used in the case study are available as open source tools.

In this case study, the forest damages caused by a storm are investigated. This is done by computing the height differences between Airborne Laser Scanning (ALS) data collected before storm and photogrammetric imagery collected after the storm. The process is fully based on service-oriented computing environment. The source data sets are retrieved from content services via standardized access interfaces, the computations needed for the analysis are performed by a service side component accessible through a WPS interface, and the results are presented in a Web browser utilizing the new 3D visualization libraries.

Figure 1: The OpenLayers-based client application for selecting the area of interest for the process. The map user interface diplays orthoimagery and cadastral parcels, requested from the service of the national mapping agency

The client user interface contains an orthoimage layer and a cadastral parcel layer as application's background map layers (see Figure 1). The orthoimage layer is retrieved through the WMS interface and the cadastral parcel layer through the WFS interface from a service provided by the national mapping agency. The user can select a specific cadastral parcel from the map and execute the analysis process for that area in order to retrieve the storm-inflicted forest damages (Figure 2).

The analysis is performed by sending the selected cadastral parcel's polygon and a metric filtering parameter to the analysis process through the WPS interface. The metric filtering can be used for including only the significant changes to the result. The filter is also useful for practical reasons, as it speeds up the calculation process in case of large datasets, and reduces the size of output file. The Java-based WPS process is hosted in the GeoServer application, and it performs the height difference computation between the two surface models for the selected area. The height difference between the surfaces is computed using various PostGIS operations. First, the height difference values are calculated for each pixel pair and after that the pixels whose height difference value exceeds the given height filter parameter are outputted as polygons. The results of the calculation are returned in a GeoJSON form that is visualized in the client application.

In the result file, the differences in the two surface models that have occurred in the area between the two data capture moments are represented as 3D polygonal volumes and are calculated on-demand from user selected cadastral parcel. The identified damages (here simply the greater than 15 m losses in height) are visualized together with the original surface models and for instance the ortophoto as the texture in a 3D view of the area, thus providing a powerful tool for exploring the effects of the storm (Figures 3-5). The surface models that are used as

www.ogrs-community.org

background images in the 3D-visualization are requested from the MapServer's WCS interface in the binary ENVI format, (comparable with the approach in Pedersen 2014). Dynamic exploration of the storm effects is now possible by freely rotating and scaling the model using mouse-driven controls. By adding light sources into the model it would be possible to simulate for instance the changes the storm caused in the sun lightning conditions in the study area.
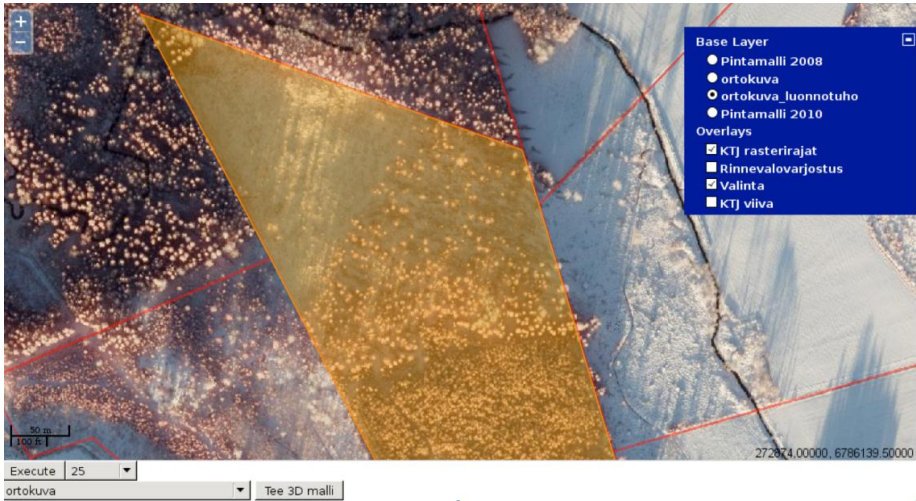

Figure 2: The application user interface showing a selected cadastral parcel. In this figure the background imagery shows the situation after the storm
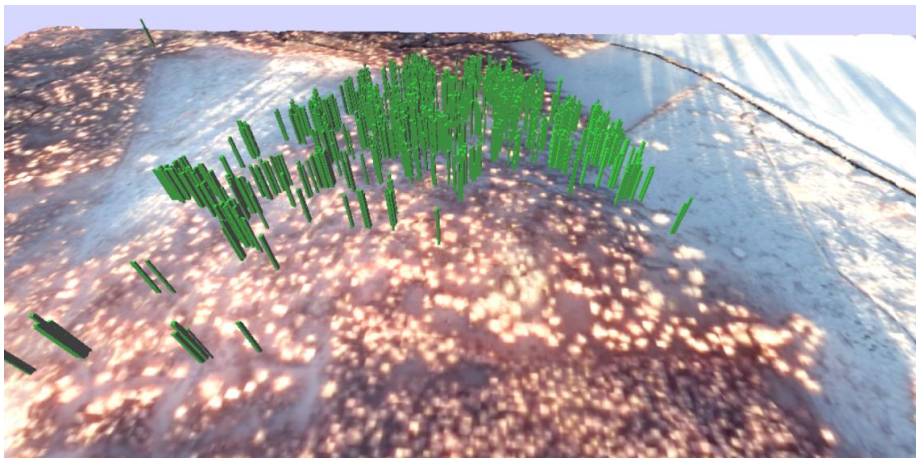

Figure 3: The results of the WPS process from the selected parcel, shown together with the digital terrain model (DTM), textured with the orthoimagery taken after the storm

OGRS 2014
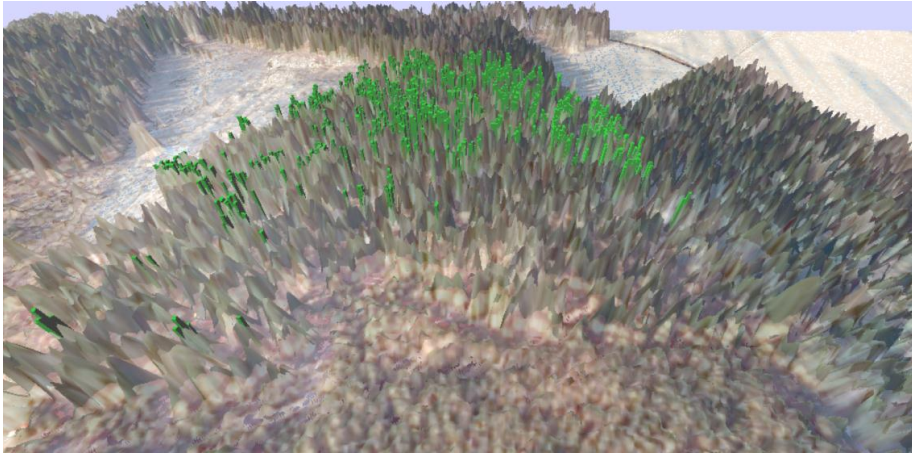
Figure 4: As in Figure 3, but now the digital surface model (DSM) before the storm, textured with the original orthoimage, is also displayed
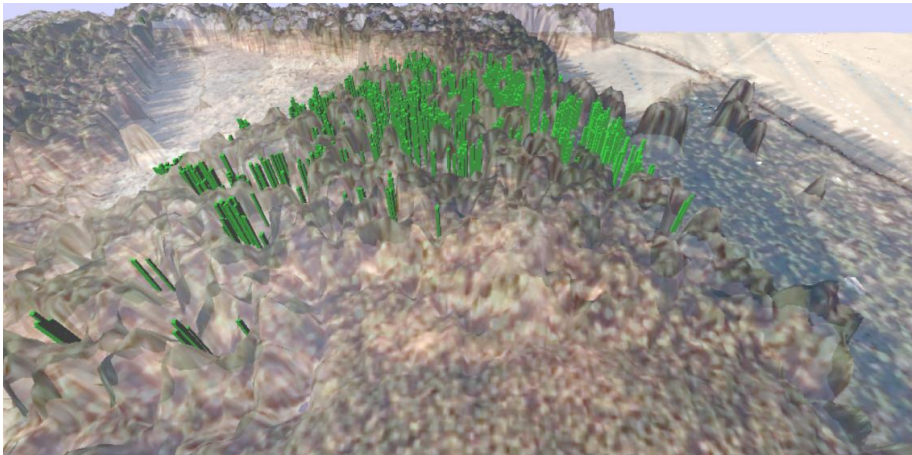
Figure 5: As in Figure 4, but now the DSM after the storm is displayed

5   Conclusions

The case implementation described in this paper demonstrates that it is currently possible to create dynamic 3D visualizations in geospatial applications using plain Web browser as the computing platform. By combining these new capabilities with the rich set of content resources and service-based spatial analysis and processing facilities, new kinds of end user applications can be created. In this paper a case study was presented, in which a forest owner is able to assess the effects of a storm

www.ogrs-community.org

in his forest by just using a Web-browser based client application to select the area of interest, letting the online analysis service to find out the difference between the surface models measured before and after the storm, and then analyzing the results in a freely rotatable and scalable 3D model. Several other use cases can be predicted in the 3D analysis of the changes occurring both in natural and in man-made environments.

Acknowledgments

References

Khronos, (2014), WebGL frameworks,
   URL: *http://www.khronos.org/webgl/wiki/User_Contributions* [30 April 2014].

Kim, H-W., Kim, D-S., Lee, Y-W and Ahn, J-S, (2014), 3-D Geovisualization of satellite images on smart devices by the integration of spatial DBMS, RESTful API and WebGL. Geocarto International. (Published online 31 Mar 2014).

OGC, (2014), Open Geospatial Consortium standards,
   URL: *http://www.opengeospatial.org/standards/* [11 April 2014].

Pedersen J., (2014), WCS i threejs.
   URL: *http://labs.kartverket.no/wcs-i-threejs/* [10 April 2014].

QGIS Plug-ins, (2014), QGIS Python Plugins Repository - Plugin: Qgis2threejs,
   URL: *http://plugins.qgis.org/plugins/Qgis2threejs/* [11 April 2014].

Rocatello E., Nozzi A. and M. Rumor, (2013) Design and Development of a Framework Based on OGC Web Services for the Visualization of Three Dimensional Large-Scale Geospatial Data Over the Web.
   URL: *http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-4-W1/101/2013/isprsarchives-XL-4-W1-101-2013.pdf* [10 April 2014].

Three.js, (2014), JavaScript 3D Library Home Page,
   URL: *http://threejs.org* [23 Apr 2014].

Morphometric signatures of landslides

Ivan Marchesini[1] , Mauro Rossi[1,2] , Alessandro Mondini[1] , Michele Santangelo[1] , and Francesco Bucci[1]

[1]CNR IRPI, Via Madonna Alta 126, 06128 Perugia, Italy
[2]Department of Earth Sciences, University of Perugia, Piazza dell'Universitá 1, 06100 Perugia, Italy

Abstract: Morphometric parameters are widely used in landslides mapping and modeling. Here we present a framework/procedure to analyze the morphological fingerprints of landslides in a territory. For this purpose we identify a set of morphometric variables and a procedure to distinguish different morphological landslide signatures. Our intent is to create a library of the landslide morphological signatures as much as possible complete. Results will be helpful to improve the ability to detect landslide on the surface, the modeling capabilities, and the knowledge of landslide processes.

Keywords: Landslide, Geomorphometry, signature, GIS

1 Introduction

Pike et al. (2008) defined Geomorphometry as the science of quantitative land-surface analysis. The term Geomorphometry has got a great impulse and a renewed interest starting from the book edited by Hengl & Reuter in 2008, which probably represents a sort of milestone for this interdisciplinary field. The primary objective of geomorphometry is the characterization of discrete surface features (landforms) through the analysis of land-surface parameters (e.g. slope, curvatures, etc.). Recently Jasiewicz & Stepinski (2013) introduced the concept of "geomorphons", a new approach (criterion) to identify/characterize surface landforms.

Morphometric analysis is widely used in landslide mapping and modeling. Geomorphologists implicitly (i.e. exploiting their own experience/capacity to analyze the shape of the slopes looking, as an example, for an upper concavity and lower convexity on a slope, that typically indicates the presence of a landslide -

Guzzetti et al. 2012) use morphometric information to map landslides in the field or through aerial photo-interpretation. Indeed spatial variation in slopes, aspect, roughness and curvature help photo-interpreters to detect and map landslides, particularly when slope movements are not recent and their surfaces not straightforward discernible by their photographic signature (i.e. by the variation in tones, textures and patterns in the image).

Landslide models can explicitly consider this information. For instance, the slope parameter is fundamental in physically-based landslide models, such as those based on the infinite slope criteria (e.g. Baum at al. 2008, etc.), or based on a more complex three-dimensional landslide schema (e.g. Mergili et al. 2013).

However, the widest use of morphometric parameters in the framework of landslide hazard assessment, is to estimate their importance as an explanatory variable in landslide susceptibility studies (e.g. Rossi et al. 2010, etc.). For this purpose mainly multivariate statistical classification models are used. Those statistical approaches evaluate the likelihood of a landslide occurring in a given area, given a set of local geo-environmental conditions which mostly include geomorphic parameters. Researchers have also attempted at quantifying morphological changes produced by landslides locally or in small areas. As an example, land-surface parameters have been used to identify active landslides (Berti et al. 2013) and, coupled with remote sensing data and techniques, to create reliable landslide event inventories (Mondini et al. 2013).

The lack of a set of common tools to identify and quantify these changes is an important limitation when trying to generalize (or extend) results obtained locally. In fact at present, no common criteria, or set of variables or analysis tools exist to globally compare such morphological changes.

2 The framework

2.1 Description

We have implemented a procedure, through an open source framework, to analyze the morphometry of landslides and non-landslides areas in a territory, and hence to identify the combination of morphometric conditions that characterize the areas where landslides have been mapped (Rossi et al. 2013). Our aim is to build a standard procedure to compare globally the landslide morphometric characteristics.

The procedure will be helpful to improve: (a) the ability to detect landslides from the analysis of Digital Elevation Models (DEMs), (b) the capabilities to model landslides in different environmental context, and (c) the knowledge of landslide processes.

The proposed framework makes use entirely of open source software and it consists of three steps (Figure 1).

www.ogrs-community.org

In the first step a Digital Elevation Model and a landslide inventory map are selected as inputs. Then a software tool extract a set of morphometric variables from the Digital Elevation Model both for different type of landslides in the inventory map, and for landslide free areas. The tool exploits GRASS GIS (GRASS GIS Development team, 2013) to calculate, for the different landslide types (classes) and for different kernel sizes, the main land-surface morphometric parameters: slope, cross curvatures, longitudinal curvature, etc. This part of the procedure exploits the GRASS GIS morphometric algorithms r.param.scale (Wood 1996) and r.slope.aspect (Horn 1981) but we are planning to introduce other algorithms, e.g. the aforementioned r.geomorphon. In the second step the Empirical Distribution Functions (EDFs) of the values of each parameter are analyzed using R (R Core Team, 2014) for each combination of class and kernel. The results of this analysis are stored in the system as R data files which represent the obtained Geomorphological Signature Function. Red connecting lines in Figure 1 identify the sequence and the tools used in the first two steps.
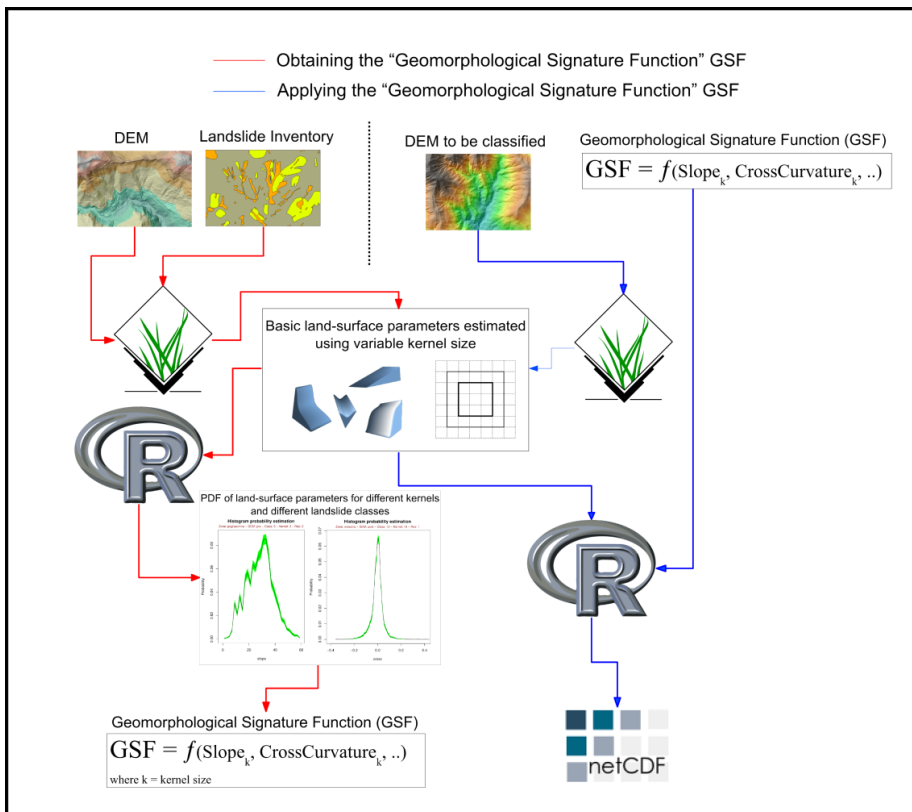

Figure 1: The procedure steps

The third step (blue connecting lines in Figure 1) deals with the application of the previously obtained EDFs and can be applied to another adjacent/similar zone where comparable slope processes and digital elevation model exist. Over this zone the basic land-surface parameter layers, the same type (slope, curvature, etc.) as used in the first two steps, must be re-calculated through GRASS GIS. Then, using R again, the stored EDF functions are exploited to estimate the probability that each single pixel of the zone belongs to a specific class among those defined in the step 1, i.e the landslide types and the no-landslide areas. This is done for each land-surface parameter layer and for each kernel size considered in the first step.

The output of the third step is a set of raster layers, stored into a netCDF file, where each raster corresponds to a combination of land-surface parameter, kernel size and landslide class. The values of the raster cells range from 0 to 1 and represent the probability of each cell for being in each landslide class given the particular combination of land-surface parameter and kernel size. The NetCDF was chosen since it is a well documented format, capable to maintain a huge amount of raster layers, appendable (in case previous results must be enhanced with more recent ones), portable, scalable and since there are a lot of R libraries able to efficiently read and write this format.

2.2 A preliminary application

The method was preliminarily applied on a small test area of the territory of the Messina Province, Sicily (Italy), where a high-resolution DEM (1m) was available. The area was subdivided in two adjacent basins.

In the first basin we have applied the proposed procedure using landslides and no-landslide areas information to identify EDF functions. For this basin a DEM and a landslide inventory map were available. Landslide inventory include different types of slope movements: rock falls, flows, slides, and complex movements. Here we present an example of the analysis performed considering small (3m), medium (33m) and large (63m) kernels sizes and few land-surface parameters: slope, longitudinal curvature, i.e. measured in direction of gradient, and cross curvature, i.e. measured in direction perpendicular to gradient).

In the second basin we used the derived EDF functions to calculate the probability of each cell for being in each landslide class. Figure 2 B,C,D,E,F show the probability raster layers for each combination of land-surface parameter and kernel size. In the figure red color represents high probabilities while blue color corresponds to low probabilities.

In figure 2A we show the probabilities associated to the five landslide classes (No-landslide, Slide, Rock fall, Flow, Complex landslide) derived for a single pixel (red square in the inset) in the test area in the Messina Province. For simplicity we drawn the probabilities corresponding to a single pixel, but similarly the probabilities associated to each other pixel can be derived. The morphometric

probability spectrum shown in Figure 2A shows that the considered pixel can belong to a rock fall landslide body.

3   Conclusions

The results of the procedure really depend on (i) the date of the mapped landslides respect to that of the DEM, (ii) the resolution of the DEM, (iii) the method used to produce the DEM, (vi) the post-processing procedures possibly applied to correct DEM artifacts, (v) the quality of the landslide inventories.

The proposed framework can be used to determine the geomorphological signature of the instability phenomena in a territory starting from the analysis of a specific set of geomorphic parameters. This information can provide a quantitative estimation of the relation existing among landslides and geomorphic parameters, and can be valuable (i) for the detection of landslides in a territory and (ii) for the landslide hazard assessment.

The strong interoperability between the used Open Source software, proved really powerful for both geomorphological and statistical analysis. Moreover it allows to exploit, in parallel, all the available CPUs of a very powerful computer, reducing the computational time. Open Source software was fundamental in our work of creating a standardized analysis framework that simplifies the collection of comparable geomorphometric results in other study areas.
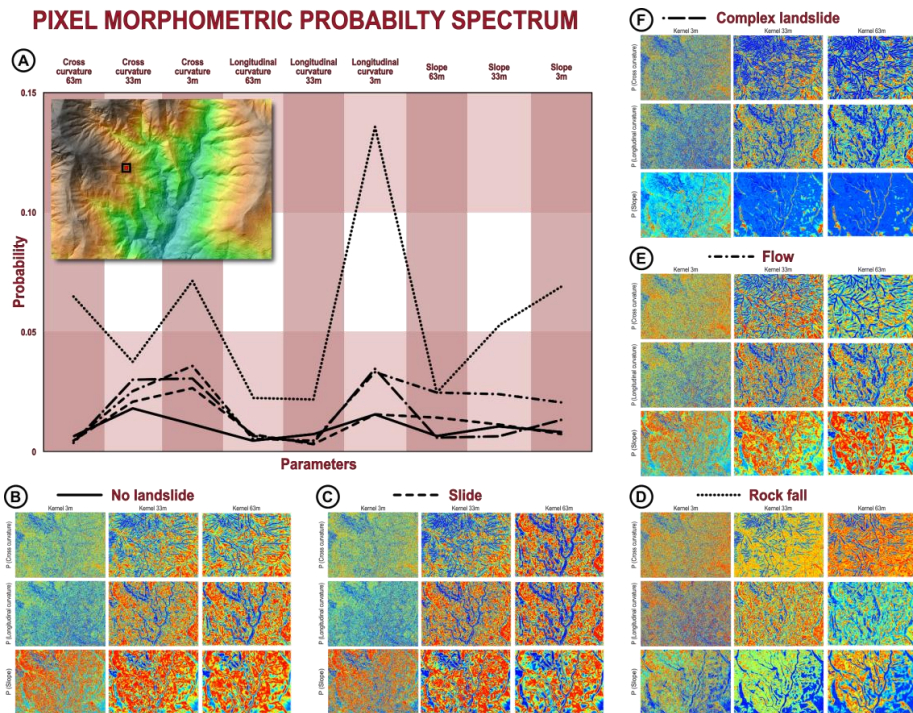
Figure 2: Example of the EDF functions and of the morphometric probability

Acknowledgments

References

Berti, M., Corsini, A. & Daehne, A. (2013), Comparative analysis of surface roughness algorithms for the identification of active landslides, Geomorphology, Volume 182, Pages 1-18, ISSN 0169-555X,
URL: *http://dx.doi.org/10.1016/j.geomorph.2012.10.022*.

Baum, R.L., Savage, W.Z. & Godt, J.W. (2008), TRIGRS—A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0: U.S. Geological Survey Open-File Report, 2008-1159, 75 p.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M. & Chang, K. T. (2012), Landslide inventory maps: New tools for an old problem. Earth-Science Reviews, 112(1), 42-66.

www.ogrs-community.org

Hengl, T. & Reuter, H.I. Eds. (2008), Geomorphometry: concepts, software, applications. Vol. 33. Access Online via Elsevier, 2008.

Horn, B.K.P. (1981), Hill Shading and the Reflectance Map, Proceedings of the IEEE, 69(1):14-47.

Jasiewicz, j. & Stepinski, T.F. (2013), Geomorphons — a pattern recognition approach to classification and mapping of landforms, Geomorphology, Volume 182, Pages 147-156, ISSN 0169-555X,
URL: *http://dx.doi.org/10.1016/j.geomorph.2012.11.005*.

Mergili, M., Marchesini, I., Rossi, M., Guzzetti, F. & Fellin, W. (2013), Spatially distributed three-dimensional slope stability modelling in a raster GIS, Geomorphology,
URL: *http://dx.doi.org/10.1016/j.geomorph.2013.10.008*

Mondini, A.C., Marchesini, I., Rossi, M., Chang, K., Pasquariello, G. & Guzzetti, F. (2013), Bayesian framework for mapping and classifying shallow landslides exploiting remote sensing and topographic data, Geomorphology, Volume 201, Pages 135-147, ISSN 0169-555X,
URL: *http://dx.doi.org/10.1016/j.geomorph.2013.06.015*

Pike, R., Evans, I.S. & Hengl, T. (2008), "Geomorphometry: a brief guide". In Geomorphometry: Concepts, Software, Applications, Edited by: Hengl, T and Reuter, H. I. 3–18. Ispra, Italy: European Commission.

Rossi, M., Mondini, A.C., Marchesini, I., Santangelo, M., Bucci, F. & Guzzetti, F. (2013), Landslide morphometric signature. 8th IAG/AIG International Conference on Geomorphology Geomorphology and Sustainability Paris, France, August 27-31, 2013. (Abstract, Poster).

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A.C. & Peruccacci, S. (2010), Optimal landslide susceptibility zonation based on multiple forecasts, Geomorphology, Volume 114, Issue 3, Pages 129-142, ISSN 0169-555X,
URL: *http://dx.doi.org/10.1016/j.geomorph.2009.06.020*

Wood J. (1996), The Geomorphological characterisation of Digital Elevation Models. Diss., Department of Geography, University of Leicester, U.K.

Decision support on distributed computing environment (IQmulus)

Angéla Olasz[1] and Binh Nguyen Thai[2]

[1]Department of Geoinformation, Institute of Geodesy, Cartography and Remote Sensing (FÖMI), Budapest, Hungary
[2]Department of Cartography and Geoinformatics, Eötvös Loránd University, Budapest, Hungary

Abstract: In archaeological cultural resource management there is a shift from setting up better data acquisition techniques to generate proper modelling methods or cross-validation between model results. Nowadays archaeological predictive modelling is extremely successful and productive, and it is being used as an effective tool in archaeological site location and explanation. In archaeological site modelling distributed file systems and parallel processing environments may offer new solutions. The Apache Hadoop environment offers the execution of large data sets across clusters of computers using the simple programming model MapReduce. This paper presents IQmulus, a distributed environment, where users can achieve rapid response to their needs not only on their own datasets but with others as well, along with distributed processing power for pre-processing and data analysis.

Keywords: Distributed file systems, MapReduce, predictive archaeological modelling, archaeological GIS, cultural resource management

1   Introduction

It as an unquestionable fact that data is of crucial importance for the geospatial sector, with emphasis on the aspects of the data acquisition, distribution, processing, analysis and visualization. There is exponential increase in the amount of geospatial information available in highly heterogeneous formats. The geospatial community considering all segments, such as industries, governmental institutes, organizations, foundations, authorities, or the academic world has to react to this huge development, from data storage needs to the implementation of new approaches for the automatic extraction of information from raw datasets.

Numerous new data acquisition techniques by both airborne and terrestrial technology have been elaborated. Firstly, in airborne surveys consider for example the expansion of UAVs (Unmanned Aerial Vehicles) and the hundreds of different sorts of instruments that can be installed on them. Along the software development in flight planning or semi-automatic real-time feature extraction, not to distinguish commercial or open source implementations, they are untraceably expanding. Likewise, innovations for instance in mobile-mapping systems with different kinds of data acquisition techniques combining LiDAR with digital imagery, allowing the users to jointly acquire laser point clouds and imagery and load the into GIS systems for applications such as inventories, maintenance and asset management from their office. All these techniques record lots of data, yet in different data models and representations; therefore, resulting datasets require harmonization and integration before deriving meaningful information from them. To summarize, huge datasets are available but raw data is almost of no value if not processed, semantically enriched and quality checked. Up to now, there are only limited automatic procedures for this. Though quickly-developing acquisition technologies provide huge amounts of data, a wealth of information is latent in many datasets.

Hereinafter, we will discuss predictive modelling methods with current technologies. After that a brief introduction on distributed systems and Apache Hadoop distributed environment. Finally a short introduction on IQmulus project as a solution for distributed decision support system.

1.1 Predictive modelling

Predictive modelling in general provides a powerful tool to predict the probability of an outcome by a statistical model created or applied to best fitting the phenomena. In geosciences for predictive modelling, the spatial data is collected for the relevant predictors (variables), a statistical model is formulated by different approaches, the predictors are made and the model is validated (or revised) by additional data or field work in a study area. Doing such predictions is accepted procedure in every scientific study where an endeavour is to get insight some unknown continuation. In archaeological applies predictive modelling is a technique that tries to predict the "location of archaeological sites or materials in a region, based either on a sample of that region or on fundamental notions concerning human behaviour" (Kohler & Parker 1986). Predictive modelling involves figuring out what environmental and social variables are attractive reasons for people choosing to live in a particular space, and using that information to predict where archaeological sites might be (Hasenstab & Resnick 1990). Is considered the location of archaeological remains in the landscape is not random, is connected to certain natural characteristics and socio-cultural aspects, briefly definite part of the landscapes were preferred by human settling. It is reasonable to assume, primarily a society relies on agricultural production a choice of settlements

location is suitable land for agriculture activity, water and the source of material for everyday usage among others. It is assumed that:

1. The human colonization of the landscape is not random,
2. The scattering of the found archaeological sites can be modeled by certain characteristics on the natural and cultural environment (Verhagen 2007).

It follows that using GIS and geostatistical analysis a certain portions of the landscape can be found where their probability of occurrence is high.

1.2 Recent solutions

Both, for Academic research in Archaeology and for CRM require and claim for methods to predict and explain the reason for location of various archaeological sites is of great importance. In GIS the predictive modeling applied in archaeology is one of the first more complex operations that were researched. CRM and academic use traditionally require two different approaches. In CRM the primary focus is to able to predict in a certain time the probability of the presence of archaeological sites, the need to understand why it is there is secondary. In academic research the central issue is to understand the reasons and mechanisms behind the choosing of a certain portion of the landscape. The predictive approach can be expressed as looking for locational factors and the interpretive as understanding locational choice factors. In the creation of predictive models also distinguished.

Inductive (CRM) or Deductive (academic) mostly differs on they applying the known sites locations or not, can also be viewed as the difference of using deduction from theory and induction from observations. But they are most adequately described as data driven and theory driven (Wheatley & Gillings 2002) Data driven models are applied to see if a relationship can be found between a sample known archaeological sites and a selection of landscape characteristics (environmental factors) if the correlation found then extrapolated to larger area (if is possible). Theory driven starts by formulating a hypothesis on the location preferences, and selecting weighting the appropriate landscape parameters (Verhagen et al. 2009).

For input data all models uses environmental factors as a thematic layer or a derived geostatistically processed data such as: elevation, slope, aspect, solar radiation, distance to water, vertical distance to water, cost distance to water, distance to water confluences, topographic variation (DEM, DSM), soils (Proximity to pottery clay), land use, distance to historic trading paths, proximity to firewood, and proximity to transport routes (Wescott & Mehrer 2007).

Decisions can then be made by comparing objectives or, at the lowest level of the hierarchy, by comparing attributes. Attributes are measurable quantities or qualities of a geographical entity or of a relationship between geographical entities.

An attribute is a concrete descriptive variable; an objective is a more abstract variable with a specification of the relative desirability of that variable. The attributes used for archaeological predictive modeling are usually a limited number of environmental variables; they can be supplemented with expert knowledge on less measurable variables. In the model created by Soonius & Ankum (1990) (see Figure 1.), the environmental variables used were selected by means of a χ2 test using the known archaeological sites (Wescott & Mehrer 2007).
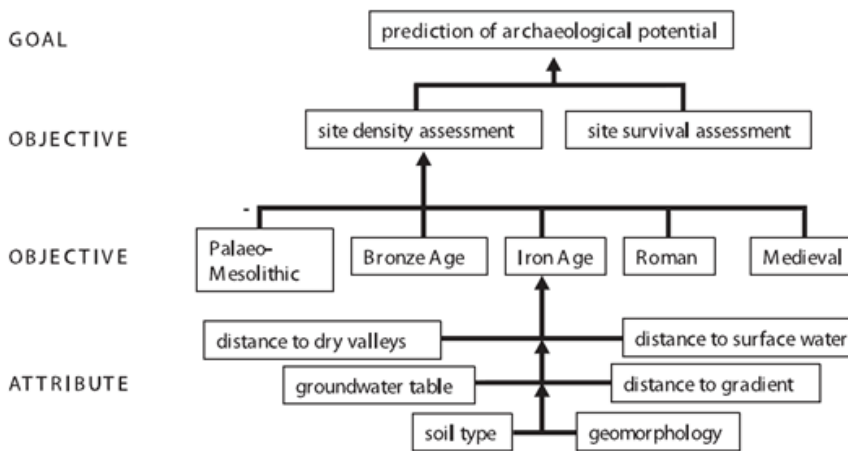
Figure 1. The hierarchical structure of objectives (Soonius & Ankum 1990)

The widely used modelling methods are developing concerning the technical progress of data-mining techniques. The base of the modeling methods are Bayesian statistics (weights of evidence), regression analysis, multi-criteria analysis, supervised classification/ unsupervised classification (or cluster analysis), fuzzy logic, artificial neural networks, and recently random forest modelling.

In the modeling process between the input data (location of the archaeological sites) and the different types of raster data (digital elevation model, soil map), is being evaluated if they have statistical correspondence - either negative or positive.

2 Distributed approach

Data may arrive from different sources independently in different formats resulting in a big heterogeneous set of data. Relational data stores are not prepared for this kind of cases. However most of the time, we are curious, if there is any relation or bond between these data, hence we are looking for interrelationship among data sets. Data is that they are stored permanently and seldom modified.

www.ogrs-community.org

The Apache Hadoop[1] is an open source framework/eco-system designed for capturing, organizing, storing and sharing, analyzing and processing data in a clustered environment giving the solution for linear scalability of both storage and computational power. The ability of storing and analyzing large amount of data on clustered computers reduces the invested cost remarkably.

In order to use or set up Hadoop, we must know the boundaries of Hadoop framework especially for decision makers and developers. Let's start with its disadvantages.

1. It's not a distributed Relational Database.
2. Security issues (no data caging)
3. It's not easy to configure, maintain and use.
4. Data management in Hadoop is different from RDBMS.
5. Linux and Java knowledge is needed to understand and use Hadoop.

If some of the above mentioned criterias is a must in the system requirements then Hadoop is not the right choice, otherwise you will get the following benefits by choosing Hadoop:

1. designed for storing and querying large data sets,
2. archived data can be resumed and used at any time,
3. real-time data analysis support,
4. batch processing,
5. support ETL (extract-transform-load) process.

Data have become a part of everyday work. Depending on application areas, some of them need high data availability, others need high computational functionality on large data sets. Scalability of any framework is a crucial point of success or failure, however not everyone can afford large hardware investments in a short period of time, not to mention the processing power is not increased proportionally with hardware investment.

Primary approach of Hadoop: "Data should be distributed and minimizing network latency by moving algorithms across clusters in a distributed manner". Hadoop have the following properties:

1. Data sits on one or more machines.
2. All nodes contain both distributed file system and data processing capability.

---

[1] http://hadoop.apache.org/

Hadoop is completely modular. Its core components are HDFS and Mapreduce which is required, others are optional depending on your need. Some of the widely used components are HBase, Hive, Pig, and Sqoop.

3   IQmulus

The objective of the IQmulus research project[2] is to enable the use of large, heterogeneous geo-spatial data sets (Big Geospatial Data) for better decision making through a high-volume fusion and analysis information management platform. The project, funded by the 7th Framework Programme of the European Union, is now after the first year from its four-year duration. To ensure user-driven development, it implements two test scenarios: Maritime Spatial Planning, and Land applications for Rapid Response and Territorial Management. The contribution of large number of users from different geospatial segments, application areas, institutions and countries have already been achieved. The incorporated project partners represent numerous different facets of the geospatial world to ensure a value-creating process requiring collaboration among academia, authorities, national research institutes and industry. Providing improvements for researchers, developers, academic staff and students allow them to sharing knowledge, implement new solutions, methods and ideas in the field of geoinformation through the development and the dissemination of a new platform. Furthermore, the research project offers new practices for end-users, following the continuous evolution the big geospatial data management requirements. As end users are among those providing geospatial support for decision makers, the whole decision making process can benefit from improvements and involvement of new practices.

User requirements are driving the development of the IQmulus solutions through discussing basic questions and problems, expectations and visions about how a high-volume fusion and analysis information management platform would be implemented and how would it provide solutions better than their actual working processes. Users were interviewed on their current workflows as well as their envision how those could be resolved in an ideal or even futuristic geospatial world. For example, applications related to Rapid Response and Territorial Management has been discussed with municipalities, disaster management, water management and spatial planning authorities or universities' research groups. In addition, they were asked to assess their existing approaches (in-depth analysis of strengths and weaknesses, e.g., lack of data, lack of computational resources, lack of high capacity operational system, lack of faster modelling approaches, lack of free

---

[2] https://www.iqmulus.eu/

www.ogrs-community.org

open source geospatial technology etc.) to find out where the IQmulus platform can yield significant improvements.

As a consequence of user requirements identification, in the current development phase the creation of concrete processing services has started. Services consist of algorithms and workflows focusing on the following aspects: spatio-temporal data fusion, feature extraction, classification and correlation, multivariate surface generation, change detection and dynamics. In the following phases, system integration and testing will be done, and the IQmulus system will be deployed as a federation of these services based on the above mentioned two scenarios. Actual deployment will be based on a distributed architecture, with platform-level services and also basic processing services invoked from clusters and conventional computing environments. Concerning workflow definition and execution, the concept is to develop a set of domain-specific languages that makes the definition of several types of spatial data processing and analysis tasks simpler. They can be compiled to run on parallel processing environments, such as GPGPUs, Cloud stacks, or MapReduce implementations such as Hadoop. The compiled algorithms can be packaged as declarative services to run on the respective environment they have been compiled for. The Apache Hadoop open source framework has already been chosen as a basis for architecture development. This framework offers the execution of large data sets across clusters of computers using simple programming models. It is supplying the next- generation cloud-based thinking and is designed to scale up from single servers to hundreds of workstations offering local computation and storage capacity. A Hadoop Distributed File System (HDFS) has already be installed in cloud environment, providing high-throughput access to application data. The Hadoop MapReduce system for parallel processing of large data sets is currently under implementation. Thanks to this solution, users can achieve rapid response to their needs used by their own datasets in cloud processing by exploit the system capacities.

Despite of all the features provided by Hadoop eco-system, large raster-based data storage and processing is not efficient on HDFS. In order to solve this problem further research and investigation is needed by applying segmentation and clustering algorithms on raster images based on HDFS and MapReduce method.

4   Conclusion

In conclusion working on the above mentioned open issues, IQmulus aims at the development and integration of an infrastructure and platform that will support critical decision-making processing in geospatial applications using Open source solutions at the first opportunity. It's obvious that IQmulus is not applicable for every decision support problems. Achievements and further development directions with benchmark data will be presented in the paper through a concrete application example arising from user requirements.

OGRS 2014

Acknowledgments

References

Hasenstab, R.J. & Resnick, B. (1990), GIS in historical predictive modelling: The Fort Drum Project. *In* Interpreting Space: GIS and Archaeology. Kathleen M. S. Allen, Stanton W. Green, and Ezra B. W. Zubrow, eds. Pp. 284-306. London: Taylor & Francis.

Kohler, T.A. & Parker, S.C. (1986), 'Predictive Models for Archaeological Resource Location'. In: M.B. Schiffer (eds.), Advances in Archaeological Method and Theory 9, Orlando, Academic Press, pp. 397–452.

Soonius, C.M. & Ankum, L.A. (1990), Ede II: Archeologische Potentiekaart, RAAP rapport 49, Stichting RAAP, Amsterdam.

Verhagen P. (2007), 'Case Studies in archaeological predictive modelling'. Doctoral Thesis, Leiden University Press,
URL: *https://openaccess.leidenuniv.nl/handle/1887/11863*.

Kamermans, H., van Leusen, M. & Verhagen, P. (Eds.) (2009), Archaeological prediction and risk management. Leiden: Leiden University Press.

Wescott L.K. & Mehrer, W.M. (2007), 'GIS and Archaeological site location modeling', Taylor and Francis, FL. 58-89.

Wheatley, D. & Gillings, M. (2002), Spatial Technology and Archaeology: the archaeological applications of GIS. London: Taylor & Francis.

Management of history of the rotation of agricultural parcels through open-source tools

Christine Plumejeaud-Perreau[1], Nathalie Long[1], Ba-Huy Tran [2], Alain Bouju[2], Bruno Lescalier[2], Karell Bertet[2], Michel Ménard[2], and Vincent Bretagnolle[3]

[1]Laboratoire Littoral, Environnement et Sociétés (LIENSs), U.M.R CNRS 7266, 2 rue Olympe de Gouges, 17000 La Rochelle, France
[2]Laboratoire Informatique Image et Interaction (L3I), Université de la Rochelle, Av. Michel Crépeau, 17000 La Rochelle, France
[2]Centre d'Etudes Biologiques de Chizé (CEBC), U.M.R CNRS 7372 & Université de la Rochelle, 79360 Villiers-en-Bois, France

Abstract: This paper presents how an open-source GIS can help to manage a spatio-temporal database dedicated to the long-term monitoring of an agricultural landscape. Based on the space-time composite model (Langran & Chrisman 1998), the system particularly focus on crop rotations and change of shape in agriculture parcels. In order to improve the performance and the flexibility of the system, a migration to an open-source solution was performed. At first, we highlight benefits of this approach as well as its limitations in terms of the ability to integrate new data sources. Then, we propose a new architecture that would facilitate the adaptation of software solutions to the extensions of our data model.

Keywords: Spatio-temporal model, Geographic information system, Crops rotation, Environment

1   Introduction

In rural areas with a predominance of agricultural activities, the study of environmental issues such as biodiversity preservation, soil erosion by water and tillage, erosive runoff, water pollution and gene fluxes may benefit from the long-term analysis of the crop mosaic resulting from farming practices. In fact, agricultural landscapes are primarily designed by farmer decisions dealing with crop choices and crop allocation at the farm scale. The arrangement, the shape and the nature of crops compose

the spatial organization of landscape which impacts ecological processes at various scales (Lazrak et al. 2010, Schaller et al. 2012). This information can be relevant when studying links between socio-economic environment and agricultural practices and subsequent spatial organization of landscape.

Recognizing the benefits of the long-term observation of agricultural practices for research on environmental issues, the CNRS Chize has established an observatory for crop rotations on the "*Zone Atelier Plaine & Val de Sevre*" workshop area. Since 1994, a Geographic Information System for the Environment (GIS-E) has been deployed in order to monitor the crop rotation of agricultural parcels.

The paper presents first the context of this interdisciplinary research around the GIS-E. The second section describes benefits of the migration of existing system to an open-source solution. The next section specifies the extensions of the data model and proposed a new architecture that would have the advantage of facilitating the adaptation of software solutions for the integration of new data sources. Finally, the conclusion summarizes the progress achieved around the GIS-E, while highlighting the problems we have identified about the system design and our perspectives.

2 Background

2.1 Environment observatory in the Chize workshop area

The need for a long-term observatory of rotation and associated biodiversity indicators has been recognized in a research program lead by the CNRS since 1994 called "*Zone Ateliers*"[1] . The observation area of 500 km^2 is designated as "*Zone atelier Plaine et Val de Sèvres*"[2], located south of Niort, in the department of Deux-Sevres in Poitou-Charentes, France.

Since 1994, land uses of 19,000 agricultural parcels are recorded from the field each year by Agripop (CNRS Chize) team and centralized in a database named "*Assolement*". An agricultural parcel is defined as a management unit, a polygon surrounded by entities having different land use in successive years (typically four). Agricultural parcel is bounded by physical limits such as a road, a river, a field path or a single field boundary. It contains only one type of crop. It differs from the cadastral parcel, but also the blocks stored in the RPG [3] which are updated every two years and distributed by IGN[4].

Besides, flora and fauna data collected in the fields for several years by Agripop team are centralized in another database. These data are reported by different researchers based on their observations through a Web interface.

2.2 The Geographic Information System for Environment (GIS-E)

The study of spatial organization changes of landscape requires a careful management of the history of the rotation of agricultural parcels. This is far from being trivial since we confront then to well-known "*split-tract problem*", the same as Howestine (Howenstine 1993) has raised for long term statistics on socio-economic tracts. New

---

[1]http://www.za-inee.org

[2]www.zaplainevaldesevre.fr

[3]Database recording the identification of agricultural parcels for the French government

[4]http://www.geoportail.gouv.fr/donnee/48/registre-parcellaire-graphique-rpg-2010
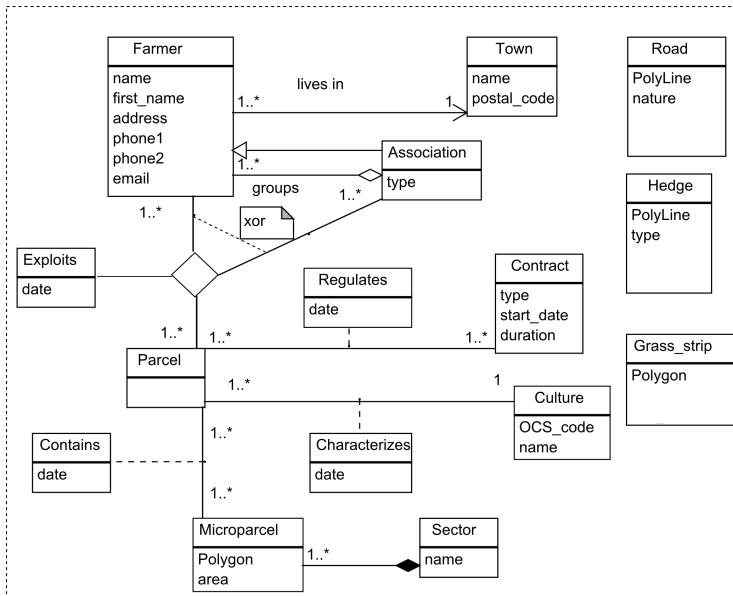
Figure 1: Space-time composite model used for rotation and changes of agricultural parcels.

models for spatio-temporal databases have been designed in order to solve problems raised by the classic *"snapshot layer"* model (Armstrong 1988) : memory consumption, no temporal links between parcells. Among them, the Space-Time Composite paradigm has encountered a great success due to its efficiency (Gregory 2002) : instead of storing each parcel's geometry for each year, the model records one layer of small geometry pieces (called *"microparcels"*) made by the intersection of all parcels layers during the observation period. The geometry of any parcel can be rebuilt "on the fly" through unioning the microparcels that are stored as belonging to it.

In the case of the presented environment observatory, this model is well suited to annual data acquisition, without revision of previous years. Indeed, researchers conduct two surveys on the ground every year to identify early and late crops (in March/April, and in June/July). All data gathered is then controlled and centralized in the server : around 24 000 microparcels are thus stored, for a total count of 19 000 parcels. There is therefore no revision of anterior information.

The schema in (Figure 1) shows the model used for our area study. The association class *Contains* plays the role of change dictionary between *Parcel* and *Microparcel*. *Farmer* or *Association* exploits parcels at various dates, because Association groups one to many Farmer(s). Some parcels are regulated by various kind of contracts, each *Contract* described by its type, its duration and a date of start (*start_date*). The rotation of cultures is described by the time stamped relation *Characterized* between *Parcel* and *Culture*. Since the kind of road (highway, main road, path) does not play the same role in the ecological processes, we keep some attributes describing it (*nature* is a summary).

The long-term observation not only requires a suitable data model, but also an

operational system with user-friendly interfaces, similar to the proposed solution by Geographic Information Systems (GIS). Each year, users must register both the new type of land use and possibly the changes of shape for each parcel. GIS tools offer many facilities for manipulating geographic entities (their properties, their forms) from ergonomic menus that users are accustomed to use. Therefore, the first software solution, developed in 2007, was based on ArcGIS coupled with an Access database through VBA scripts. Applying the space-time composite model described above, the system allowed the acquisition and data analysis of crop rotation. Annual rotations are also exported to a Web server which is connected to the fauna and flora database. Since 2005, the acquisition of rotation data is performed in the field from tablet PCs having a copy of the crop rotation database.

2.3 New requirements and limitations

Nowadays, in order to open the information system to the Web and to facilitate the analysis between landscape evolution and biodiversity changes, the incorporation of fauna and flora data into "Assolement" database is highly considered. Furthermore, information about the linears (road, hedges, rivers, railways, power lines...) which allows assessing the level of horizontal and vertical fragmentation of the landscape is still stored as snapshot layers. This necessarily involves an extension of the data model and a modification of utilities handling the model.

However, the new ESRI products (ArcGIS from version 9.0) do not longer support VBA scripts but Python, which prevents any evolution of this solution. Moreover, in terms of analysis, the performances of the current solution are relatively poor because Access cannot provide the same performance as a DBMS with a spatial extension. For example, one needs 2 minutes to extract the historical of land use of parcels.

3 An open-source solution

Following the analysis of the existing solution, an open-source platform was taken into account. Data was first transferred from Access to a PostgreSQL database, thus, the PostGIS spatial extension could be exploited. Then, Quantum GIS cross-platform was chosen, as an open-source solution, for implementing a customized interface (plug-in) based on our spatio-temporal database, named "Assolement".

Based on QGIS 1.8 and its Python API, the new solution covers all features offered by the previous one, ArcGIS associated with Access. The new plugin allows users to get a view of all parcels cultures per year, to edit information about current contracts, associations, regulations associated with the parcels and to collect new cultures or geometries each year in the database through its interface (Figure 2). In term of performance, this new system proves a noticeable improvement (only 10 seconds to get the history of cultures for all parcels versus 2 minutes in the previous system). The database size is around 400 MB, reducing the previous one that was bigger due to use of "home-made" geometry types used by Access.

Despite a high level of reactivity of the new platform, it turns out that the relationship between Quantum GIS and the spatio-temporal model raises some new challenges. First, a functional layer by a query joining three tables simultaneously ("*Parcel*", "*Microparcel*" and "*Contains*") can not be constructed. One solution is to build a table for storing all information necessary for the select queries, thus constituting a
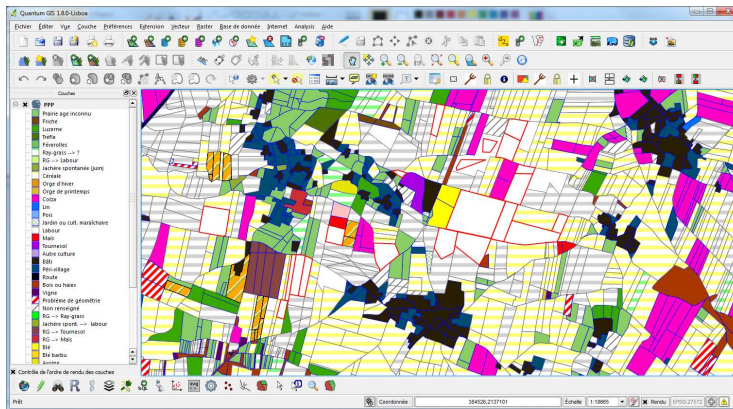
Figure 2: "Assolement" plugin interface: view on the 2013 layer.

model view, and offering a point of view through this snapshot layer. In this approach, updates on this table should be propagated to the model tables. Such method cannot ensure data consistency. Solving this problem would require changes in data sources connections management of Quantum GIS. Therefore, an update of plug-in API to Quantum GIS 2.2 (released in February 2014) should be taken into consideration.

Next, the applied data model does not integrate data describing the different linears which fragment the landscape, as well as fauna and flora data. To analyze landscape evolution and ecological corridors, it is necessary to incorporate information about roads, trees hedges, grass strips, waterways, railways, power lines...These objects must be modeled and stored with their dates of appearance and disappearance, and also with different topological relationships between them and the parcels.

Finally, the current model is implemented in a relational database optimized for Aggregate queries (Robinson et al. 2013). If a view based on spatio-temporel graph is generated, as presented by (Mondo et al. 2010), and the user interfaces are addressed in this view, the system could gain more flexibility and performance. Indeed, we can integrate progressively new data sources (fauna and flora database) and offer different connections to the new model entities through the spatio-temporal graph.

4   Towards a graph-based model

We intend to develop a modeling approach based on a graph representation which offers more flexibility for data manipulation. For instance, the modeling events occurring in parcel, like division, merge or redistribution (Claramunt & Thériault 1995, Plumejeaud et al. 2011) is planned to accelerate the search queries. Indeed, one disadvantage of the space-time composite model, as implemented here, is not knowing the duration of parcels shape in advance. We also consider building a new interface for data connection in the form of a graph within Quantum GIS, allowing exploitation data associated with the spatio-temporal model.

Ontologies and taxonomies will be used to facilitate the representation of knowledge (concepts, properties) and the integration of multiple sources. This approach is

well illustrated by the *"Linked Data"*[5] movement which encourages evolution of the Web towards a global data space. In addition, this approach allows capturing domain knowledge in the form of rules using SWRL [6] or RIF[7].

We believe this approach is more flexible to add new concepts and properties. In a first step, a platform like D2RQ will be used in order to access relational databases as virtual and read-only RDF graphs. Then, a new kind of browsing for environmental data could be developed based on tools like SEWELIS [8] (Ferré 2009) or a faceted browser on top of RDF data. SEWELIS uses faceted search, a recent navigation approach which defining a summarization and navigation data structure that provides rich views over data, and guides users from view to view. Guided navigation returns rich feedbacks to users, and prevents them to fall in dead-ends. This approach would allows exploiting association rules among features, extracting correlations among data or verification of user assumptions.

5   Conclusion

The works presented in this article are part of an interdisciplinary project *"Géo Connaissances des milieux naturels"* which aims to enhance the data exploitation gathered by the *"Zone Atelier Plaine & Val de Sèvre"* observatory. Since 1994, the observatory has been built a database on crop rotation based on space-time composite model that is well suited to the problem being treated. It was necessary to make evolve the Geographic Information System for Environment in order to allow comparative analyzes between crop rotation, landscape and biodiversity through the integration of other data sources. This change affects both the software infrastructure and data model.

As a first step, the system was evolved to an open solution based on the relational database manager system PostgreSQL with its PostGIS spatial extension. Afterwards, a user interface was developed through the Quantum GIS plug-in. Despite its high level of service, the proposed solution is still closely linked to the views per dataset of the same type. This prevents to handle the consistency of all information. Thus, the migration of the software infrastructure does not prove enough to deal with the challenges raised by the opening of new data sources.

In a second step, we describe the main benefits of the model migration to a knowledge graph as well as the mapping process between different data sources. This approach could provide greater flexibility and performance to the GIS-E, thus, ensuring its responsiveness to the user requests. Our perspectives concern the deployment of this solution, and its evaluation including an opening on the Web of our system.

References

Armstrong, M. P. (1988), Temporality in spatial databases, *in* 'GIS/LIS'88', pp. 880–889.

Claramunt, C. & Thériault, M. (1995), Managing time in gis an event-oriented approach, *in* J. Clifford & A. Tuzhilin, eds, 'Recent Advances in Temporal Databases', Workshops in Computing, Springer London, pp. 23–42.

---

[5]http://linkeddata.org
[6]Semantic Web Rule Language
[7]Rule Interchange Format
[8]SEmantic WEb Logical Information System

www.ogrs-community.org

Ferré, S. (2009), Navigating the Semantic Web with Logical Information Systems, Technical Report PI 1934.

Gregory, I. N. (2002), 'Time-variant gis databases of changing historical administrative boundaries: A european comparison.', *T. GIS* **6**(2), 161–178.

Howenstine, E. (1993), 'Measuring demographic change: The split tract problem', *The Professional Geographer* **45**(4), 425–430.

Langran, G. E. & Chrisman, N. R. (1998), 'A framework for temporal geographic information', *Cartographica: The International Journal for Geographic Information and Geovisualization* **25**(3), 1–14.

Lazrak, E., Mari, J.-F. & Benoît, M. (2010), 'Landscape regularity modelling for environmental challenges in agriculture', *Landscape Ecology* **25**(2), 169–183.

Mondo, G. D., Stell, J. G., Claramunt, C. & Thibaud, R. (2010), 'A graph model for spatio-temporal evolution', **16**(11), 1452–1477.

Plumejeaud, C., Mathian, H., Gensel, J. & Grasland, C. (2011), 'Spatio-temporal analysis of territorial changes from a multi-scale perspective', *International Journal of Geographical Information Science* **25**(10), 1597–1612.

Robinson, I., Webber, J. & Eifrem, E. (2013), *Graph Databases*, O'Reilly Media, Incorporated.

Schaller, N., Lazrak, E., Martin, P., Mari, J.-F., Aubry, C. & Benoît, M. (2012), 'Combining farmers' decision rules and landscape stochastic regularities for landscape modelling', *Landscape Ecology* **27**(3), 433–446.

Geographic knowledge discovery from sparse GPS-data with R – spatio-temporal examination of Amazonian river transports

Henrikki Tenkanen[1], Maria Salonen[1], Matti Lattu[2], and Tuuli Toivonen[1,3]

[1]Department of Geosciences and Geography, University of Helsinki, Finland
[2]Department of Teacher Education, University of Helsinki, Finland
[3]Department of Biosciences, University of Helsinki, Finland

Abstract: Technological developments have made it viable to collect movement data of humans, vehicles and animals. A low-cost GPS-based monitoring system was developed to collect movement data of riverboats in Western Amazonia. The pilot monitoring system provides real-time location and time data that indicate where and when the collaborating vessels are navigating. We developed a tool to identify individual trips from the data and compute motion parameters such as speed, change of speed and the direction of movement. The monitoring system was piloted by analyzing spatio-temporal characteristics of local riverboat transportation in Western Amazonia during the year 2012. The aim of this case study was to explore the seasonal and directional variation in the transportation characteristics along the rivers. Furthermore, we analyzed the relationship between the water height of the rivers and the travel speed of individual journeys. Results of the analysis show that navigation along the rivers indeed has seasonal and directional variation, and that the river morphology seems to affect the movement patterns of the vessels. Our case study also proved that our tools to collect and analyze the movement data functioned as planned.

Keywords: Data mining, GPS, transportation, Western Amazonia, R

1 Introduction

A vast amount of spatio-temporal data has become available with the fast development of information technology and different monitoring systems over the last two decades. Position-aware devices are one of the most dominant sources for collecting movement data of various targets such as vehicles and people (Pelekis et al. 2012). Spatio-temporal information that is derived from the tracking devices enable to build movement patterns from the targets, and to calculate measurable motion parameters such as speed, change of speed or the direction of movement (Laube & Imfeld 2002, Dodge et al. 2009).

For this study a specific low-cost GPS-based monitoring system called Amazonian Riverboat Observation System (AROS) was developed to collect movement data of the local riverboats in the departments of Loreto and Ucayali in Peruvian Amazonia. AROS uses SPOT Personal Trackers[1] to provides real-time GPS-data with coordinates and timestamp that indicate where and when the collaborating vessels are navigating. Also a specific tool called Trajectory Reconstruction and Analysis Tool (TRAT) was developed to analyze the mobility data. The main purpose of TRAT is to identify individual trips from the data and calculate different movement parameters from the GPS-data that was obtained with AROS.

To test the use of AROS and TRAT, we analyzed different spatio-temporal transportation characteristics based on data from year 2012. The aim of the study was 1) to technically test AROS monitoring system for capturing the spatio-temporal characteristics of riverine transportation and 2) to test TRAT for detecting individual journeys from AROS data and calculate different movement parameters. The third aim was to study the transportation patterns in the study area based on developed tools. Here, our aim was to find out if there is seasonal and directional variation in transportation characteristics along the rivers, and if there can be find connection between water height of the rivers and travel speed of individual journeys.

Monitoring the movements of vessels along Amazonian rivers for long period of time enable us to evaluate how much the river dynamics (i.e. high water/low water) effect on transportation patterns. Such information is particularly useful, as the entire transportation system in the Peruvian Amazonia is highly dependent on river transportations (Rodriguez Achung 1994). There are practically no roads in the area of the size of Germany and the rivers form the main transportation network for the people living in the area. Local economies and interaction between populations depend on public transportation loosely organized with long distance river launches.

---

[1] http://www.findmespot.eu/en/index.php?cid=101 [accessed 12.5.2014]
www.ogrs-community.org

2   Materials and methods

2.1   Data collection system

Data used in this study was collected with an experimental data collection and monitoring system AROS that resembles the automatic identification system (AIS) that is used to track professional maritime vessels (see e.g. Demsar & Virrantaus 2010). AROS is based on low cost satellite messengers (approx. 100€ each) that enable to determine the location of the devices by GPS satellite system and to send their location information via communication satellites to a database (see Figure 1).
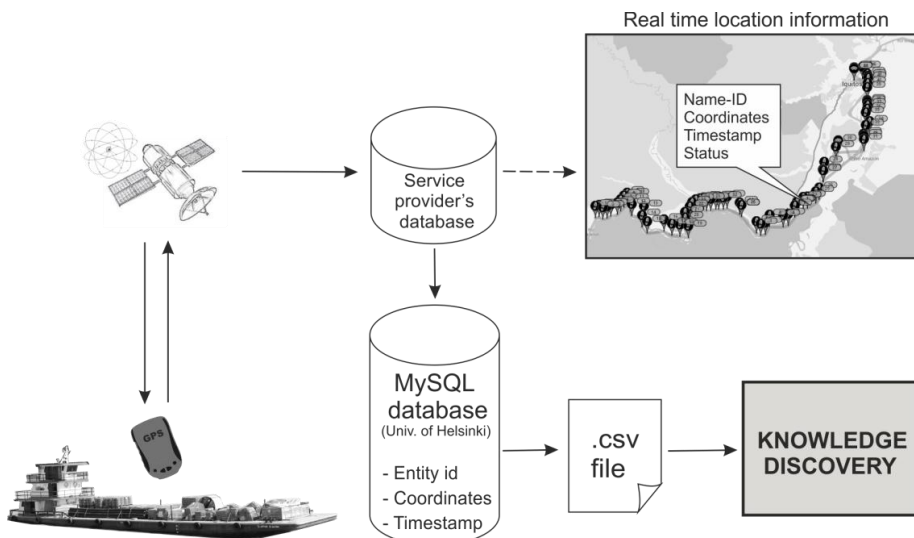

Figure 1: Concept of the pilot data collection system that utilizes GPS satellite messenger that is able to both receive and send location information

Altogether five satellite messengers were used to collect GPS-data of different boats. The messenger devices send their location information every 10 minutes to a service provider's database from where the data is automatically retrieved via XML-feed to a MySQL database at the University of Helsinki for permanent storage. Data consists of seven attributes: id, latitude and longitude coordinates, timestamp, Unix time and message type (test/tracking). The dataset used in this case study consisted of 11 572 GPS observations collected with AROS.

2.2   Data analysis

Analyzing mobility data and reconstructing trajectories from GPS-points is not a straightforward procedure (see e.g. Marketos et al. 2008). For detecting individual

journeys (i.e. single trip from origin harbor to destination harbor) and extracting knowledge from such trajectories, an analytical tool (TRAT[2]) was developed to fully automatize the knowledge discovery and visualization processes. TRAT is distributed as open-source code and it is developed for R. It uses geospatial packages available for R, and a package called Muste (Sund et al. 2012) that is used as user interface and tool for data management.

GPS-data from AROS is relatively sparse, and therefore it was necessary to develop an analytical approach that utilizes training dataset for calculating different motion parameters. This was done by enriching the original GPS dataset with ancillary data that contains information about network distance from the origin harbor, sinuosity index of river etc. (see Figure 2). Each AROS observation was spatially joined to the closest reference point which enabled us to calculate network distance and speed between consecutive points along the meandering rivers, determine the navigation direction and detect individual journeys from the data. Detecting an individual journey from the data mass is based on a specific algorithm that determines if an observation belongs to the same trajectory as earlier one or not. This classification is intrinsically based on following criteria: direction of movement, stationary time, temporal gap between observations, and spatial gap between observations.
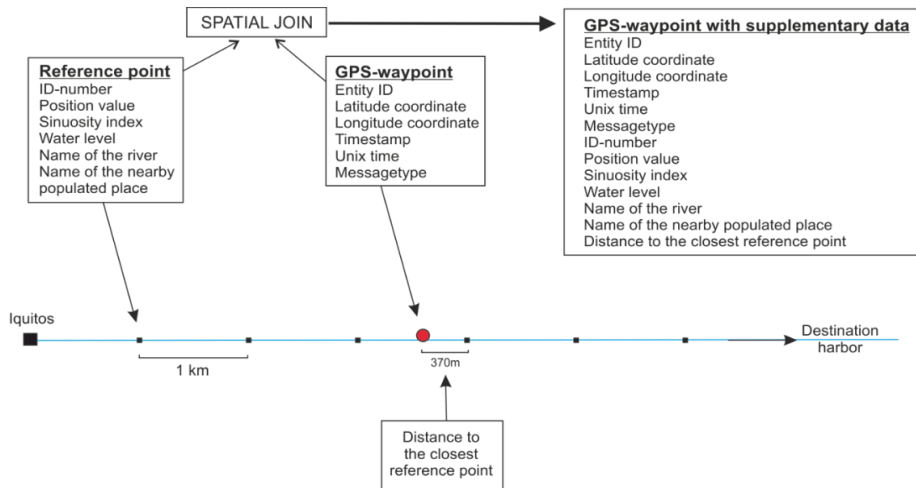

Figure 2: The final dataset is enriched by spatially joining each GPS-waypoint to the closest reference point

3   Results

Here we present few key findings from our case study and give examples how mobility data can be visualized. Figure 3 represents individual journeys that were detected from AROS data (year 2012) and classified by main navigation direction. Interactive 3D visualization of trajectories in space-time cube, where x- and y-axis represent space and t-axis represents time, enabled us to assess the quality of TRAT for detecting individual journeys.
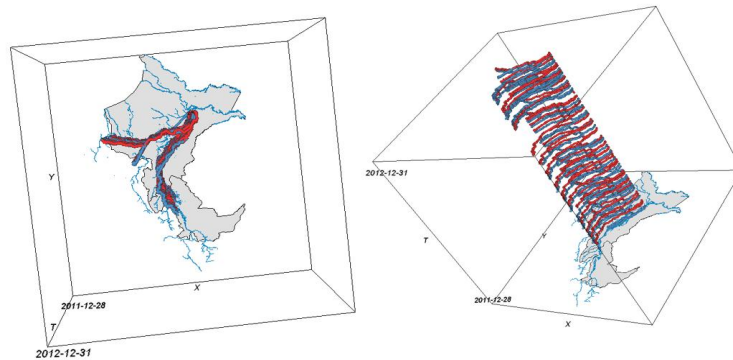

Figure 3: Individual journeys identified with TRAT and represented in space-time cube. Data was collected with AROS in 2012

Figure 4 represents an example of movement profile of the journeys which shows the travel speeds of the journeys (y-axis) along the vessel's navigation path (x-axis). Results show that navigation along the river has clear directional variation since navigating downstream is over 40% faster compared to upstream navigation during high water and intermediate. However, during low water there is not difference between navigation directions. Results show also clear seasonal variation in travel speeds, and hence navigation was over 30% faster during high water compared to low water (navigating downstream). Navigating upstream was fastest during low water but seasonal differences were considerably lower compared to downstream navigation.

Figure 5 represents the average speed of individual journeys (red lines) along circular axis that represents time. Black lines inside the red ones represent the percentage of the journey (from total route length) that was captured by AROS and thus can be used to assess the quality of the data. Blue polygon in the middle of the radar plots represent daily water levels of the river where vessels are navigating.

Fitting simple regression model between average travel speed of the journeys and water levels of the river revealed that there seems to be strong connection

between travel speed and river height when travelled downstream (R^2=0.73) along highly meandering river. On other cases that were studied, the results suggest that there is no connection between travel speed characteristics and river height.
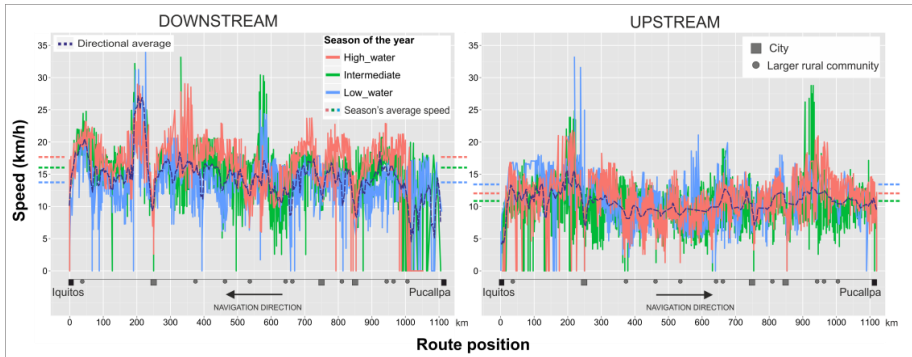

Figure 4: Movement profile representing the travel speeds of the vessels along the navigation route. Journeys were classified by season and navigation direction.
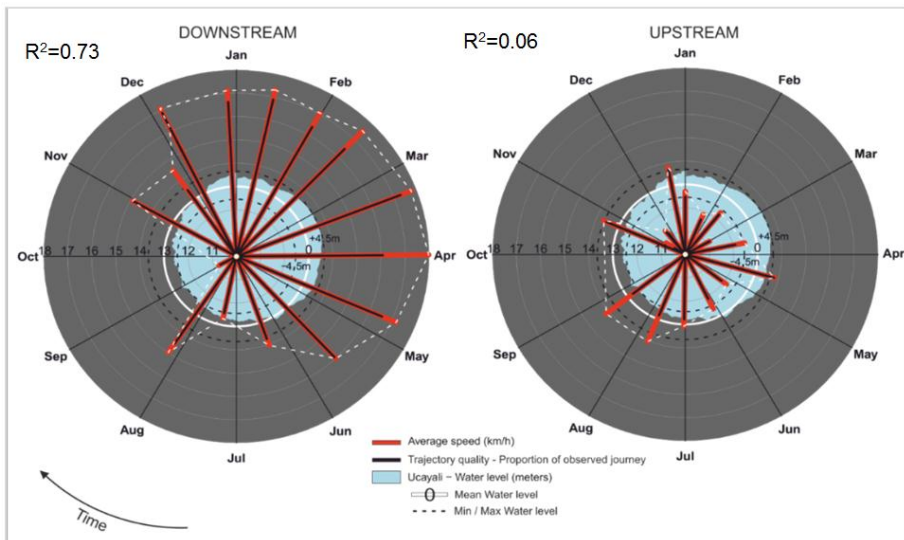

Figure 5: Radar plot representing the average speeds of individual journeys vs. water level. The R^2 values from linear regression (y=speed, β=water level) were 0.73 and 0.06 for downstream and upstream data respectively.

4 Discussion

Monitoring the movements of Amazonian riverboats with low-cost monitoring system AROS proved to work well as means for capturing the seasonal variation in movement characteristics along the rivers. Even though the sample frequency of AROS is relatively low (1 obs. / 10 minutes), it was high enough for our purposes because the travel speed of the vessels along the rivers is relatively slow (max 25 km/h). The satellite communication of the system was also functioning well and was a prerequisite for the system as there is no mobile network available in the area.

Travel speed calculations of the data analysis tool TRAT proved to be mainly accurate. The only exceptions were few cases where travel speed was higher than expected because of erroneous spatial join to the training dataset (i.e. river network). Assessing the accuracy of TRAT to detect individual journeys from the data proved that TRAT functioned as desired, thus 100% of the individual journeys and main direction of movement (upstream/downstream) were correctly detected from the data. Also the results of the case study seem to be in accordance with previous studies (cf. Salonen et al. 2012) that also imply that AROS and TRAT function as planned.

Transportation plays a key role when trying to find the factors affecting on development of a certain location. It, together with the overall accessibility of a location (i.e. how well the location can be reached) has a significant role in local livelihoods, land use change, deforestation, and conservation effectiveness in the region. Results of this study provide more accurate data for studies focusing on previously mentioned topics in the study area. Active development of openly available analytical tools is of paramount importance especially in areas such as rural Peru where local scientists can confront difficulties in getting hold of commercial GIS software that can analyze novel data sources such as mobility data. AROS and TRAT offer an excellent low cost solution to study global accessibility issues. By making a few modifications to the tool they could be used in other contexts such as studying global transportation patterns of professional vessels. R proved to be great platform for developing tools to analyze AROS data since it provides large collection of available packages that can be directly used for developing GIS tools that can handle GPS-data and visualize the results in an understandable manner.

Acknowledgments

References

Demsar, U. & Virrantaus, K. (2010), Space–time density of trajectories: exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science*, 24(10), 1527–1542.

Dodge, S., Weibel, R. & Forootan, E. (2009), Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems* 33, 419–434.

Laube, P. & Imfeld, S. (2002), Analyzing relative motion within groups of trackable moving point objects, *in* Egenhofer, M. & D. Mark (eds.), Geographic Information Science 2478, Springer, Berlin-Heidelberg, Germany, pp. 132–144.

Marketos, G., Frentzos, E., Ntoutsi, I., Pelekis, N., Raffaetá A. & Theodoridis, Y. (2008), Building Real-world Trajectory Warehouses, *in* 'MobiDE '08: Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access', ACM Press. New York, NY, USA, pp. 8-15.

Rodriguez Achung, M. (1994), Crecimiento urbano de Iquitos: condicionamientos estructurales en la decaca del ´70 y sus perspectivas. Documento técnico No 08, IIAP, Iquitos, Loreto, Peru, 103 pp.

Pelekis, N., Andrienko, G., Andrienko, N., Kopanakis, I., Marketos, G. & Theodoridis, Y. (2012), Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems* 38, 343-391.

H2Network: A tool for understanding the influence of urban mobility plans (UMP) on spatial accessibility

Adam Gouge, Erwan Bocher, Nicolas Fortin, and Gwendall Petit

Institut de Recherche en Sciences et Techniques de la Ville (FR CNRS 2488),
Ecole Centrale de Nantes, Bâtiment T, 44321 Nantes, France

Abstract: We examine the impact of Urban Mobility Plans (UMP), imposed by the LAuRE law of December 30, 1996, on spatial accessibility to principal destinations, such as hospitals or schools. We use the open-source Geographical Information System OrbisGIS, paired with its Open Geospatial Consortium complaint spatial database H2 and its extensions H2GIS and H2Network, to produce accessibility indicators. We cross these indicators with demographical data to produce service area maps and numerical statistics. These indicators may be used to evaluate the effectiveness of UMPs.

Keywords: Urban mobility plans (UMP), spatial accessibility, shortest paths

1 Introduction

The Law on air and the rational use of energy (LAuRE), published on December 30, 1996, states that *everyone has the right to breathe air that does not harm their health.* In particular, it requires cities of more than 100,000 inhabitants to establish an Urban Mobility Plan (UMP[1]) in order to reduce the impact of public transportation on the environment. This law also requires the UMP to be reevaluated every five years, though it does not specify how.

In collaboration with Nantes Métropole, the *Institut de Recherche en Sciences et Techniques de la Ville* (IRSTV[2]), the French Research Institute in Urban Sciences and Techniques, initiated a research project known as Eval-PDU (2008-2012) in an effort to evaluate the environmental impacts of public policies on public transportation. The

---

[1] http://www.installationsclassees.developpement-durable.gouv.fr/Urban-Travel-Plan-PDU.html, retrieved April, 2014.

[2] http://irstv.fr, retrieved April, 2014.

methodology includes numerous approaches, including numerical simulations evaluating air quality, noise pollution, energy consumption, and greenhouse gas emissions with a chain of physically-based models, based on alternative and comparative scenarios.

In this paper, we examine how to evaluate UMPs geographically. In effect, changing the kinds of transportation offered to the public (e.g., introducing a new tram line or installing a public bicycle system) or modifying the structure of the road network to favor public transportation (pedestrianizing a street, introducing bus lanes) influences how people choose to get around on a daily basis. Everyday activities such as going to work can quickly become a burden if the changes introduced by the UMP are not sufficiently evaluated and their impacts anticipated (the distance from point A to point B could change; access to public services could decrease).

To address these issues, we propose a methodology and an open-source tool which, starting from geographic data representing the public transportation network of a city, allows one to produce indicators representing accessibility to principle destinations, such as train stations, airports, universities, hospitals, gyms, etc. We further clarify these accessibility indicators by crossing them with demographical data, thus quantifying the impact of a UMP on a city's inhabitants by studying their spatial distribution relative to principal destinations.

2 Methodology

The very definition of a UMP, by modifying public transportation options available to citizens as well as traffic flow, implies a restructuring in the way roads are shared. This reorganization has direct repercussions on citizens' transportation choices around and, by extension, on their quality of life. When we realize how much time people spend in public transportation—an hour and a half per day on average for inhabitants of the Ile de France region; over two hours for 20% of them—the vital necessity of tools for quantifying and characterizing the impact of UMPs becomes clear.

A number of travel-related problems can be cast in the language of graph theory, which examines relationships among discrete entities. The entities are represented as a set V of *vertices* or *nodes* and the relationships as a set E of *edges* (Newman 2003). The resulting graph is the ordered pair $G = (V, E)$. In a travel context, the edges are often road segments and the nodes are intersections or points of interest.

A variety of techniques and indicators for studying the structure of the graph are well-known. Some examples include degree distributions, clustering, network correlations, random graph models, shortest path calculations, and closeness and betweenness centrality. (Rodrigue et al. 2013, Newman 2003). In particular, Dijkstra's algorithm provides the solution to finding the most efficient travel route and the information that can be concluded from this calculation, such as identifying the closest facility or defining service areas based on travel time (Rodrigue et al. 2013). Yet these results are not widely implemented in Geographical Information Systems (GIS) or other cartographic contexts.

In this paper, we examine the notion of accessibility, defined as "the measure of the capacity of a location to be reached by, or to reach different locations. Therefore, the capacity and the structure of transport infrastructure are key elements in the determination of accessibility" (Rodrigue et al. 2013). In our case, we study accessibility via three different H2Network functions:

1. `ST_Graph`: From a table of geographical data, we construct a mathematical graph by assigning numerical ids to edges and nodes, resulting in two new tables. An edge may be enriched with additional information that controls how shortest paths are calculated, such as its orientation or weight. For example, one can specify whether a road is one-way or bidirectional, as well as its length or estimated travel time.

2. `ST_Accessibility`: We identify a subset of destination nodes (such as all hospitals). Fix any node v. Calculate the shortest distance from v to each of the destination nodes. Take the minimum of these distances, and designate the corresponding destination as the closest destination. Repeat this for every node v in the graph.

3. `ST_ShortestPathTree`: Calculate the part of the graph which can be reached starting from a given start node and limited by a given time or distance.

We use the information obtained during these calculations to produce a service area map, for various services such as train stations or schools. This map includes isolines whose intervals are expressed in metric or temporal units. Finally, the isoline map is crossed with demographical data in order to quantify the number of inhabitants affected by changes in the road network introduced by a UMP. Using this information, we can estimate, for example, the percentage of inhabitants situated at a distance of less than 500 meters from a train station, or at a travel time of greater than 20 minutes from a school.

3   Implementation in H2Network

The methodology described above is implemented in the open-source Geographical Information System OrbisGIS (Bocher & Petit 2012), which uses the relational database management system (RDMS) known as H2[3]. In 2013, the OrbisGIS team made the choice to switch to H2 from the previous database they had invented, called GDMS (Bocher et al. 2008). This choice brought greater modeling flexibility, improved performance and better interoperability with other databases. When the choice was made, H2 did not yet have a Geometry data type. The team collaborated with the H2 team to add basic support for Geometries, and eventually created their own open-source spatial extension of H2 called H2GIS[4].

H2GIS contains multiple spatial functions, including those specified by the OGC's Simple Features for SQL (SFSQL) 1.2.1 standards (Herring 2010, 2011). It includes additional spatial functions the team has developed that extend and complement the SFSQL functions. Examples include:

- `ST_CompactnessRatio`, which computes the perimeter of a circle whose area is equal to a given polygon's area, and returns the ratio of this computed perimeter to the polygon's perimeter (useful for urban morphology), and

- `ST_Delaunay` and `ST_ConstrainedDelaunay`, which calculate (constrained) Delaunay triangulations (useful in hydrology).

Both H2 and H2GIS are fully written in Java. They can be used in embedded mode in order to give spatial database capabilities to any Java application with a small mem-

---

[3]http://www.h2database.com, retrieved April, 2014.
[4]http://www.h2gis.org, retrieved April, 2014.

ory footprint. Moreover, H2GIS is based on a robust and well-known open source library known as the JTS Topology Suite[5], which provides an API for modeling and manipulating 2-dimensional linear geometry, as well as numerous geometric predicates and functions.

H2Network is the network analysis component of H2GIS. It provides functions for constructing a mathematical graph from geographical data representing a road network and performing shortest path calculations and other network analysis on the resulting graph. Some of its functions were mentioned in the previous section.

4   Case study

We take as study area the 24 municipalities of Nantes Métropole (566 500 inhabitants). The road network is modeled as 53,616 segments (polygonal chains); the data comes from the national topographical database BD TOPO® produced by the *Institut National de l'Information Géographique et Forestière* (IGN[6]), the French National Geographic Institute. The study area is divided into 229 IRIS zones, which are defined by the *Institut national de la statistique et des études économiques* (INSEE[7]), the French national statistics institute, to be statistical units comprising a population of 2,000 inhabitants. We represent them as polygons.

We study accessibility with respect to three scenarios:

1. Base case: The original road network as defined by the BD TOPO® data.
2. Busway: Scenario 1 with a new central bus line (the Busway) crossing Nantes from north to south and reducing light vehicles' access to the city center.
3. Bridge: Scenario 2 with a new bridge (the Eric Tabarly Bridge) which joins two major neighborhoods of Nantes, Ile de Nantes and Malakoff, otherwise naturally separated by the Loire river.

Overall, our results show a spatial redistribution of distances, which are notably more pronounced in Scenario 2. In effect, the accessibility calculations for light vehicles (cars) in this scenario show an increase in distances when approaching the center, as compared to Scenario 1. If we relate these distances to the inhabitants, our analysis shows that the number of inhabitants affected is relatively more important due to the fact that neighborhoods near the city center are more densely populated.

Invitation

We would like to invite you to attend and participate in our workshop entitled *Road network analysis with H2Network: Applications of the spatial database H2GIS*. We will demonstrate and apply the various analyses evoked in this paper.

---

[5] http://sourceforge.net/projects/jts-topo-suite, retrieved April, 2014.
[6] http://www.ign.fr/, retrieved April, 2014.
[7] http://www.insee.fr, retrieved April, 2014.

www.ogrs-community.org

Acknowledgments

References

Bocher, E., Leduc, T., Moreau, G. & González Cortés, F. (2008), GDMS: an abstraction layer to enhance spatial data infrastructures usability, *in* '11th AGILE International Conference on Geographic Information Science-AGILE 2008'.

Bocher, E. & Petit, G. (2012), 'OrbisGIS: Geographical information system designed by and for research', *Innovative Software Development in GIS* pp. 23–66.

Herring, J. (2010), OpenGIS Implementation Specification for Geographic information - Simple feature access - part 2: SQL option, Technical report, Open Geospatial Consortium.

Herring, J. (2011), OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture, Technical report, Open Geospatial Consortium.

Newman, M. E. (2003), 'The structure and function of complex networks', *SIAM review* **45**(2), 167–256.

Rodrigue, J.-P., Comtois, C. & Slack, B. (2013), *The geography of transport systems*, Routledge.

---

[8]http://belgrand-gebd.ifsttar.fr, retrieved April, 2014.

OTB, remote sensing image library: past and future

Manuel Grizonnet, Julien Michel[1] and Mohammed Rashad KM[2]

[1]CNES - DCT/SI/AP, 18, avenue Edouard Belin, 31401 TOULOUSE CEDEX 09, France
[2]Lab for Spatial Informatics, IIIT Hyderabad, India (work performed during CNES internship)

Abstract: The open source image processing library ORFEO ToolBox (OTB) was initiated by the French Space Agency in 2006 in the frame of the ORFEO Pliades program. OTB is based on numerous open-source libraries which are integrated in the OTB framework giving a homogeneous interface for the user. OTB is available in QGIS through the Sextante plugin. The OTB project will be made in near future a more community governed with an official steering committee and public membership.

Keywords: ORFEO ToolBox, CNES, Monteverdi, Pleiades

---

1 Introduction

The open source image processing library ORFEO ToolBox (OTB) was initiated by the French Space Agency in 2006 in the frame of the ORFEO Pliades program. This presentation will give an overview of the project evolution and expectations for the future.

2 OTB: methodological component of the ORFEO program

The main objective of the OTB at the beginning in 2006 was to capitalize new results of the methodological research to ORFEO thematic groups which prepared the arrival of Pliades satellites (launched in 2011 and 2012). ORFEO constellation is a dual Earth observation system with metric resolution, in which Pliades (France) is the optic component and Cosmo-Skymed (Italy) is the radar component.

This component is made of two "small satellites" (mass of one ton) offering a spatial resolution at nadir of 0.7m and a field of view of 20km. Their great agility enables a daily access all over the world, which is a critical need for defence and civil security applications, and a coverage capacity necessary for the cartography kind of applications at scales better than those accessible to SPOT family satellites.

Figure 1: Right angle and parallel lines detection on harbor area

OTB was developed in the frame of the methodological component of this program and aim to provide cutting edge algorithms for remote sensing image manipulation (Christophe & Inglada 2009).

Among other, OTB provides:

- Image access: optimized read/write access for most of remote sensing image formats, meta-data access, visualization;
- standard remote sensing pre-processing: radiometric corrections, orthorectification;
- Filtering: blurring, denoising, enhancement;
- Feature extraction: interest points, alignments, lines;
- Image segmentation (Michel et al. 2012): region growing, watershed, level sets, mean-shift;
- Classification: K-means, SVM, Markov random fields;
- Change detection;
- Information extraction for integration in GIS and mapping systems.

These functions are documented in the OTB Software Guide (ORFEO ToolBox Development Team 2014*b*) which integrates description of algorithms and methods with source code examples.

To encourage full access to the details of all algorithms, CNES decided to distribute the OTB as an open source library. OTB is distributed under the CeCILL v2 license (compatible with GPL). This way CNES hopes to benefit from the contribution of many specialists to help grow the practical use of satellite imagery in general.

3   Different access for different users

The OTB C++ library targeted initially mostly developers but moved slightly to a more user oriented project which dedicated tools build upon the core library (Monteverdi, OTB Applications).
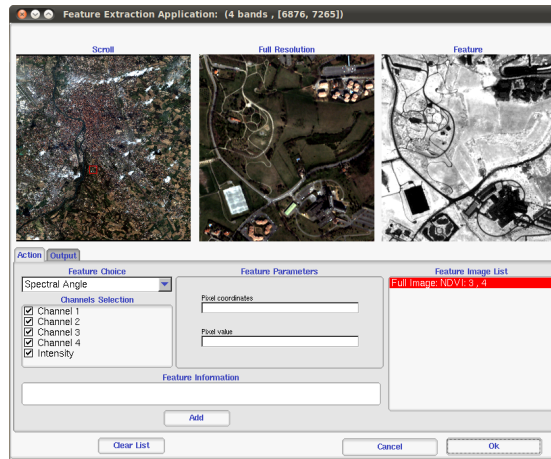
www.ogrs-community.org

Figure 2: Feature extraction module in Monteverdi

In 2009, the Capacity Building program leaded by the CNES Department of Strategy and Programs funded the development of an end-users-oriented software on top of the Orfeo ToolBox, named after the composer of the Orfeo opera, Monteverdi. This new software provided a simple graphical interface to visualize and dynamically chain most of the processing offered by the OTB in a user-friendly environment.

While initially designed for training courses and capacity buildings activities, Monteverdi usage grew far beyond this initial scope and met a great success for OTB users in the ORFEO community, encouraging new non-developing users to gain interest in OTB. During the 2010 to 2012 period, Monteverdi became one of the main platforms for integration of OTB functions and a lot of modules have been developed, including most of the old OTB-Applications which have been integrated to Monteverdi.

The release of Monteverdi 1 at the end of year 2009 changed a lot of things as most of non-developer users, who were looking for quite a long time for an application providing Orfeo TooBox algorithms under a unified graphical interface. As a consequence, Monteverdi continued to grow in the past three years and it has now more than 40 modules available. However, if Monteverdi 1 was indeed a great tool to perform numerous remote sensing and image processing tasks provided by OTB in a minute, it was not well suited for heavier and longer processing, scripting and batch processing. The implicit processing pipeline that the users can build by chaining modules can become very heavy to update, resulting in freezes or lags, a big issue with Pliades images which can be very large. Last, the user interface could be better, with a lot of mouse operations needed to perform some actions and no history or undo functions.

4 OTB applications : one code, multiple targets

OTB Applications was perhaps the older package of the Orfeo ToolBox suite after the OTB package itself. Since the Orfeo ToolBox is a library providing remote sensing functions, the only applications that were distributed at the beginning were the exam-
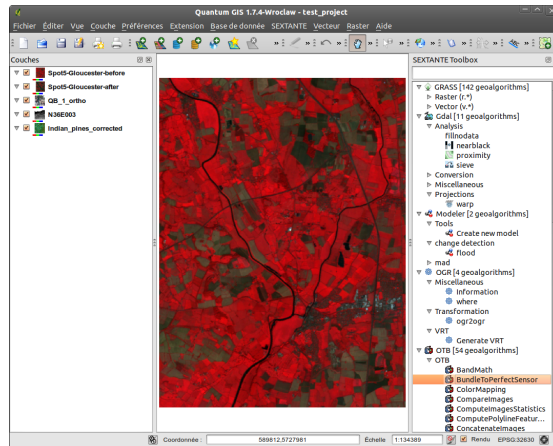
Figure 3: Access to OTB applications through Quantum GIS processing

ples from the Software Guide and the tests. These applications are very useful for the developer because their code is very short and only demonstrates one functionality at a time.

In many cases, a real application would require :

- Combining together two or more functions from the Orfeo ToolBox
- Providing a nice high level interface to handle : parameters, input data, output data and communication with the user

The main idea of this framework is the following: each application is written once for all in a shared library (also known as plugin). This plugin can be auto-loaded into appropriate tools without recompiling, and is able to fully describe its parameters, behavior and documentation (ORFEO ToolBox Development Team 2014*a*) and also save the state of all the parameters of an application in an XML file and run the application again with the same set of parameters, overriding just the one you want.

The tools to use the plugins can be extended, but Orfeo ToolBox shipped the following:

- Command-line launcher
- A graphical launcher, with an auto-generated QT interface, providing ergonomic parameters setting, display of documentation, and progress reporting,
- A SWIG interface, which means that any application can be loaded set-up and executed into a high-level language such as Python or Java for instance.

Moreover OTB is available in QGIS through the Sextante plugin. Sextante provides access to a huge number of processing modules, and acts as a geo-processing environment that can be used to call native and third-party algorithms from QGIS.

5   Monteverdi2

Monteverdi success went far beyond this initial scope since it opened the OTB world to a wide range of users who needed a ready to use graphical tool more than a library of
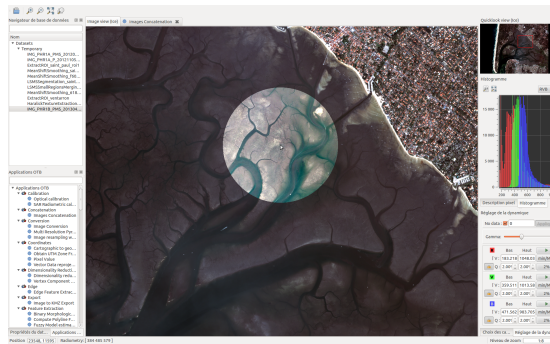
www.ogrs-community.org

Figure 4: Monteverdi2 main window

components to write their own processing chain. With this 4 years of lifetime, we have a lot of feedback regarding how useful the tool was, but also regarding what should be improved to move toward greater usability. We therefore decided to rework the Monteverdi concept into a brand new software, enlightened by this experience.

The development started in 2012 with regular release and an iterative development process whit short loop which include user feedbacks.

We planned to continue to improve Monteverdi2 version in 2014 by integrating (among other things) a more efficient image viewer using GPU shading and rendering.

6   And now?

OTB is based on numerous open-source libraries which are integrated in the OTB framework giving a homogeneous interface for the user. In particular, the core of the OTB system is coming from ITK (Johnson et al. 2013) which has proven its efficiency for medical image processing. Most of the algorithms for segmentation and registration have been well tested in this context.

ITK move recently to a new major version called the 4.0 series (simply refer as ITKv4).The ITK project went through a refactoring process to revise, simplify and accelerate the library.

The OTB team worked hard to make the next OTB release compatible with this new version. This migration suppose major API changes but also brings a lot of new functionalities to the library.

ITKv4 refactoring have also separated the library into a modular structure which now contains over 100 modules. Modularization is a common design answer to the problem of managing complexity and growth and ITK classes are now organize into conceptual and cohesive groups. It has also facilitated the use of add-ons with ITK without having to include them in the toolkit. Modularization is something that the OTB development team would like also to address in the next versions.

The end of the ORFEO program will mark a major change in the OTB history as it was mostly funded by the Pliades project since 2006. Even if CNES decide to continue the development of the library in the next years, the project will necessarily evolve to target new missions developed by the CNES in particular future high resolution optical satellite.

Finally, CNES believe that it is in the best interest of the OTB project at this time to turn the current organization to a more community governance organization with an official steering committee, and to make public its membership.

CNES think that it should help to guarantee that OTB remains company neutral and open and planned to initiate this organization in 2014.

References

Christophe, E. & Inglada, J. (2009), 'Open source remote sensing: Increasing the usability of cutting-edge algorithms', *IEEE Geoscience and Remote Sensing Newsletter* **March**, 9–15.

Johnson, H. J., McCormick, M. & Ibanez, L. (2013), The ITK software guide, third edition. updated for ITK 4.5, Technical report, The Insight Software Consortium.

Michel, J., Grizonnet, M., Jaen, A., Harasse, S., Hermitte, L., Guinet, J., Malik, J. & Savinaud, M. (2012), Open tools and methods for large scale segmentation of Very High Resolution satellite images, *in* 'OGRS2012 Symposium proceedings', OGRS, pp. 179–184.

ORFEO ToolBox Development Team (2014*a*), The ORFEO ToolBox cookbook, a guide for non-developers, updated for OTB-4.0, Technical report, Centre National d'Etudes Spatiales (CNES).

ORFEO ToolBox Development Team (2014*b*), The ORFEO ToolBox software guide, updated for OTB-4.0, Technical report, Centre National d'Etudes Spatiales (CNES).

A participatory WebGIS platform to support biodiversity inventory in the Geneva cross-border area

Stéphane Joost[1], Raphaël Baumann[2], Olivier Ertz[2],
Jens Ingensand[3], Ivo Widmer[1], and Daniel Rappo[2]

[1] Laboratory of Geographical Information Systems (LASIG), School of Architecture, Civil and environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[2] School of Management and Engineering Vaud (HEIG-VD), Institut des technologies de l'information et de la communication, Département COMEM+ (COMmunication-Engineering-Management), Yverdon-les-Bains, Switzerland
[3] School of Management and Engineering Vaud (HEIG-VD), Institut géomatique, gestion de l'environnement, construction et surveillance d'ouvrages (G2C), Yverdon-les-Bains, Switzerland

Abstract: The context is the URBANGENE project which aims at estimating the impact of urbanization on biodiversity using genetic information. The paper discusses the opportunities to involve citizens in a conservation process through the development of a participatory platform based on user-friendly, specially adapted and reusable crowdsourcing tools.

Keywords: Urban biodiversity, citizen Science, crowdsourcing, participatory WebGIS, social media

1 Introduction

Green spaces and biodiversity have an extremely important role in urban areas. They have a clear impact on water flow, on micro-climate regulation, on air quality, and on the health of residents. In recent years, the growth of the population combined with the reduction in the number of persons per household has led to a high consumption of green areas (Rietsch 2014). As there is empirical evidence showing that interacting with nature in cities delivers measurable benefits to people (Aldous 2007, Fuller & Irvine 2010), this progressive urbanization implies a strong

fragmentation of semi-natural environments, what clearly endangers urban biodiversity and causes decline in the quality of life of the population.

The URBANGENE project (urbangene.epfl.ch) will estimate the impact of urbanization on biodiversity using genetic information. Similarly to a study carried out in Marseille in 2011 (Schoville et al. 2013) it will measure the diversity and biological connectivity in three species (one plant, one butterfly, and one amphibian) in order to obtain objective information (quantitative measures) on their ability to migrate between habitats and to adapt to the urban environment. In the context of the planning and of the management of the urbanization process, important goals of the project are to establish a diagnosis of the current state of biodiversity dynamics and identify the ecological, socioeconomic and sociodemographic factors influencing gene flow and adaptation of plant and animal species in urban areas.

Six-hundred individuals of the broad-leaf plantain plant species were sampled during summer 2013, and meadow brown or small white butterflies will be sampled during summer 2014. As regards amphibians, the sampling phase took place during spring 2014. But in this case, it was necessary to target private ponds to complete the official list of sampling areas provided by the Swiss Confederation and by the State of Geneva. To this end, a WebGIS-based participatory platform was developed in order to make it possible for Geneva residents to indicate on an interactive map the geographic location of ponds, and to transmit information on the species they were able to observe in this area.

2   Crowd-sourcing for biodiversity inventory

The recent evolution of the Internet has generalized crowd-sourcing approaches also in the field of geographic information through Public Participation Geographic Information Systems (PPGIS) initiatives (Merrick 2003, Elwood 2006). After the first age of the Web was about one-directional publication of informational content, the second age started with the ability to comment on blog posts and is going on its way by the involvement of web users in the production of raw data and of structured content. OpenStreetMap is one famous and relevant project, a collaborative and community-based action for digitizing and mapping the world. We may notice that such actions are often associated with open data licensing.

Related to the field of biology and biodiversity monitoring which corresponds to the URBANGENE project, an interesting example is The Atlas of Living Australia (www.ala.org.au). The type of collective effort to constitute such an atlas is named "citizen science", which is a systematic collection and analysis of data, the development of technology, the testing of natural phenomena and the dissemination of these activities by researchers on a primarily avocational basis, as defined by the OpenScientist website (http://www.openscientist.org/). As mentioned by Cohn (2010), the collaboration between scientists and volunteers is likely to enhance the ability to collect scientific data, but it is also expected that public

www.ogrs-community.org

participants may contribute valuable information as they learn about biodiversity (in the present case) in their local communities.

Anyway, in such contexts, it is key to rely on robust computer platforms and on user-friendly applications to facilitate the task of contributors when they input their data (Haklay & Tobon 2003). From a technical point of view, even if such crowd-sourcing platforms may be based on classical Content Management Systems (CMS), they are usually tailor-made according to the investigated topic. Moreover, we must admit that there is a plethora of plug-ins to publish maps based on CMS plug-ins, but there are currently only few CMS offering integrated and ready-to-use crowd-sourcing solutions (Palazzolo & Turnbull 2011).

3   The URBANGENE web-based PPGIS platform

Given the above technological state, the present project aims at exploring possibilities to develop a CMS oriented approach based on a generic plug-in. The advantage of such a plug-in would be to provide an adaptable and configurable tool to match the needs of any underlying investigated topic (e.g. configuration of a custom geodata-type, editing tools for point/line/polygons, survey builder, etc). The developments will benefit from the concrete URBANGENE experimentations and help to lay down the relevant questions useful to frame the adaptability of a generic plug-in. To this end, a first step consisted of the creation of a standalone web application (Figure 1) based on a HTML5/CSS3 development making use of an open source stack of JavaScript libraries providing useful online mapping standard features. Indeed, Leaflet - the core mapping framework - is a modern library for user-friendly interactive maps, it is designed with simplicity, performs well and gives importance to usability.

Leaflet is able to access different base maps (e.g. Openstreetmap, etc.), to organize information layers, to handle geographic objects, to pan and zoom, etc. This library has several plug-ins or controls, among which the "Leaflet markercluster" enabling the aggregation of points when a large geographic scale is displayed. Leaflet also permits to access third party data (e.g. Google geocoding service) and facilitates user input with an auto-completion functionality. The jQuery UI library permits to comfortably manage the elements of the user interface. It is used in particular to provide and manage interactions like the "drag & drop" functionality also available on tactile devices (jQuery UI Touch Punch). Finally, the jQuery.kinetic plugin improves user experience through fluid map manipulations.

On the server side, PHP is used while collected data are stored in a database managed by PostgreSQL/PostGIS. The prototype developed here implemented a very simple data model. A single table was used to collect the coordinates (X, Y) of a pond, the name of the species observed around, a type of pond (private or public), an email address in the case the owner provides the authorization to sample in his private property, a name (user inputs are anonymous with the opportunity to

provide a name or a pseudonym to appear in the list of acknowledgments) and an optional section with 6 short general questions on biodiversity perception.

Open source software components are used both at client and server sides because it allows to access the underneath models and algorithms without any limitation (Rey 2012). This is especially relevant for the evaluation of the Leaflet JavaScript library in comparison to the famous OpenLayers. Both being became mapping librairies of reference (Steiniger & Hunter 2013), the first is relatively recent and promosing while the second is currently progressing towards a third major release. Indeed, the capacities of the underlying cartographic model and the flexibility to control user interactions are critical requirements for the creation of a user-friendly crowdsourcing application. Also, the level of extensibility is an important aspect to consider for a generic plug-in approach. Such an exploration and evaluation helps to choose the adequate webmapping library able to answer the objectives at best. It will be pursued during the next steps of the project. Finally, the project itself is planned to be released under a free license and managed as a GitHub project (github.com/comem/urbangene).

In a successive phase of the project, we will evaluate how existing plugins - e.g. WordPress Form Maker as a survey builder, Drupal Geofield to create geodata-types - or other components may be integrated so as to create specific crowd-sourcing features.

Figure 1: User interface of the URBANGENE Web PPGIS application. The map shows the Geneva cross-border area (the border between France and Switzerland appears in grey). On the right, the panel displays observed species to be indicated by users and additional questions.

4   Tentative results and conclusion

In March 2014, advertised by efficient media coverage (Overney 2014), the URBANGENE PPGIS web-platform gave the opportunity to Geneva cross-border area residents to participate in the inventory of private ponds. This permitted to collect an initial list of 99 target areas (on April 23, 2014) likely to facilitate the task of biologists. Indeed, collaborators were able to start the sampling of common toads along 5 main urban development directions. In a successive step, DNA will be extracted, and then sequenced to enable the computation of genetic diversity indices to be compared with the urbanization process in the metropolitan area.

As the urban natural environment and its biodiversity belong to the residents, such an action is a good opportunity to give people a chance to get involved in a conservation process. The feedback received during the first month of data collection shows that the project called the attention of urban residents on an important but often unrecognized issue in urban planning. It also created an involved community in the domain (www.facebook.com/urbangene).
From a technological point of view, the development of a generic CMS plugin will facilitate the adaptation of the application to various use cases, and favour its reuse for the deployment of future biodiversity PPGIS initiatives.

References

Aldous, D.E. (2007), Social, environmental, economic, and health benefits of green spaces. *In* Proceedings of the International Symposium on Horticultural Plants in Urban and Peri-Urban Life, 171-85.

Cohn, J. P. (2008), Citizen Science: Can Volunteers Do Real Research? *BioScience* **58**, 192–197.

Elwood, S. (2006), Critical Issues in Participatory GIS: Deconstructions, Reconstructions, and New Research Directions, *Transactions in GIS* **10**, 693–708.

Fuller, R.A., Irvine, K.N. (2010) Interactions between People and Nature in Urban Environments. *In* Urban Ecology, Gaston, K.J., Ed., Cambridge University Press: Cambridge, UK, pp. 134–171.

Haklay, M., Tobón, C. (2003), Usability evaluation and PPGIS: towards a user-centred design approach, *International Journal of Geographical Information Science* **17**, 577–92.

Merrick, M. (2003), Reflections on PPGIS: A view from the trenches. *URISA Journal* **15**, 33–40.

Multispecies investigation of the impact of the urbanization process on genetic diversity in the Geneva cross-border area,
   URL: *http://urbangene.epfl.ch/* [accessed on January 31th, 2014]

Overney, J. (2014), Contribute to Science by Locating Geneva's Wildlife Habitats. EPFL News.

URL: *http://actu.epfl.ch/news/contribute-to-science-by-locating-geneva-s-wildl-4/* [accessed 12.5.2014].

Palazzolo, A., Turnbull, T. (2011) *Mapping with Drupal, Navigating Complexities to Create Beautiful and Engaging Maps*, O'Reilly Media, Cambridge (MA).

Rietsch, R. (2014), Informations Statistiques n° 4 – Janvier 2014,
URL:
http://www.ge.ch/statistique/tel/publications/2014/informations_statistiques/autres_them es/is_population_04_2014.pdf

Schoville, S. D., Widmer, I., Deschamps-Cottin, M., Manel, S. (2013) Morphological Clines and Weak Drift along an Urbanization Gradient in the Butterfly, Pieris rapae, *PLoS ONE* **8**, e83095.

Rey, S. J. (2012), Open source spatial analysis: lessons for research and education from PySAL, *in* OGRS2012: Proceedings of the second Open Source Geospatial Research & Education Symposium, Lulu.com Self-publishing, Olivier Ertz, Stéphane Joost, Marj Tonini, Yverdon-les-Bains, Switzerland, pp. 66–71

Steiniger, S., Hunter, A. J. S. (2013) The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption, *Computers, Environment and Urban Systems* **39**, 136–50

www.ogrs-community.org

Reititin: an open source tool for analysing accessibility by public transport in Greater Helsinki

Juha Järvi[1], Maria Salonen[2], Perttu Saarsalmi[2], Henrikki Tenkanen[2], and Tuuli Toivonen[2,3]

[1]BusFaster Ltd. juha@busfaster.com
[2]Department of Geosciences and Geography, P.O. Box 64, 00014 University of Helsinki
[3]Department of Biosciences, P.O. Box 65, 00014 University of Helsinki

Abstract: This paper presents a public transport routing system suitable for complicated reachability analyses. Its main goals are to be flexible, easy to extend and allow creating public services where large numbers of people can calculate routes and examine reachability of different locations, with no server software or installation needed.

Keywords: Routing, public transport

1 Introduction

Accessibility analysis is considered an appropriate way to assess interactions between transportation and land use (Bertolini et al. 2005, Silva & Pinho 2010). Different distance measures typically form an integral part of accessibility indicators. As one such measure, travel time is considered to correspond well to people's perceptions of friction of distance (Frank et al. 2008, Mavoa et al. 2012). Traditionally travel time has been calculated using privately owned car as the subject. Concern over the environmental and social sustainability of land use and transportation solutions has in recent years highlighted the need to incorporate different modes of transport in accessibility analyses. Particularly, there is a need to properly analyse accessibility using public transportation (Salonen & Toivonen 2013). Until recently, however, adequate data and tools for detailed public transport analysis have been lacking.

During the last years Finland, as many other countries in Europe, has witnessed a fast development in data availability. Helsinki region has been one of the active promoters of open administration. As part of this process several important spatial and non-spatial data resources have been opened for public use. One of the most used resources has been the public transportation Route Planner service (Reittiopas), provided by Helsinki Regional Transport Authority (HSL). The route and schedule database underlying Route Planner is accessible via an application programming interface (API) and through a dump file in Kalkati.net XML format. This type of data has proven to be useful not only for application developers but also for researcher and urban planners interested in assessing the spatial patterns of public transport accessibility.

The MetropAccess-project (University of Helsinki, Department of Geosciences and Geography) aims at developing comparable tools for analysing spatial patterns of accessibility and travel time using different travel modes. In this paper, we present one of our key analysis tools, Reititin, which is developed in collaboration with BusFaster Ltd. The tool calculates optimal public transportation travel chains for extensive sets of origin-destination pairs using the Kalkati.net XML dump provided by HSL and OpenStreetMap data.

2 Technology

The tool is written in the JavaScript language to work on all common operating systems and with the goal to embed it on a website, allowing any interested person to make route calculations en masse. Large analyses with thousands of routing calculations using different combinations of starting locations, times, enabled transport modes and other options can take hours to complete. It would be impractical to allow free unrestricted access for everyone to execute such queries on our servers, while we'd also like to allow casual users to make analyses for example to select a hotel with good connections to specific locations. Users are unlikely to install software on their own computer just to see if the results would be useful to them. If the calculations are instead performed in a web browser, the entire service can be provided through any low-cost or free online file storage service. Then there is practically no limit to the number of concurrent users and each user is able to run analyses as complicated as their own computer is able to handle. Results can be shared online together with selected analysis options encoded in a link to the service, allowing the analysis to be modified and repeated by others.

The web browser environment is constrained in terms of memory and calculation speed, and the entire street network and public transport schedule must be transferred to the user for batch routing. For this, data files must be kept small to reduce the downloading time.

2.1 Algorithms

Schedules are compressed by identifying redundancies between different days of the week, delta coding and using the LZ77 algorithm (Ziv & Lempel 1977). This compression shrinks both memory usage and data file size. The road map is simplified and topology stored together with delta coded geometry by referring to existing road points when a road intersects another previously stored road. This processing can reduce Kalkati or Google Transit Feed Specification format to under 2% and OpenStreetMap pedestrian routing maps to under 20% of original compressed size without affecting routing results. This process lasts a few minutes for the Helsinki metropolitan region.

In the routing, optimal single source shortest paths are calculated using Dijkstra's algorithm (Dijkstra 1959) modified to track time of the day in addition to the cost being optimised. The road network for pedestrian routing and public transit route network are connected at public transit stops, which are conceptually a part of the road network. Points along public transport routes passing by the stop are all considered separate locations.

Cost of moving from a routing graph node representing the sidewalk near a bus stop to another node representing a bus at the stop depends on the waiting time, which can be computed only when the pedestrian's earliest possible arrival time at the stop is known. Therefore many graph edge costs are computed only when the best route to their start point is found. Instead of storing locations in the priority queue used by Dijkstra's algorithm, we store objects encapsulating several types of information.

Events such as arriving at an intersection or stop are represented as JavaScript objects representing a visit at a location and containing the event's cost, time, location in the routing graph, previous location and a visit method. Routing is based on traversing the priority queue and calling visit methods of events found. Costs and arrival directions for all locations are kept in a separate structure without modifying the routing graph allowing multiple parallel route searches. Different modes of transport are represented by subclasses of visitor objects making them easy to modify independently of each other.

In object oriented programming terms, there are classes for visitors for transit vehicles, transit stops and road network nodes. Adding a new type of transportation is relatively simple. For example a bicycle sharing system could be represented by adding a new type of road node visitor that does not allow entering bus stops, has faster connections between nodes and is connected to the regular road network only at bike lending points.

Routes are optimised for cost represented in increments of 100 milliseconds spent sitting inside. The cost of walking is then typically set higher than actual elapsed time due to the effort spent and exposure to the weather. Transfers between public transport lines are penalised by 3 minutes in addition to actual waiting time because they add uncertainty and risk to the user if a vehicle is delayed or arrives

too early, or the user makes a mistake while trying to find the correct location on time.

The priority queue in Dijkstra's algorithm is simply an array with a slot for each possible time step, which can contain one or more events stored in a linked list. If maximum route cost is 5 hours, array length in 100 millisecond increments is 180 000 items regardless of road network size. Complexity of the algorithm is O(V+E+T) where V is the number of road intersections or public transport stops, E the number of connections between them and T the maximum route cost. The discrete time steps speed up calculations for large cities while causing an average error of 50 milliseconds at every intersection, which is negligible compared to other random factors such as traffic lights.

2.2   Output

To produce Pareto optimal routes in large batches, the calculation is repeated with several initial departure times. The latest possible departure time still giving the earliest possible arrival is then found for each origin-destination pair. This way producing an OD matrix takes a very small constant lookup time per element and routing time scales linearly with road network intersections, number of origins and different departure times tested. Calculating routes from a single origin and departure time to any number of locations in the Helsinki metropolitan area takes about 3 seconds in a web browser.

As an output, Reititin produces detailed travel chain descriptions between the origin and destination coordinates. The descriptions include information on travel times and distances for each mode of public transport and necessary walking stretches. Also all stops and lines are listed. The tool enables creation of new public transportation routes and routing along them. This feature can be used to analyse alternative transport scenarios for future, and testing the effect of potential new public transport routes. It is also possible to conduct temporal analyses by comparing travel time patterns for current and older routes and schedules.

3   Conclusion

The outputs of routing analyses can be and have been used for various purposes. The basic outputs – travel time matrices - enable detailed analyses of service accessibility (Jäppinen et al. 2013, Salonen et al. 2012, Salonen & Toivonen, 2013) and its temporal variation (Saarsalmi 2014). Furthermore, the detailed travel chain descriptions make it possible to assess theoretical CO_2 emissions for each route (Lahtinen et al. 2013).

In order to make the most use of the positive trend of open data policies, it is essential that tools that are created to handle the data sources are also open access. Reititin provides an example of such development and it will hopefully provide a useful tool for urban planners of the region.

www.ogrs-community.org

References

Bertolini, L., le Clercq, F. & Kapoen, L. (2005), 'Sustainable accessibility: a conceptual framework to integrate transport and land use plan-making. Two test applications in the Netherlands and a reflection on the way forward', *Transport Policy* **12** (3), 207–220.

Dijkstra, E. (1959), 'A note on two problems in connexion with graphs', *Numerische Mathematik* **1**, 269–271.

Frank, L., Bradley, M., Kavage, S., Chapman, J. & Lawton, T.K. (2008), 'Urban form, travel time, and cost relationships with tour complexity and mode choice', *Transportation* **35** (1), 37–54.

Jäppinen, S., Toivonen, T. & Salonen, M. (2013), 'Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: An open data approach', *Applied Geography* **43**, 13-24.

Lahtinen, J., Salonen, M. & Toivonen, T. (2013), 'Facility allocation strategies and the sustainability of service delivery: Modelling library patronage patterns and their related CO2-emissions', *Applied Geography* **44**, 43-52.

Mavoa, S., Witten, K., McCreanor, T. & O'Sullivan, D. (2012), 'GIS based destination accessibility via public transit and walking in Auckland, New Zealand', *Journal of Transport Geography* **20** (1), 15–22.

Saarsalmi, P. (2014), 'Päivittäistavarakaupan spatio-temporaalinen saavutettavuus pääkaupunkiseudulla', Pro gradu -tutkielma, Helsingin yliopisto, Geotieteiden ja maantieteen laitos.

Salonen, M., Toivonen, T. & Vaattovaara, M. (2012), 'Arkiliikkumisen vaihto-ehdoista monikeskuksistuvassa metropolissa: Kaksi näkökulmaa palvelujen saavutettavuuteen pääkaupunkiseudulla', *Yhdyskuntasuunnittelu* 2, 50, 8-27.

Salonen, M. & Toivonen, T. (2013), 'Modelling travel time in urban networks: comparable measures for private car and public transport', *Journal of Transport Geography* **31**, 143–153.

Silva, C. & Pinho, P. (2010), 'The structural accessibility layer (SAL): revealing how urban structure constrains travel choice', *Environment and Planning* A 42 (11), 2735–2752.

Ziv, J. & Lempel, A. (1977), 'A universal algorithm for sequential data compression', *IEEE Transactions on Information Theory* **23**, 337–343.

Lahti Urban Laboratory: a web-based map platform for urban planning

Jiri Kadlec[1] and Juhani Järveläinen[2]

[1]Department of Civil and Environmental Engineering, Aalto University, Lahti, Finland
[2]Department of Environmental Sciences, University of Helsinki, Lahti, Finland

Abstract: The Lahti Urban Laboratory interactive map website is a collection of data and tools for visualization and modeling of storm water runoff in the urban environment. The whole platform has been built as a distributed system using geospatial web services (Web Map Service, Web Feature Service) and open source components (PostGIS, MapServer, TinyOWS). This paper presents three use cases of the urban laboratory tools: experimental catchments visualization, land use editing, and surface runoff modeling.

Keywords: Urban planning, catchment, PostGIS, web GIS

1   Introduction

The hydrological cycle in urban areas is strongly affected by urbanization process (Lee et al. 2003). Urbanization leads to an increase of impervious surfaces (such as asphalt roads or car parks) with high runoff to rainfall ratios (Schueler 1994). Water flowing from these impervious surfaces is known as stormwater runoff. In Finland, stormwater runoff is typically not treated by wastewater treatment plants. Instead, it flows through a separate network of pipes and culverts directly into receiving water bodies and lakes (Valtanen et al. 2013). By collecting heavy metals, suspended solids and other pollutants from the impervious surfaces, urban stormwater runoff may have harmful effects on water quality in receiving lakes (Nordeidet et al. 2004).

Sustainable urban planning decisions require extensive online collaboration between urban planners, hydrologists, and public stakeholders (Brooking & Hunter

2013). The hydrologist is responsible for developing and testing a mathematical model that is capable of predicting storm water runoff and pollutant transport under varying land cover conditions (Modaresi et al. 2010). The urban planner is responsible for maintaining up-to-date spatial data including future land cover change scenarios. Finally, the public is concerned about their environment including the urban forests, streams, rivers and lakes. Despite the rapidly increasing availability of geographic information on the web, traditional desktop-based geographic information systems (GIS) and specialized hydrological model software tools are predominantly used in the hydrological model setup, often requiring time-consuming exchange of large datasets between the urban planner, GIS expert and hydrologist and resulting in information loss (Bach et al. 2014). The limitation of desktop GIS can be overcome by moving the system to a distributed web-based information platform (Brooking & Hunter 2013, Wing & Goodwin 2011). This paper explores options how to efficiently visualize the effects of urban planning decisions on stormwater runoff quantity and lake water quality using a web-based online information platform. The demonstration has been tested for Lahti in Finland.

2 Application design and use cases

The web based interactive map application consists of three separate web pages (use cases): Experimental catchments visualization, Land use editing tool, and stormwater runoff calculation tool. The interfaces are readily available in any location with Internet access. Up-to-date source geographical information (urban plans, land cover maps, drainage network) is maintained by different data owners: Lahti city department, Lahti Aqua utilities department, Finnish environmental institute, and Aalto University. It is published by using the Web Map Service (WMS), Web Feature Service (WFS), or GeoJson standard data format and shared in the three front-end interfaces. The "Lahti urban water laboratory" is built completely based on open source software For the front-end interactive maps, the OpenLayers tools are used. For the back-end calculations, spatial overlay operations in the PostGIS database are utilized. In addition, the TinyOWS and MapServer software packages are used to support the land use editing interface.

3 Experimental catchments visualization

As part of hydrological research in Lahti, three experimental observation sites with real-time observations are maintained. The first site is located in a natural forest area with a planned construction. The second site is in a suburban residential area. The third site is in a city centre residential area. The time series data from field sensors are transmitted to a relational database and served using the Sensor Observation Service (SOS) or Javascript Object Notification (JSON) format. The open-source *Flot* Javascript toolkit is then used to display the one, two or three time
www.ogrs-community.org

series of precipitation, discharge or turbidity in a dynamic interactive chart. Schema of components of the observation network is shown in Figure 1. Users can visually compare how three small catchments behave in different types of recent rainfall or snowmelt events.
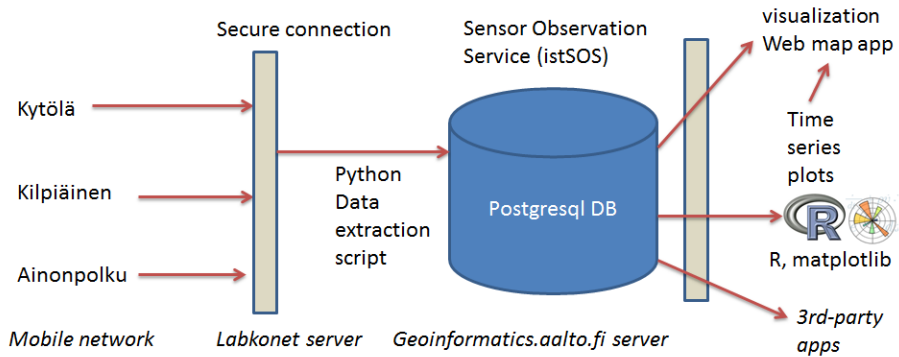

Figure 1: Transfer of real-time observations from experimental stations to the end user interface

4 Land use editing tool

In urban planning, it is important to be able to rapidly compare the impacts of a land use change decision on the environment, including stormwater runoff quantity and quality. The services-oriented architecture of the Lahti Urban Laboratory platform supports interactive editing of all land use maps that are stored as internal data sources. It also supports the creation of new land use maps and their combination with the original maps. The simultaneous editing of maps is enabled using the WFS Transactional (WFS-T) protocol. Let us consider a use case, where a part of an urban forest is planned to be replaces by a low-density residential area. Multiple users can edit the "future land use" map simultaneously over the internet. Both the polygon areas (geometry) and the polygon attributes (land use class) may be edited. An example of using the map editing tool is shown in Figure 2. First, the user selects a land use class (for example "Industrial"). Then, the user draws a polygon in the map for indicating a future industrial area. After the edits are saved, the new polygon is inserted into a database and becomes available as part of the WFS and WMS service and a new "map layer" in the end-user interactive map client interfaces. In one of the interfaces, the user can switch between the old and the new land use scenarios and decide which scenario has lesser impact on stormwater runoff and on lake water quality.

Figure 2: Land use editing tool with land use categories for urban planning. In this example, the user is drawing a new "industrial area" polygon to replace a forest.

5   Stormwater runoff calculation tool

In the stormwater runoff calculation web page, the user can select an interactive map of present or future land use in the Lahti region. The data sources include official data sources (National Land Survey of Finland, City of Lahti), volunteered geographic information (OpenStreetMap), data collected by researchers (Land cover map created by interpreting air photographs and edited by the land use editing tool), and urban planning documents (Strategic plan and detailed area plans), displayed together with a map of lakes and their drainage catchments. The user selects the land use dataset, rainfall amount, receiving lake and indicator: surface runoff volume, sediments, nutrients, or heavy metals. The land use map data from external data sources is accessed through the Web Feature Service (WFS) or Web Map Service (WMS) protocols. In case of WFS, the data must be "polygon" geometry type. In case of WMS, the "GetFeatureInfo" operation must be supported to enable listing of attributes for each land use polygon. Land use map data collected internally by Lahti researchers is stored in a PostGIS database. Stormwater runoff is estimated as a function of rainfall and land use by using the event mean concentration method (Park *et al* 2009). The schema of the calculation is shown in Figure 3.
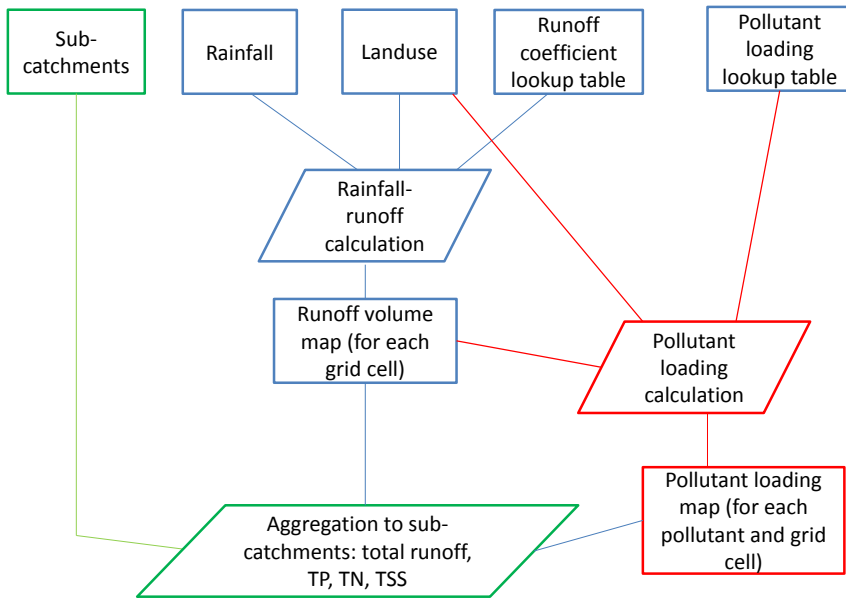
Figure 3: Schema of surface runoff volume and pollutant loading calculation

For each land use data source, a lookup table is created. The lookup table links each land use category with an "impervious area percentage" value and an "event mean concentration" factor. The "impervious area percentage" is the estimated percentage of asphalted roads, roofs, car parks and other impervious surfaces within the polygon. The "event mean concentration" is specified separately for each pollutant and it characterizes the pollutant's typical concentration during a stormwater event. A hydrology expert can edit the lookup table according to local knowledge. The actual runoff coefficient (runoff volume / rainfall volume ratio) is then calculated as a function of rainfall height and the stormwater runoff factor. For rainfall, the rain can be assumed as uniform for the whole region, or as spatially varying. If the rainfall map is spatially varying, then a runoff raster is created by overlaying the rainfall map and the runoff coefficient map. Next, a pollutant loading raster is created for each pollutant indicator (heavy metals, nutrients, suspended solids) by overlaying the runoff raster and the land use polygons. Finally, the runoff raster and pollutant load rasters are overlaid with the drainage catchments map to calculate total runoff volume and total pollutant loading (mass) for each catchment. Additional indicators are also displayed for each catchment, for example the percentage of impervious area. The resulting maps are made available as a dynamic WMS (web map service). The MapServer open source software is used for handling the WMS service. An interactive map client is displaying the map

in a web browser. The client was created using the Javascript and OpenLayers tools. An example of the dynamically calculated output map (stormwater runoff coefficient in catchments in a suburban part of the city) is shown in Figure 4.
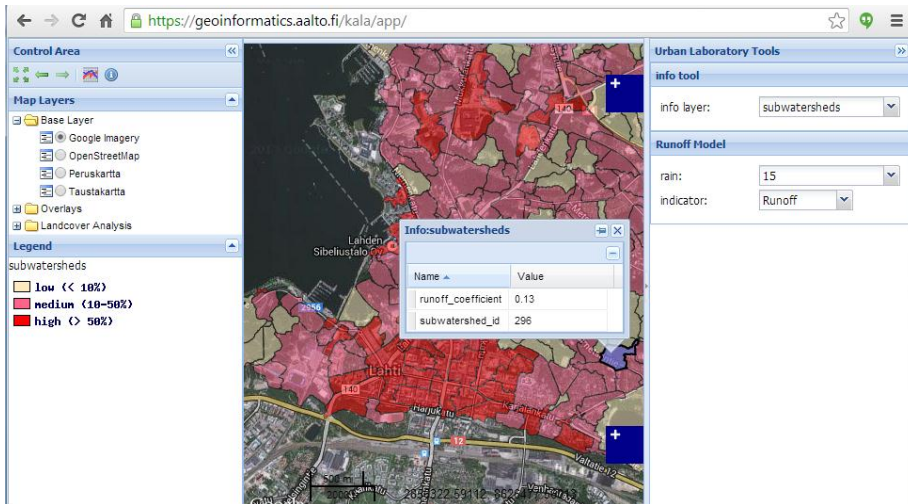

Figure 4: Stormwater runoff coefficient calculated for selected sub-watersheds in the city

6   Discussion and conclusion

The interactive web map application can be used to present alternatives of urban development and rapidly compare and contrast the impact of urban planning decisions on the water environment. In contrast with other similar applications such as SwampGIS (Wing & Goodwin 2011), the "Lahti urban water laboratory" is built completely based on open source software. In particular, the Apache server, PostGIS, MapServer, TinyOWS and OpenLayers technologies are utilized. One potential advantage of using the open-source deployment model is customizability and maintainability. For example, hydrologists can easily modify the built-in event mean concentration model parameters in the look-up tables or modify the built-in rainfall-runoff regression equations using latest results from field experimental sites (Valtanen et al. 2013). In addition, the components of the system are loosely coupled using the services-oriented architecture closely following the Open Geospatial Consortium (OGC) standards. With the increasing availability of geospatial data from the environment and urban planning offices as standard geospatial web services, we expect that it will become simpler to develop similar interactive web-based urban planning tools for other cities in Finland in the near future.

Acknowledgments

References

Bach, P.M., Rauch, W., Mikkelsen, P.S., McCarthy, D. & Deletic, A. (2014), 'A critical review or integrated urban water modelling – Urban drainage and beyond', *Environmental Modelling & Software* **54**, 88–107.

Brooking, C. & Hunter, J. (2013), 'Providing online access to hydrological model simulations through interactive geospatial animations', *Environmental Modelling & Software* **43**, 163–168.

Lee, J.G. & Heaney, J.P. (2003), 'Estimation of urban imperviousness and its impacts on storm water systems'. *Journal of Water Resources Planning and Management* **129**(5), 419–426.

Modaresi, R., Westerlund, C., & Viklander, M. (2010), 'Estimation of pollutant loads transported by runoff by using a GIS model Case study: Luleå city centre', *NOVATECH 2010*.

Nordeidet, B., Nordeide, T., Åstebøl, S.O., & Hvitved-Jacobsen, T. (2004), 'Prioritising and planning of urban stormwater treatment in the Alna watercourse in Oslo', *Science of the total environment* **334**, 231–238.

Park, M.H., Swamikannu, X. & Stenstrom, M.K. (2009), 'Accuracy and precision of the volume–concentration method for urban stormwater modeling', *Water research* **43**(11), 2773–2786.

Schueler, T. (1994), 'The importance of imperviousness', *Watershed protection techniques* **1(3)**, 100–111.

Valtanen, M., Sillanpää, N. & Setälä, H. (2013), 'Effects of land use intensity on stormwater runoff and its temporal occurrence in cold climates', *Hydrological Processes* **28**(4), 2639–2650.

Wing, M.G. & Godwin, D. (2011), 'SWAMP GIS: A spatial decision support system for predicting and treating stormwater runoff', *Journal of Spatial Hydrology* **11**(2), 21–32.

Valtanen, M., Sillanpää, N. & Setälä, H. (2013), 'Effects of land use intensity on stormwater runoff and its temporal occurrence in cold climates', *Hydrological Processes* **28**(4), 2639–2650.

This book contains the papers presented during the Open Source Geospatial Research and Education Symposium (OGRS 2014) held in Helsinki, Finland, on 10 to 13 June 2014. OGRS is a forum for researchers and educators to share knowledge, new solutions, methods, practices, ideas and trends in geospatial achieved with the development and use of free and open source software (FOSS). OGRS wants to demonstrate the many positive impacts of FOSS on research and education.

OGRS

9 789526 057071

**BUSINESS +
ECONOMY**

**ART +
DESIGN +
ARCHITECTURE**

**SCIENCE +
TECHNOLOGY**

CROSSOVER

**DOCTORAL
DISSERTATIONS**