

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/215666589>

A prototype digital cadastral information system for the survey of Kenya

Conference Paper · January 2010

CITATION

1

READS

1,647

3 authors:



David Ndegwa Kuria

Dedan Kimathi University of Technology

77 PUBLICATIONS 634 CITATIONS

SEE PROFILE



Nicholas Mwangi

Kenyatta University

1 PUBLICATION 1 CITATION

SEE PROFILE



Moses Murimi Ngigi

Dedan Kimathi University of Technology

30 PUBLICATIONS 107 CITATIONS

SEE PROFILE



AGSE 2010

Applied Geoinformatics for Society and Environment
Geoinformática Aplicada para la Sociedad y el Medio Ambiente
Third International Summer School and Conference

Hochschule für Technik Stuttgart
University of Applied Sciences

Volume 109 (2010)

Franz-Josef Behr, Anakathil Purushothaman Pradeepkumar, César A.
Beltrán Castañón (Editors)

Applied Geoinformatics for Society and Environment
Geoinformática Aplicada para la Sociedad y el Medio
Ambiente

AGSE 2010

Publications of the

Stuttgart University of Applied Sciences
Hochschule für Technik Stuttgart

Volume 109 (2010)

Universidad Católica de Santa María, Arequipa, Peru

ISBN 978-3-940670-19-9

Publications of the Hochschule für Technik Stuttgart Stuttgart University of Applied Sciences

Volume 109 (2010)

ISBN 978-3-940670-19-9

Conference Web Site. <http://applied-geoinformatics.org/>

Authors retain copyright over their work, while allowing the conference to place their unpublished work under a Creative Commons Attribution License, which allows others to freely access, use, and share the work, with an acknowledgement of the work's authorship and its initial presentation at this conference.

Authors have the sole responsibility concerning all material included in their respective contributions.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover Design: Christian Schmidt, Franz-Josef Behr

Typesetting: A. Shamila Jayasekare

Program Committee

Chair: Prof. Dr. Franz-Josef Behr, Stuttgart University of Applied Sciences, Germany

Members:

Prof. Dr. Dietrich Schröder, Stuttgart University of Applied Sciences, Germany

Dr. Anakkathil Purushothaman Pradeepkumar, School of Environmental Sciences, Mahatma Gandhi University, Kottayam , Kerala, India

Dr. David Kuria, Jomo Kenyatta University of Agriculture and Technology, Kenya

Prof. Dr. Michael Hahn, Stuttgart University of Applied Sciences

Dr. V. P. A. Weerasinghe, University of Kelaniya, Sri Lanka

Organizing Committee

Prof. Dr. Franz-Josef Behr, Dr. A.P. Pradeepkumar, Mirka Zimmermann M. Sc. and the staff members of the Faculty of Geomatics, Computer Science and Mathematics, Stuttgart University of Applied Sciences, Germany

César A. Beltrán Castañón, PhD. Professor at the University of San Agustín and Santa Maria Catholic University - Arequipa, Peru, and the staff members of Santa Maria Catholic University

Table of Contents

Keynotes	15
Internationalization of Higher Education in Latin America: Language, Prospects and Problems of Peru.....	3
A.P. Pradeepkumar and Franz-Josef Behr	
Monitoring of Vegetation and Change of Landuse in Chile.....	5
Victor Sandoval V.	
Indoor Real Time Location with Active RFID – System Preciscion and Possible Applications	7
Paul Rawiel	
Photogrammetry, Earth Observation Systems, Information Extraction	15
LiDAR Assisted CityGML Creation.....	17
L.M Malambo and M. Hahn	
Low-Cost Airborne Platforms and Remote Sensors for Precision Farming.....	25
Andres Flores, Gonzalo Cucho, Ernesto Gordillo, Marlon Aucallanchi, Renán Rojas, Jorge Barba, Aurelio Rodríguez, Roberto Quiroz, Adolfo Posadas, Carolina Barreda and Hildo Loayza	
Euler Deconvolution of Potential Field Magnetometer and Its Use in Delineation of Regional Faults with Data from Northern Peru.....	31
C.J. Solano Salinas and E. Borda	
Eliciting Conflict in Expert’s Decision for Land Use Classification.....	37
V.K Panchal, Sonakshi Gupta, Nitish Gupta and Mandira Monga	
Designing Airborne LiDAR Data Interoperabiity Standards for Multiple Organizations – Canada's first experience in New Brunswick.....	43
P. W. Adda, D. Coleman; D. B.Finley	
Environmental Issues, Sustainable Development, Millennium Development Goals	49
GIS-Based Landuse/ Land Cover Change Characterization in the Humid Tropical Meenachil River Basin, Kerala, South India.....	51
Vincy Mary Varghese, Brilliant Rajan, A.P.Pradeepkumarand Riju Stephen	
Forest Fire Risk Analysis and Management System Using Geoinformation Technology	57
Subin K. Jose, Santhosh Kumar, M.M. Jesnamol, N.K. Sreeraj, G.Madhu and Babu Ambat	
Forest Disturbance Analysis Using Geoinformatics in Pamba Range of Periyar Tiger Reserve, Kerala, India	63
Abin Varghese , John C. M., Punnen Kurian and Thomas A P	
Options for Biodiversity Monitoring within REDD+	68
Steffen Entenmann	
The ArtWET tool: a Georeferenced Approach Assessing Runoff Related Pesticide Pollution in Surface Waters across Europe	69
David Elsaesser and Ralf Schulz	
Geographic Cataloging of Ethnobotanical Information – A Review of Ethno medicinal uses of <i>Saraca asoca</i> (Roxb.) De wilde in India	75
R. Jaishanker, Arathyram R.S, V. Saroj Kumar, and C.S.P Iyer	

Morphometric Analysis and Prioritization of Microwatersheds in Valiyathodu Sub-Watershed of Manimala River, Kerala, India.....	77
V.B. Rekha, A.V. George, K.R.Vineeth and U. Johnson	
Geoinformation for Tourism Development.....	83
Mirka R. de Zimmermann and Maria Fernanda Serra	
WHO's Public Health Mapping and GIS Programme.....	87
Imran Hossain, Tania Puspita Firdausy and Franz-Josef Behr	
Monitoring of fruit fly (Diptera Tephritidae) and their relation with biotic and abiotic factor in Santa Elena region, Ecuador.....	93
Alvaro Cañadas, Viviana Duque, Oscar Calahorrano and Christian Garófalo	
Fixing Spatial Development Strategies for a City with GIS.....	95
Zahid Hasan Siddiquee	
Internet based Applications, Open Source Solutions	101
OpenLayers: Open Geodata in an Open Source Map Browser.....	103
Franz-Josef Behr	
Open Source Desktop GIS - Ready to Use for Teaching at Universities?	1034
Dietrich Schröder	
Geocoding: Fundamentals, Techniques, Commercial and Open Services	111
Franz-Josef Behr	
OpenStreetMap: Data Model, Crowdsourcing, and Technology	123
Franz-Josef Behr	
Data Visualization Using Google Maps: the Hard Way and the Easy Way.....	125
Shailesh Shrestha and Franz-Josef Behr	
Technology Innovations	133
Mobile Mapping – A Carfinder Application for a Modern Carsharing System.....	135
Paul Rawiel	
Car Navigation Systems: Technical Principles and Trends.....	136
Cesar Amador Ocampo Lecaros	
Spatial Data Infrastructures, E-Cadastre	143
Spatial Data Infrastructures – Status 2010	145
Hardy Lehmkuhler	
Redefinition of Coordinate Reference for GIS: Expected and Unexpected Effects.....	146
Hardy Lehmkuhler	
Disponibilidad de información geoespacial para la investigación mediante la Implementación de una IDE	147
Eddy J. De la Cruz Aquino	
Modelo de integración y preprocesamiento de información para el descubrimiento de conocimiento en bases de datos federadas utilizando ontologías	155
Geovanna Evelyn Espinoza Taype	
GIS for Municipalities: A Review of the State-of-the-Art in Germany	165
Dietrich Schröder	

A Prototype Digital Cadastral Information System for the Survey of Kenya	167
David Kuria, Nancy Mwangi and Moses Murimi	
Workshops.....	175
Geostatistics – an Introduction	177
Hardy Lehmkuhler	
Workshop on Open Source Desktop GIS.....	179
Dietrich Schröder	
Workshop on “R” for Geospatial Analysis	192
Dietrich Schröder	
Internet Mapping – An Introduction	193
Franz-Josef Behr	
Poster	195
Using GIS and Remote Sensing for the Assessment of the Potential Impacts of Flooding on Roads Network in Makkah "Mecca", Saudi Arabia.....	197
Meraj N. Mirza and Ramze A. Elzahrany	
Plataforma de Procesamiento de Imágenes Satelitales para la Detección de Contaminación Minera.....	198
Juan Carlos Gutierrez Cáceres ¹ and César Beltrán Castañón	
Sistema de Seguimiento de Cuidado de la Salud, basado en tecnología Geoinformática: Caso madres gestantes	
César Beltrán Castañón and Agueda Muñoz del Carpio Toia	
Developing a Community Forum using Ushahidi.....	200
Emmanuel Bobb Anya Lansana, M. Abdullah Sohl, Franz-Josef Behr	
A Framework Experiences of a Distance Education Program for GIS in Latin America. UNIGIS 2000 – 2010.....	201
Carla M. Valdospinos Navas	

Preface

The Applied Geoinformatics for Society and Environment (AGSE) series of seminars was started in 2008 with the objective of taking the benefits of geographic information technologies to a wide canvas of applicable areas, particularly in the realms of development, environment and social well-being. AGSE has always sought to align itself with the laudable aims of the United Nations Millennium Development Goals and through its workshop and conference activities tried to inculcate the spirit of the MDGs in the work of the participants. The stress has always been on empowerment, be it gender-based or nation-based; and this is evidenced by the high proportions of women as well as an equally high proportion of scientists from the developing world participating in the two previous AGSEs (AGSE2008 in Trivandrum, India and AGSE2009 at Stuttgart, Germany). Also alternate viewpoints have their own niche in the AGSEs. The motto thus is ‘to offer an interdisciplinary, international forum for sharing knowledge about the application of Geoinformatics with focus on application and on developing countries’.

It is apt that the AGSE returns to a developing nation with high biodiversity this year, the UN’s International Year of Biodiversity. Peru with its outstanding annual growth rate of 5.3%, and very high earnings from tourist arrivals is a cradle of ancient civilization, as well as a natural biodiversity laboratory. Peru hosts one of the world’s most varied and exotic ecologically significant landscapes across terrain ranging from the sea level to the Andean high peaks. Natural phenomena and Peru’s fame and fortune are closely interlinked, the Humboldt current bringing in rich fish resources while the El Niño brings with it misery. Peruvian sense of nationality goes back to pre-historic times with the nation taking pride in Pachacútec, the emperor who expanded and consolidated the Inca Empire. And it is apt that the AGSE2010 is being hosted in Arequipa, La Ciudad Blanca, by the Universidad Católica de Santa María which is at the forefront of technology as well as the host of *Juanita* at its Museo Santuarios Andinos, thus melding two diverse streams of human endeavour. The AGSE2010 and JINIS2010 (the XVII Jornada Internacional de Ingeniería de Sistemas), which take place in parallel, bring to the forefront the technological acumen that Peru has developed over the years.

The papers in this AGSE2010 cover scientific and application oriented tracks classed into 12 broad themes of current relevance to geoinformatics like Environmental Issues, Sustainable Development, Millennium Development Goals; Open Source Solutions and Tools; Internet-based Applications; Neogeography and Geovisualisation; Photogrammetry; SDI; LiDAR data analysis; GIS and health; GI Policy and Society, e-Government; and alumni experiences. The conference certainly will have a Latin American flavor spiced with international inputs from countries like India, Bangladesh, Ecuador, Saudi Arabia, Germany, Kenya, Canada, Sri Lanka, USA etc. GIS-based papers addressing biodiversity and sustainable development in developing countries, fire hazard risk management, disaster management, share space with papers elucidating state-of-the-art GIS technologies to tackle river pollution, papers extolling Free and Open Source GIS, city GIS etc. Material from the workshops on Internet mapping, Open source Desktop GIS, Geostatistics and R, Process management, quick WMS server setup add to the pedagogic nature of the proceedings volume.

The Third Applied Geoinformatics for Society and Environment summer school-cum-seminar takes forward an idea that germinated in late 2007, for a renewed internationalization effort in the Hochschule für Technik Stuttgart (Stuttgart University of Applied Sciences – SUAS), with a commitment to ‘collaborative science for the betterment of society and environment’. A proposal for establishing a continuing series of summer schools-cum-seminars, rooted in the concept of knowledge sharing across boundaries and barriers, and with the active collaboration of the alumni of this university under the aegis of Stuttgart Active Alumni Group (SAAG) was suggested by Prof. Dr. Franz-Josef Behr of the Department of Geomatics, Computer Science and Mathematics. The success of the first and second AGSEs funded by DAAD and other German agencies in collaboration with partners worldwide, lead naturally to the organization of this third summer school of the series.

This conference would not have been possible without the collaboration and support of different people and organizations.

The editors would like to express their gratitude to Universidad Catolica de Santa Maria, Arequipa, for the preparation and hosting of the conference! “Gracias” for providing such a nice venue!

The editors would also like to express their gratitude to the Faculty of Geomatics, Computer Science and Mathematics, Stuttgart University of Applied Sciences and to the reviewers for their efforts to improve the quality of the symposium and the proceedings in the narrow time frame. Special thanks to Prof. Dr. H. Lehmkuhler, Prof. Dr. P. Rawiel, Prof. Dr. D. Schröder for all support and active participation!

The support of all sponsors (see website), especially the German Academic Exchange Service (DAAD) with additional financial support by the Federal Foreign Office, INPHO GmbH, GeoEspacial Tech, German Society for Geodesy, Geoinformation and Land Management., German Cartographic Society e.V. (DGfK), Knödler-Decker-Stiftung, and the Faculty of Geomatics, Computer Science and Mathematics of Stuttgart University of Applied Sciences is gratefully acknowledged. Special thanks to Mirka Rodriguez de Zimmermann for her proactive and dedicated work as well to Miguel Guerrero and to Cesar Amador Ocampo Lecaros for their valuable support. Again, A. Shamila Jayasekare has done a great work while formatting and typesetting the submissions.

Prof. Dr. F.-J. Behr

Dr. A. P. Pradeepkumar

Prof. Dr. César A. Beltrán
Castañón

Stuttgart University of Applied
Sciences, Germany

Mahatma Gandhi University,
Kottayam, India

Universidad Catolica de Santa
Maria, Arequipa, Peru

Alumni Representative

Representative STUTT GART
Active Alumni Group (SAAG)

Palabras de bienvenida

Estimados amigos,

Es muy grato para nosotros a nombre del Programa Profesional de Ingeniería de Sistemas de la Universidad Católica de Santa María, Arequipa – Perú, poder recibir a diferentes expertos, alumnado y gente interesada alrededor del mundo para intercambiar conocimiento e ideas en la “Third International Workshop and Summer School on Applied Geoinformatics for Society and Environment 2010”, estamos seguros que los objetivos planteados serán alcanzados, como anfitriones esperamos que su estadía en nuestra ciudad sea de lo más complaciente.

Sean todos ustedes bienvenidos!, Arequipa lo espera!

José Sullá Torres, Director

Programa Profesional de Ingeniería de Sistemas
Universidad Católica de Santa María

Keynotes

Internationalization of Higher Education in Latin America: Language, Prospects and Problems of Peru

A.P. Pradeepkumar^{1*} and Franz-Josef Behr²

¹ Disaster Management Division, School of Environmental Sciences, Mahatma Gandhi University, Kottayam, Kerala, India 686 560

² Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart, Schellingstraße 24, D-70174 Stuttgart, Germany
franz-josef.behr@hft-stuttgart.de

KEYWORDS:

Internationalization, Peru, Latin America, language, globalization

ABSTRACT:

Internationalization is a word that has become clichéd by its application in and out of context and to almost all situations where more than one nation is involved. But despite these criticisms, the fact remains that an international experience often defines a turning point in a person's life. The whole perspective of the person involved undergoes a change, which becomes the most important outcome of the internationalization process, rather than the much touted economic or sociocultural benefits. Internationalization imparts a holistic perspective in a person and places him on a high vantage point of critical observation, which many of his peers may not have the fortune to experience. Internationalization has a national realm as well as an individual realm and often the nation's policies decide the level to which academic internationalization can progress, whereas at the individual realm often national policies have little influence (deWit et al., 2005).

Thus internationalization has a rustic, earthy realism about it which is often masqueraded by the glitzy glamour of international organizations and their arms of internationalization. Pradeepkumar and Behr (2009) opined that internationalization is not a new phenomenon, and the visits of Chinese scholars to the Indian universities of the BC period were truly events that predated globalization and internationalization as we know it today. Even in those days, internationalization and globalization went hand-in-hand. Knowledge transfer, even in the days of instantaneous communication, is effective only when one can personally experience excellence, and no political or geographic barrier can prevent this; just as was in the case of Hsüan-tsang (who went on to establish the Weishei school of Buddhism in China) the Chinese scholar in residence at the Taxila University and Nalanda University in India. Cross-border scholarly migration has taken place in ancient times at most centres of higher learning in the world (Stearns and Langer, 2001). This activity is currently termed 'networking' and Latin American nations leverage this amongst themselves and internationally to attain excellence in certain research and academic centres.

It's been 460 years since the founding of Peru's oldest and most prestigious university, UNAM or Universidad Nacional Mayor de San Marcos, Universidad del Perú (UNAM), and has Peru made good use of this early advantage? It is in this context that we take a look at internationalization of education in Latin America with special emphasis on Peru. The level of internationalization varies from country to country in Latin America (de Wit et al., 2005). In Peru, as in most developing countries internationalization had meant receiving development aid and technology transfer in all sectors, including education. But this approach is slowly giving way to cooperative agreements and partnerships, not just with the developed nations, but also with the developing world (Butters et al., 2005). Internationalization of education in Peru has not really lead to quality improvement, because most of it has involved foreign entities setting up universities in Peru and running courses that need low investment, like sociology, law and business administration. This has meant that no qualitative improvement in the S&T capabilities of students have taken place due to this internationalization, which actually is glorified globalization in action.

Peru's indigenous population is a treasure trove of little studied languages. International collaborations in the study of such languages are essential and this is an area of internationalization which can lead to better elucidation of the cultural and sociological uniqueness of the natives and thereby enrich the world's understanding of its cultural diversity. The move by Consejo Nacional para la Autorización de Funcionamiento

* Corresponding author.

de Universidades (CONUFA)'s to start a new university for the indigenous population of Northern Amazon region which will deal with topics like forestry management and medicinal plants which are relevant to the local people, in their own languages like Aguarana, Ashaninka (pre-Andean Arawak linguistic family), and Shipibo-Conibo is a move in the right direction and it is in such ventures that international cooperation can be fruitful.

The maximum fee paid by Peruvian students for higher education is \$3,600 and 33% of students work while studying in the public universities compared to 28% in the private ones. This can be correlated with the fact that students who study in the government universities come from a less affluent background compared, but it also cannot be ignored that probably it is the former that maintains higher quality control over intake and provides better quality education (only 16% of applicants are successful in government universities while 74% are successful in private ones (INEI and ANR, 1997)). Thus internationalization of education in Peru has to take into account the realities of the educational aspirations of the people, rather than allow the profit motive to decide the course of foreign involvement. Introspection on the factors that can lead to student motivation and improved enrolment often reveals that excellence in academic matters is the prime motivator, as the Pontifical Catholic University of Peru in Lima found out in the latter half of the 90s, when it was faced with a bleak future and examined the reasons for low student enrolment (Salmi, J., 2009). A critical look at the GATS, Mercosur agreements are required so that internationalization does not get subsumed by the doctrines of globalization in Peru.

References

- Pradeepkumar, A.P. Behr, F-J. 2009 Globalization and internationalization: European and German perspectives on higher education under GATS in Behr F-J, Schroeder D, Pradeepkumar A P *Proc. Intl Applied Geoinformatics for Society and Environment Conf (AGSE2009), Stuttgart, 13 - 18 July, 2009.*
- Butters, L.J.C., Quiroga, L.B. and Dammert, P. L. 2005 Internationalization of higher education in Peru in: deWit, H., Jaramillo, I.C., Gacel-Ávila, J. and Knight, J. (editors) *Higher Education in Latin America: The International Dimension* World bank, Washington D.C., pp311–328
- deWit, H., Jaramillo, I. C., Gacel-Ávila, J. and Knight, J. (editors) 2005 *Higher Education in Latin America: The International Dimension* World bank, Washington D.C., pp311–328
- INEI (Instituto Nacional de Estadística e Informática) and ANR (Asamblea Nacional de Rectores). 1997. "I censo nacional universitario 1996. Síntesis estadística." Lima.
- Salmi, J. 2009. *The Challenge of Establishing World-Class Universities*. The World Bank, Washington D.C., 136p.
- Stearns, P. N. and Langer, W. L. 2001. *The Encyclopedia of World History*. 6th edition. Houghton Mifflin, London. 1057p

Monitoring of Vegetation and Change of Landuse in Chile

Victor Sandoval V.

Dr. y M.of, Sc. en Inventarios Forestales y Cartografía
Universidad de Goettingen
Actualmente profesor de las cátedras de Geomatica, Muestreo y Monitoreo
Universidad Austral de Chile

A fines de la década pasada, el esfuerzo combinado entre diversas organizaciones gubernamentales y de investigación chilenas, dio como resultado el primer Catastro y Evaluación de los Recursos Vegetacionales Nativos de Chile, que generó información detallada acerca de las superficies pertenecientes a las distintas clases de uso del suelo, además de una tipología de la estructura y composición de los recursos vegetacionales del país, con énfasis en las áreas boscosas. Este proyecto y sus respectivos monitoreos, corresponden a los estudios más completos y recientes que existe acerca de la descripción del paisaje chileno. La metodología empleada, incluyó la realización de una cobertura aerofotográfica del territorio nacional en escalas 1:250.000, 1:100.000 y 1:50.000 En ese contexto y considerando la experiencia obtenida en el levantamiento periódico de información regional en los últimos 15 años, se exponen algunos elementos relacionados a la cartografía y monitoreo del uso del suelo en Chile y se plantean opciones futuras para estimar el cambio de uso del suelo nacional. El objetivo es mejorar los datos que este proceso provee, para finalmente entregar a la sociedad información que facilite la toma decisiones que afectan finalmente a nuestros recursos naturales y habitantes.

At the end of the last decade, the combined efforts of various governmental and research organizations in Chile resulted in the First Cadastre and Evaluation of Native Vegetation Resources of Chile, which generated detailed information about the areas belonging to different classes of land use, and a typology of the structure and composition of the vegetation resources of the country, emphasizing forested areas. This project and its corresponding monitoring refer to the most recent studies in existence about the description of Chilean landscape. The methodology used included cartography and aerial photography of the country at scales of 1:250,000, 1:100,000 and 1:50,000. In this context and considering the experience gained from the periodic survey of regional information in the last 15 years, some elements related to mapping and monitoring land use in Chile are set, and future options to estimate the change of use of national soil are stated. The aim is to improve the data provided by this process, to finally present society information that facilitates making decisions that will ultimately affect our natural resources and inhabitants.

Indoor Real Time Location with Active RFID – System Precision and Possible Applications

Paul Rawiel

University of Applied Sciences Stuttgart
Schellingstr. 24, D - 70174 Stuttgart
Phone: 0049 (0)711 8926-2529, Fax: 0049 (0)711 8926 2666
e-mail:paul.rawiel@hft-stuttgart.de

KEYWORDS

Indoor Positioning, RFID, Real Time Location System

ABSTRACT:

Today with GPS a broad public has access to a widely available and easy to use location and navigation system. E.g. many cars have a navigation system on board that makes it easier and faster to find a special location. But with GPS, navigation stops where GPS is not available as inside buildings or in big cities where the signal is blocked by high-rise buildings. Today other manmade points of reference than satellites such as broadcast stations for mobile communication exist. In many cases a continuing navigation system for pedestrians that has to fulfill other requirements would be desirable. Especially inside buildings, where GPS is not available there is a gap for precise reliable possibilities to locate and track persons or objects or to guide a user through the building according to his needs. An active RFID system is an option to close this gap. This work presents such a location system based on radio frequency identification (RFID) to locate objects inside a building. The results of studies about the system precision [7] and possible applications are presented in this paper.

1 Introduction

When talking about positioning and navigation, most people have satellite based systems in mind such as NAVSTAR GPS, GLONASS or the coming European system GALILEO. Other technologies are not well known to the general public. The public is getting more and more used to comfortable navigation systems, and the demand for such systems not available only in the car or under the free sky, but also inside buildings to get guided through a museum, an exposition or something alike will become stronger [2]. But not only navigation could be an interesting application. Another application could be to locate important and expensive equipment for example in a hospital. Or not only equipment – in a hospital it can also be of vital interest to locate persons fast and with high reliability. In big production facilities it sometime costs a lot of time and money to find a special tool to execute a special production step. Car manufacturers send their cars through several quality checks after they are completely assembled. It may be of interest where a car is located during these quality checks and for how long it remains at a certain station or where it can be found on the parking area. Another possible application could be the automatic control of guideless vehicles. For this, it is necessary to track the vehicle precisely and to control its movements from a remote station automatically. In this work a system that allows us to track a number of objects in real time and some studies about its precision are presented [7]. Additionally some possible applications out of different areas like logistics, manufacturing or health care are shown.

1.1 Indoor Positioning

Different methods using radio frequency methods have been used to do indoor location for example passive UHF RFID, Bluetooth, or WLAN (see [7]). With UHF RFID with its passive RFID Tags and with the Bluetooth approach Vorst et. al. [7] reached an accuracy of around half a meter. The WLAN approach comes to a lower accuracy of around 4 meters. UHF RFID as well as Bluetooth is limited in the range that can be reached while WLAN has here an advantage. In this paper a system that uses active RFID tags that send a UWB signal (ultra wide band). UWB works with a very high base frequency of approx. 6-8.5 GH. The positioning with UWB is based on signal run time differences and/or angles of arrival at different receivers. Due to the enormous economic attraction of RFID, positioning with active transponders (UWB) is brought to market under the label RFID. From a technical point of view, positioning with UWB is a wireless application of telemetry [3]. For a closer look on telemetry senders see [1].

1.2 Time Difference of Arrival (TDOA)

TDOA is a procedure for positioning using time difference of arrival of a signal with at least three receivers in 2D and four receivers in 3D. A geometrical interpretation for 3D is as follows: If the time difference of arrival at two receivers is known, the position of the sender is any position on a hyperboloid with the focal point in a receiver and the axis on the connecting line between the receivers. With an additional receiver, the position of the sender is on the intersection curve of two hyperboloids. An additional receiver and the intersection of the previous curve with the new hyperboloid result in a point as position of the sender. The synchronization of the time-pieces of the receivers plays a major role for the accuracy of the procedure. In the system used in this work, the receivers are connected to each other with a special timing cable.

1.3 Angle of Arrival (AOA)

In order to make angle measurements possible, each receiver must be equipped with a two-dimensional antenna array. That means a lining up from four to twelve single antennas in horizontal and vertical direction, which may not lie further away from each other as a wavelength [4].

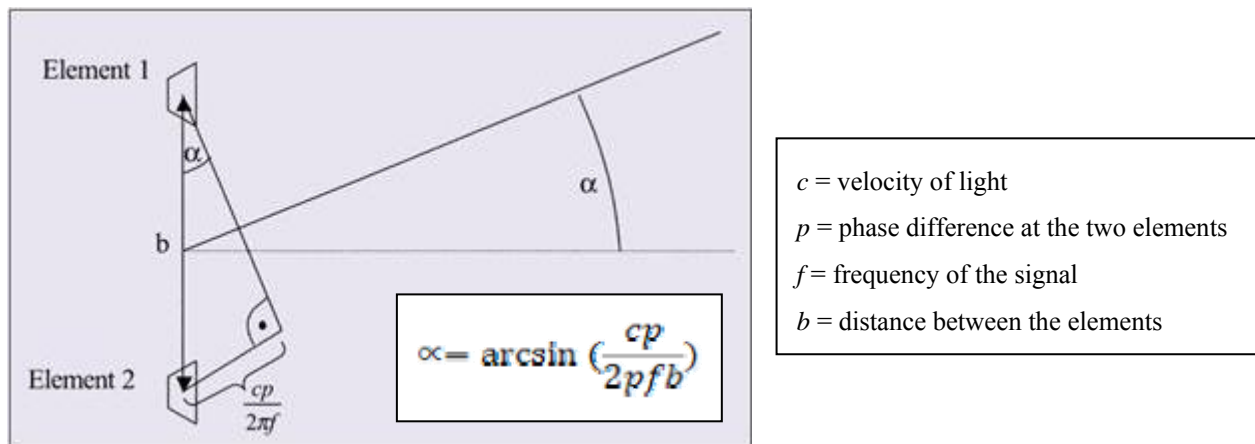


Figure 1: Angle measurement with two elements [4]

An advantage of AOA is that only two receivers are needed to determine a position. The disadvantage is the accuracy that differs very much with the spatial arrangement of receivers and the sender.

1.4 The Real Time Location System

The system that was used is a complete solution of the company Ubisense with four sensors that were installed in a laboratory of the University of Applied Sciences Stuttgart. The sensors are connected with two types of cables. One is the previously mentioned timing cable the other is a network cable through that the sensors communicate via router with a server which hosts the Ubisense Platform Control. A local coordinate system was established in which the further location and navigation activities were performed. Therefore the sensors were measured with a total station with reflectorless distance measurement. Additionally, the sensors were calibrated through the determination of tags on known positions. The tags have to be initialized as well and make them known to the system. With the Site Manager a big variety of settings can be done to configure the system for a special application in a special environment. For example, zones can be defined and events can be triggered. There can be fixed zones and moving zones attached to a tag. Thus, if a tag enters or leaves a zone, a special event can be handled.

1.5 Precision Considerations

Different filters can be used to make the position of a tag more stable. For the figure below, a tag was moved on a previously determined track that served as reference track, to which the calculated positions of the system can be compared. The blue points are the measured values while the red line is the given track.

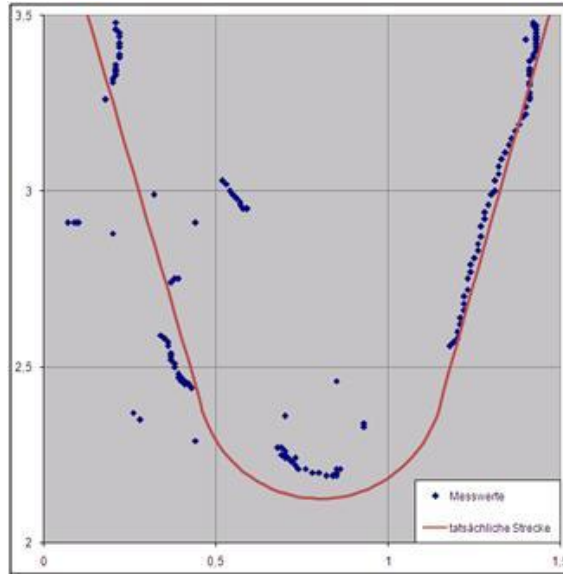


Figure 2: Accuracy test of the system [2]

On the right side the measured positions match very well with the reference positions, while on the left side significantly higher deviations can be observed. This probably is due to the fact that the left part of the track led close along the wall. So here multipath effects play a bigger role for the accuracy of the position. The filter applied here is fixed height information filtering: The tag is free to move horizontally, but vertical motion is constrained to be close to a fixed height above the cell floor. The motion model is of position and velocity with gaussian noise on velocity. If a tag is not seen, it is assumed to continue moving with the same horizontal speed and direction, with uncertainty increasing over time, but to keep "close" to the same fixed height [5]. In other experiments a reference track was measured precisely with a theodolite and different types of tags have been moved in different conditions with different velocities along the track. Different filters have been applied and the results have been compared.

The following figure shows the result of a first experiment on the 6 m reference track. Here, one can only see the residuals in the orthogonal direction to the track. In direction of the track, errors cannot be detected. On the 6m track, 70 position measurements were done on each run. The average distance to the reference track hereby 0.046m in the horizontal position and 0.088 m in the vertical position. The standard deviation of a horizontal position was of 0.041 m, in the vertical position of 0.137 m [7].

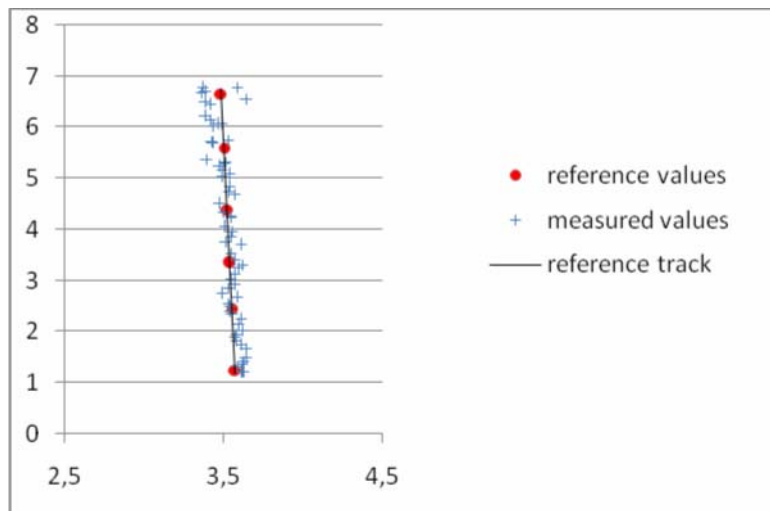


Figure 3: Accuracy measurements of the system under good conditions [7]

Another experiment was no intervisibility between the tag and the sensors, where more multipath influences could be expected led to the following results:

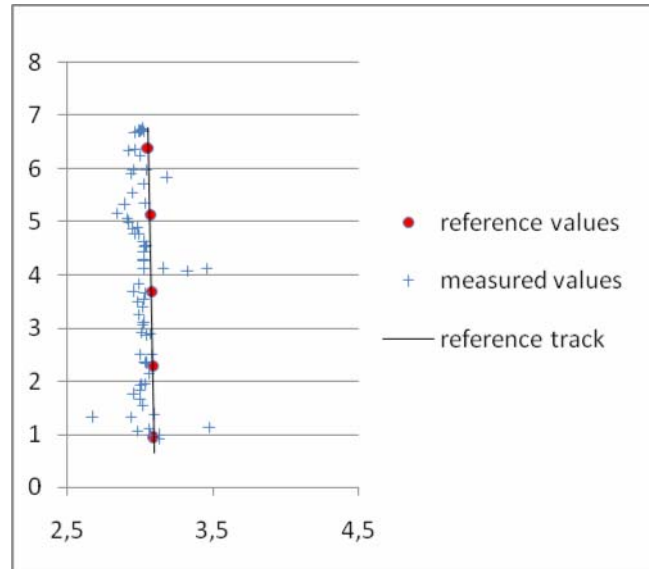


Figure 4: Accuracy measurements of the system with multipath effects simulated [7]

The average distance to the reference track in this case was 0.096m in the horizontal position and 0.091 m in the vertical position. The standard deviation of a horizontal position was of 0.090 m, in the vertical position of 0.287 m [7].

The experiment that was carried out to determine if the velocity of the tag has any influence on the precision brought the result, that there is no significant increase inaccuracy with increasing velocity [7]. A bigger problem is the multipath effects that can be dramatic in especially in industrial environments where rooms maybe filled with metal racks or tube systems. Under normal conditions, an accuracy of about 15 to 30 cm should be reachable, which is enough for a big variety of applications.

2 Possible Applications

2.1 Logistics Materials Flow – Forklifts



Figure 5: Forklifts[7]

In large storing facilities, goods that are stored on pallets or in other carriers have to be located. That can be done with a special order after which special kinds of goods are always stored in well defined locations, so they can be found when needed. A more flexible way of storing can be done, if the movements of the forklifts are continuously tracked. A combination with conventional RFID, where a reader is placed at the forklift, that registers every pallet or carrier that is picked up and dropped off, the drop-off-location can be stored in the system, so the locations of the goods are always known and less manual location and identification is required. As a side effect, misplaced forklifts can easily be found [7].

2.2 Logistics Quality Management – Barcode Scanners

It always may happen that goods are damaged on transport or even get lost. Therefore the goods usually are tracked using the barcode readings at each handover along the transport. In large storing or cross-dock facilities, the gates where goods leave and come in are monitored by video cameras. So in case of reclamation because of loss or damage, the transport company has a possibility to deliver a proof of correct handover in time with the video sequence that registered the handover of the items in question. The problem is that terabytes of video sequences are produced, without any location information about the goods. Much time has to be spent to find the video sequence with the wanted documentation of the transport. The active RTLS can help here by tracking the barcode scanners that register the incoming and outgoing items. A combination of the scanned item ID with time and the position of the barcode scanner can help to identify the video sequence that monitored the shipment in much shorter time [7].

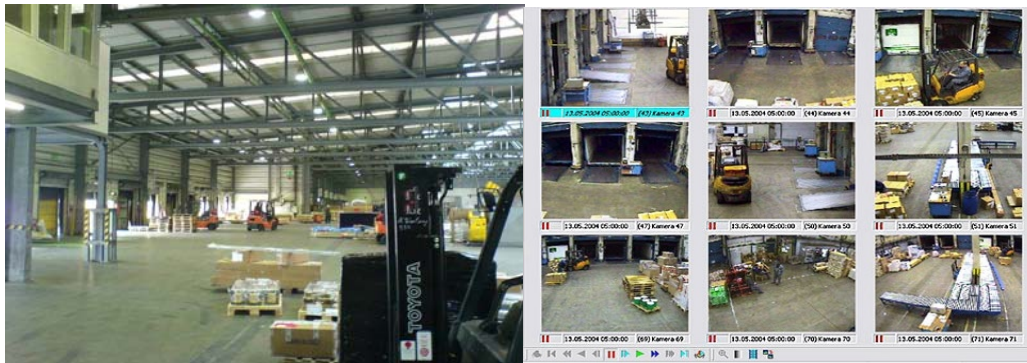


Figure 6: Video monitoring of storing facilities [7]

2.3 Automatic control of guideless vehicles

In a study project of the University of Applied Sciences a vehicle in form of a model railway engine was controlled using the RFID-based location system and the .NET API that is provided by the system. An open source server was used to communicate with the railway to control the actions of the model engine and the turnouts of the rails [2].

The digital model railway equipment is to be controlled automatically by a computer on basis of location data of the tags. That means the railway engine has to be connected to a tag as well as different wagons according to their load have to be connected to tags. The system determines the destination to where to navigate according to the load that is transported. That means e.g., the locomotive recognizes automatically that currently coal is loaded and therefore routes to the coal-fired power plant. For the control of the railway equipment an open source server (DDW-Server) that generates signals and control commands was used. Due to the .NET-API of the Ubisense system that is written in C#, the Client was written in C# as well.

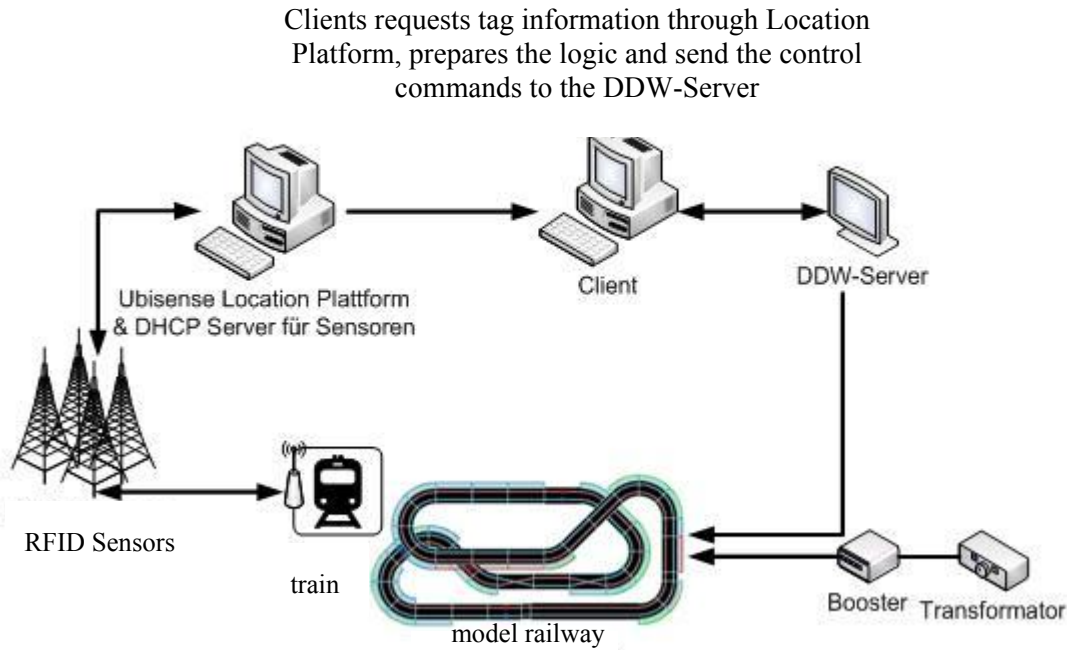


Figure 7: System configuration [2]

In order to send the desired control commands to the DDW server, the simple Railroad Command Protocol (SRCP), which is based on TCP/IP, is used [2].

3 Conclusions

The implementation of the indoor real time location system at the Stuttgart University of Applied Sciences worked without major problems. A first application was implemented and worked fine. The accuracy of around 15 cm that can be reached with the system brings some limitations for applications with it. In the example of the automatic control of a model railway the tracks can only be separated from each other if they run in a certain distance from each other. Freely controllable vehicles need some additional equipment to recognize obstacles on the way. For other applications in logistics, health care or production, the accuracy of 15 cm or even less is enough to improve processes in terms of quality and time.

Personal navigation can be well realized with the system. Pedestrian navigation with its special requirements can be brought from under the free sky where GPS navigation is established to inside buildings and thus make it possible to realize continuous navigation systems inside buildings. The accuracy hereby is limited to around 15 cm. For navigation or to locate persons or objects, this should be sufficient. For exact control of vehicles other components have to be added to support the system.

References

- [1] Bensky, A.: Short-range wireless communication, fundamentals of RF system design and application, LLH Technology Publishing, USA 2000
- [2] Haug, P.: *Echtzeitortung mit Hilfe eines aktiven RFID-Systems am Beispiel der automatischen Steuerung einer Modelleisenbahn*, Bachelor thesis at the Stuttgart University of Applied Sciences, unpublished, Stuttgart, 2009.
- [3] Finkenzeller K.: *Rfid-Handbuch – Grundlagen und praktische Anwendungen von Transpondern, kontaktlosen Chipkarten und NFC*, 5th Edition, Carl Hanser Verlag Munich, 2008
- [4] Ingensand, H., Bitzi, P.: Leica Geosystems, Technologien der GSM-Positionierungsverfahren, AVN 8-9, Wichmann, 2001
- [5] Ubisense: *Ubisense LocationEngineConfig Manual*, unpublished, Ubisense Limited, 2008

- [6] Ubisense (Meike Porrmann): *Ubisense RTLS – an Overview*, Presentation unpublished, Ubisense Limited, 2009
- [7] Viehmann, B.: *Genauigkeitsuntersuchung zum aktiven RFID – Real Time Location System von Ubisense*, Bachelor thesis at the Stuttgart University of Applied Sciences, unpublished, Stuttgart, 2010
- [8] Vorst, P. et. al.: *Indoor Positioning via Three Different RF Technologies*, Proceedings of the 4th European Workshop on RFID Systems and Technologies (RFID SysTech 2008), Freiburg, Germany, 2008

**Photogrammetry, Earth Observation Systems,
Information Extraction**

LiDAR Assisted CityGML Creation

L.M Malambo^a and M. Hahn^b

^{a,b} Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany

^a lonesome.malambo@hft-stuttgart.de, ^b michael.hahn@hft-stuttgart.de

KEY WORDS:

CityGML, 3D modeling, GIS, LandXplorer, LiDAR, Urban Planning

ABSTRACT:

This report describes ways of creating multi level building models using the CityGML model. A fusion of various data such as LiDAR (Light Detection and Ranging) data and aerial imagery is used to derive the necessary 3D information for CityGML creation. Laser scanning data has been used to provide height information for both digital terrain modeling and building models. High resolution imagery in form of Google images (orthophotos) have been used as auxiliary data for the digitization of object footprints. Acceptable results for the various CityGML levels of detail for the Stuttgart University of Applied Sciences (HfT Stuttgart) campus have been produced.

1 Introduction

1.1 Background

For centuries maps have been used to provide man with information about places, relative distances between places and events in a particular locality. Over the years maps have moved from the realm of cartographers to virtually all sectors of development. Technologies such GIS, remote sensing, and in particular the Internet, have popularized the use of maps and heightened the need for spatial information.

While plane maps or 2D information dominates the spatial information industry the use 3D information is emerging as a new trend in many disciplines. With "What You See Is What You Get" approach which is driving our information needs today, 3D models have increased in popularity and are now being used in many sectors such as urban planning, tourism and entertainment.

3D city models have found use in urban planning. While they are generally used for visualization of the built environment, there are used for other sophisticated applications including disaster management, mapping of buildings and their heights, simulation of new buildings, change detection, virtual reality etc¹.

There are several data models for the creation of 3D city models including X3D, GML, KML, CityGML, VRML, STL etc. This report explores and describes ways of creating CityGML models using data such as LiDAR (Light Detection and Ranging) data and aerial imagery. LiDAR, which is an optical remote sensing technology, provides 3D point cloud data for objects which can be used to produce as-built models of the object. Imagery such as those available in Google maps and orthophotos are important supporting data in this process. The Stuttgart University of Applied Sciences campus has been used as a demonstration site for this modeling.

1.2 About CityGML

CityGML has emerged as a tool or method for modeling 3D city structures and objects. CityGML defines classes and relations for the most relevant topographic objects in cities and regional models with respect to their geometric, topological, semantic and appearance properties (Kolbe & Bacharach, 2006). Topographic objects here include not just building structures but also elevation, vegetation, water bodies, roads and more.

CityGML was designed as an open data model and XML-based format and is implemented as an application schema of GML3, the extensible international standard for spatial data exchange developed within the Open

¹ 3D City & Urban Modeling: <http://www.satimagingcorp.com/svc/3d-city-and-urban-modeling.html>

Geospatial Consortium (OGC) and ISO TC211 (Wikipedia, 2010). CityGML models comprise both complex and georeferenced 3D vector data along with the semantics associated with the data. In contrast to other 3D vector formats, CityGML is based on a rich, general purpose information model in addition to geometry and appearance information (Open Geospatial Consortium Inc, 2008).

CityGML represents objects in varying levels of detail (LOD). The current implementation supports five consecutive LOD starting LOD 0 to LOD 4, with objects becoming more detailed with increasing LOD as depicted in figure 1.

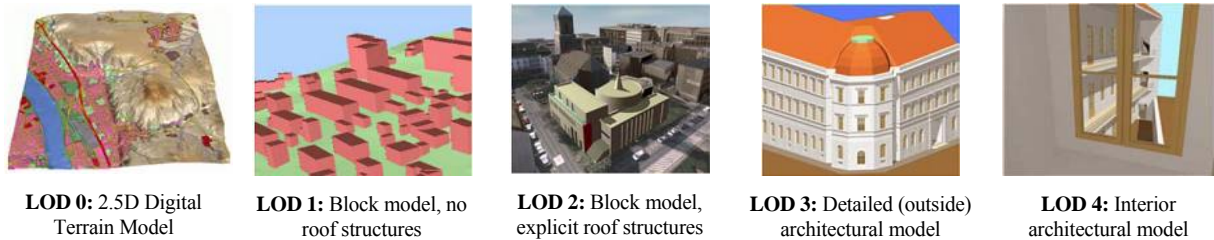


Figure 1: The five (5) LODs defined by CityGML (courtesy of Stadler & Kolbe, 2008)

2 CityGML Creation Procedure

Several approaches are being used in the creation of CityGML ranging from manual to automatic ones. Manual methods mainly involve the extrusion of building footprints from 2D vector files or the systematic creation of 3D building geometries using CAD (e.g. AutoCAD 3D) or GIS (e.g. ArcGIS, ArcScene) tools. Semi-automatic and automatic methods employ photogrammetric methods such as building extraction from orthophoto or satellite imagery (Theng, 2006) or close range methods (Falkowski, 2009).

At a basic level a 3D city model consists of a digital elevation model (ground height) and 3D building models (building heights) (VEPS, 2010). Since both are data in 3D space, LiDAR data has proved to be a good source for collecting such information. The strength of LiDAR data is that it provides accurate 3D data for surfaces and objects.

Even though LiDAR data can provide good 3D information, a fusion with other types of data such as high resolution imagery and vector data makes CityGML creation easier. In this light a procedure combining airborne and mobile LiDAR data, imagery and vectors files has been adopted.

Autodesk's LandXplorer Studio has been selected as the final and main software for the CityGML creation. LandXplorer creates models by extruding building footprints or by aggregation of several geometries or by construction of generic buildings from finished models. Third party formats such as shapefiles can be used in this software.

The following procedure has been used for the creation of the various CityGML LODs:

- 1) Classify airborne LiDAR data into ground points and building points. The ground points serve as input for Digital Terrain Model (DTM) generation in ArcGIS and the building points for the measurement of building heights with ArcGIS or LP360 viewer.
- 2) Use an orthophoto or a Google Maps image, georeferenced to the LiDAR data, as background for digitizing object footprints i.e. buildings, roads, trees in ArcGIS. For building footprints, 2D shapefiles are used for block models (LOD 1) and 3D shapefiles for complex models (LOD 2 and 3).
- 3) With the geometry created, the CityGML model is generated in LandXplorer as follows
 - a. LOD 0 - created from DTM generated from airborne LIDAR ground points.
 - b. LOD 1 - created by extrusion of 2D building footprint shapefiles.
 - c. LOD 2 - created from both 2D and 3D building footprint shapefiles. Roof shapes are explicitly defined by the geometry information in the 3D shapefiles. Facades are created as in b.

- d. LOD 3 - created by aggregation of various 3D faces from CAD and 3D shapefiles captured by (mobile) LiDAR-aided drafting.

3 Data and Software

Various data sources and software tools have been used in this project. The following shortly outlines those used. The LiDAR point clouds are stored in the LAS format which is supported by most 3D mapping tools.

3.1 Data

The following data has been used for the creation of various LODs of HfT Stuttgart CityGML models:

- i. Airborne LiDAR data for DTM creation and as a source of 3D information for buildings.
- ii. Mobile LiDAR data as a source of 3D building facade information.
- iii. Google image as an auxiliary background in the digitization of building footprints.
- iv. Facade pictures for realistic facade texturing.

3.2 Software

- i. A number of software packages have been used in the creation of the 3D models:
- ii. TerraSolid for Airborne LiDAR classification.
- iii. ArcGIS 9.3 for DTM creation, digitizing of building footprints.
- iv. LP360 viewer for visualization of both airborne and mobile LiDAR data.
- v. Cyclone for the clipping and cleaning of LiDAR data.
- vi. AutoCAD Civil 3D with Cloudworx extension for the easier digitizing of facade entities on LiDAR data.
- vii. LandXplorer Studio Pro 2009 as the main software for CityGML creation.

4 CityGML Data Preparation

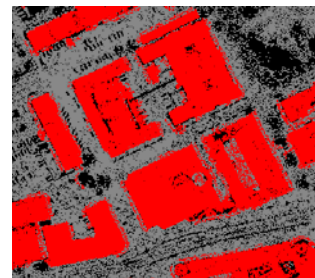
This section briefly describes some of the data processing required in the adopted CityGML creation procedure. These include LiDAR data classification, DTM generation and object footprint digitization.

4.1 LiDAR Data Preprocessing

The airborne LiDAR data has been classified into building and ground points using the TerraSolid software. Default setting for the classification of ground and building points have been used. Figure 2 visualizes the classification result.



Unclassified LiDAR data



Classified LiDAR data: ground (gray) and buildings (red)

Figure 2: LiDAR data classification

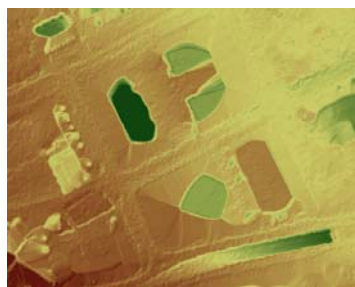
4.2 DTM generation

For the interpolation of the DTM ESRI ArcGIS has been used. ArcGIS has a number of DTM interpolation routines including Inverse Distance Weighting (IDW), Krigging, and Nearest Neighbor etc. These routines, which take either point or raster data, can be accessed from either the 3D Analyst or Spatial Analyst toolbars.

The first step in the generation of the DTM involves the conversion of the LIDAR data from LAS format to a Multipoint Shapefile which ArcGIS can handle. This *LAS to Multipoint* tool is located under the *Conversion* toolset of the 3D Analyst Extension. To convert a LAS file one specifies the input LAS file, the output point shapefile, the average point spacing, the class code etc. The class code is particularly important because it controls which LiDAR points are converted i.e. ground, building or all points. Once this is done the resulting point file can then be added to the program and used as input for the interpolation.

One obvious source of problem from this result (Figure 3) is the presence of big voids in the building locations. These voids are not correctly interpolated due to the absence of height data. Although, these holes may not cause major problems as they are covered by the building models for LOD 1 to 4 some corrective measures have been taken.

These voids have been partially corrected by first interpolating the DTM with imported ground points and then filling the resulting DTM using filling methods which are common in hydrological applications (Matkan et al, 2007). IDW has been used for the creation of the DTM because of its faster processing for a huge number of points. The filling step corrected most of the void areas as seen from the resulting height range and the disappearance of the green (lowest) areas in the output DTM.



a) Height range: 195.861-313.216m



b) Height range: 297.083-313.216m

Figure 3: Result: a) Initial DTM result b) DTM after filling

4.3 Object Footprint Digitization

Object footprints have been digitized using ArcGIS with a Google Maps image as a background to aid the digitizing, the image having been georeferenced to the LiDAR data.

Building footprints are captured as polygons outlining the extents of the building and roof faces. Depending on the level of detail 2D or 3D shapefiles are created.

2D shapefiles are sufficient for flat and block models in LOD 0 and LOD 1 respectively. In capturing these footprints it is necessary to define a *height* field in the shapefile attribute table. The height column is populated either manually i.e. inputting measured height values for each building or automatically by creating a spatial join with the LiDAR building points shapefile where a mean, highest, lowest value from the LiDAR data may be used.

For explicit roof structures as in LOD 2 and 3 the 3D shapefiles are more appropriate. In this case more detail is captured to completely define the shape of each roof (see figure 4). Each roof face is digitized with each of its vertices assigned an elevation value to completely define its 3D form/structure. Again this can be done either manually or automatic. The manual option is achieved by editing the height values in the *Sketch properties* of each polygon while the automatic option can be done by use of a *Functional surface*² or performing a 2D/3D conversion. While the automatic method may sound more interesting it is not a perfect process and usually requires more manual vertex editing.

² Available in Functional Surface toolset of the 3D Analyst Extension



a) LOD 0 -2 2D Building outlines



b) LOD 2 -3 3D Building outlines(all faces)

Figure 4: Building outlines for various CityGML LODs

Road objects are also captured as polygons. Since these are simply draped on the terrain in the city model 2D shapefiles are sufficient. Only major roads passing through and around the university have been captured. Trees on the other hand are better captured as points. The type of each tree e.g. oak, chestnut, etc may be recorded in the shapefile attribute table.

4.4 CityGML Creation in LandXplorer

LandXplorer Studio Professional offers an interactive, real-time authoring environment that allows one to visualize 3D geospatial data and to effectively create, analyze, manage and distribute geospatial information³.

Creating a model in LandXplorer begins by the creation of a new blank project. All object data for the city model are then added as layers, one at a time, until the model or scene is completely defined.

The creation LOD 0 is the simplest of all levels which only requires the insertion of a DTM. An existing LandXplorer terrain model or other third party formats such as TIFF, IMG or ASCII grids may be used. The inserted terrain model is usually colored gray. To give it a more realistic look georeferenced imagery e.g. orthophotos, Google map screen shots, satellite imagery etc, is draped on it. More information may be provided by draping 2D vectors files e.g. roads trees. Labels may also be inserted to give more information about the features in the model.

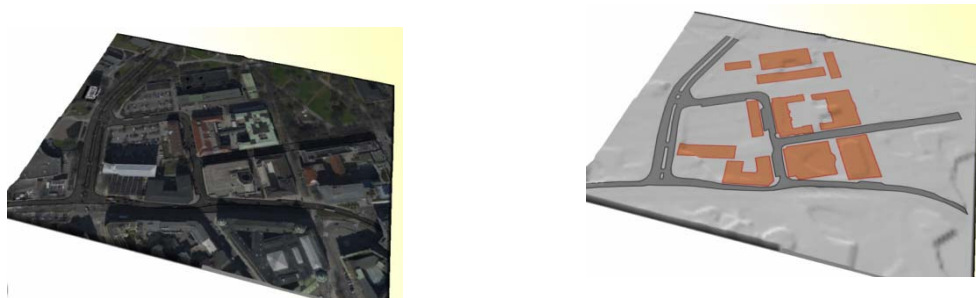


Figure 5: LOD 0 - With and without image draping resp.

The LOD 1 block model is created by extrusion of 2D shapefiles of the building footprints. The extrusion is based on the building height inputted manually or from data held in a shapefile attribute table.

³ AutoDesk (2009): LandXplorer Manual



Figure 6: HfT Stuttgart LOD 1 CityGML (without and with facade texturing resp.)

To improve the appearance of the building, roof textures and facade textures may be inserted as shown in figure 7. LandXplorer allows the use of terrain textures as roof textures. So in the figure the roof textures are derived from the image draped on the terrain. For facade textures a uniform facade texture element has been used. Lastly, more objects such as roads (gray in figure) and trees may be inserted to realize a complete scene.

In the LOD 2 building models are created by merging of block models (LOD 1) and detailed roof structures from 3D shapefiles. This is adopted because of LandXplorer inability to automatically produce facades from 3D shapefiles (the hanging roof problem). Roof and facade textures are then inserted to improve the appearance of the model. For LOD 2 detailed facade textures were captured as part of a student project and manually rectified using the Gimp software. Trees have also been included as shown in the figure (Figure 7) below.



Figure 7: LOD 2- Some views of Building 1 of HfT Stuttgart

The procedure for creating LOD 3 is still under evaluation. One idea that has been fairly tested is the use of the Leica Cloudworx extension for AutoCAD. The extension enables AutoCAD to handle LiDAR data and also offers a number of functions such as creating sections, hiding and unhiding section of the LiDAR data and others which can help to draft facade objects such as windows, doors etc. Mobile LiDAR data has been used in this case. Figure 8 shows some preliminary results of a facade captured using this method.



Figure 8: LiDAR-aided drafting

One of the unresolved problems in this procedure is finding a suitable export format which LandXplorer can accept or handle. The STL format has been used but the results are not that satisfactory.

5 Discussion and Conclusion

Some procedures to create various CityGML levels of detail using mixture of tool and data from various sources have been outlined. Airborne and mobile LiDAR data has been used as a source of 3D information. Other data

such as images have been used to complement the LIDAR data e.g.in the digitizing of object footprints. This is a good example of a fusion of data and tools to solve the city modeling problem.

While this demonstration has been done for the HfT Stuttgart campus i.e. a small area, the procedures used can as well applied to large areas. Many cities already have GIS data of building footprints so LOD 1 and 2 may readily be created if the relevant height information is added to the data. A good example is the LOD 1 and 2 CityGML data set created for the city of Frankfurt (Russ, 2009).

Lastly, CityGML models can benefit a lot of industries as stated in the introductory remarks. Therefore, procedures for their creation should continually be developed and improved. Currently, most of the successful software solutions in this respect are commercial and come at high prices. Data for CityGML creation also comes at a cost. However, a number of open source alternatives are now on the increase as shown by the numerous post of the CityGML website (<http://www.citygml.org/>) which lists various sources for CityGML software and data. So the possibility of using open source tools in the procedures outlined here is good.

References

- Matkan, A. Shakiba, A. and Ashorlo, D. (2007). *An Evaluation of Hydrological Pit Filling Methods in the Ekbatan Watersheds*. Environmental Sciences Vol.4, No.3 , Spring 2007, pp. 49-60.
- Open Geospatial Consortium Inc. (2008). *OpenGIS City Geography Markup Language (Citygml) Encoding Standard*. Berlin: Open Geospatial Consortium Inc.
- Russ, C. (2009). *Gebäuderekonstruktion Aus Lidar-Daten Frankfurt/Main*. Virtual City Systems: http://www.virtualcitysystems.de/fileadmin/pdf/case_studies/CaseStudy_Frankfurt_LoD1_LoD2_web.pdf (accessed 12 July 2010)
- Stadler, A. & Kolbe, T. (2008). *CityGML - Modelling our Environment*. from Technische Universität Berlin: http://www.igg.tu-berlin.de/uploads/tx_ikgpublication/cityGml_Stockholm.ppt (accessed 10 July 2010)
- Theng, L. B. (2006). *Automatic Building Extraction from Satellite Imagery*. Engineering letters, http://www.engineeringletters.com/issues_v13/issue_3/EL_13_3_5.pdf (accessed 14 July 2010)
- VEPS. (2010). *3D City Models*. Virtual Environment Planning: <http://www.veps3d.org/site/249.asp> (accessed 12 July 2010, from)
- Wikipedia. (2010). *CityGML*. <http://en.wikipedia.org/wiki/CityGML> (accessed 08 July 2010)

Acknowledgement

The authors would like to acknowledge the following students in the Photogrammetry and Geoinformatics master course at HfT Stuttgart for their contribution to the results of this project: Imran Hossain and Shakti Gahlout.

Low-Cost Airborne Platforms and Remote Sensors for Precision Farming

Andres Flores ^{1,*}, Gonzalo Cucho ¹, Ernesto Gordillo ¹, Marlon Aucallanchi ¹, Renán Rojas ¹, Jorge Barba ¹, Aurelio Rodríguez ², Roberto Quiroz ², Adolfo Posadas ², Carolina Barreda ² and Hildo Loayza ²

¹ Sección de Electricidad y Electrónica, Pontificia Universidad Católica del Perú, Lima, Peru, dflores@pucp.edu.pe

² Production Systems and the Environment, International Potato Center, Lima, Peru, R.QUIROZ@cgiar.org

KEYWORDS:

Earth observation systems, Remote Sensing, Agriculture, Electronics, Image Registration

ABSTRACT:

A low-cost remote sensing system for the acquisition and processing of georeferenced multispectral images of crop fields for precision agriculture applications has been developed. The system has been implemented using various types of low-cost airborne platforms with electronic equipment consisting of multispectral cameras and GPS. The acquired images are processed using algorithms that extract the information and determine the status of crops.

1 Introduction

Observation of the earth by remote sensing has been in recent years applied to solve multiple problems. It is used to make decisions about several aspects involving human activities of land (Lillesand, T, 2007; Sabins, F., 1997). It is currently applied in multiple applications, one of which is precision farming. Remotely sensed data can provide spatial and temporal information necessary to address issues in farm management decisions. Proper management can be crucial to the efficient use of inputs such as water, fertilizers and pesticides for food security and can also contribute to the protection of the environment. Air- and space-borne platforms have been extensively used for this purpose. Satellite imaging has developed rapidly in recent years and it is possible now to access land data with very high spatial and temporal resolutions, but the main drawback of optical remote sensors is that they can not register land surface responses with overcast skies. This is the case for winters in coastal regions of Peru and during the rainy cropping seasons elsewhere where crop reflectances to the EM spectrum can not be registered by sensors onboard of commercial satellites. Aerial imaging is currently a possible solution to effectively monitor crops; one example is the Aerial Photography Field Office (APFO) in the U.S. that provides this type of service to farmers. The main goal in this project is to acquire low-cost images from air-borne platform and sensors. Balloons and model aircrafts and helicopters have been built to study its application under different geographic and climatic conditions. An embedded multispectral acquisition system with GPS, that provides images with geographic information, has been developed. The project is a joint effort by the Pontificia Universidad Católica del Perú and the International Potato Center both located in Lima, Peru.

2 Platforms

2.1 Balloon Based System

Two types of balloons are being used as platforms for this application. A propane-based hot air balloon suited with a thin capillary tube transporting the gas from the ground was developed by the International Potato Center. Notwithstanding the advantage of its low cost, the system was very difficult to control in windy conditions which limited its functionality as a monitoring platform in the coastal region of Peru.

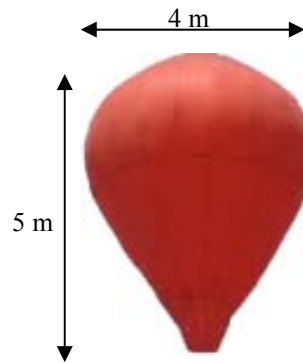


Figure 1: Hot air balloon dimensions.

A better performance was obtained using a Helium based balloon which functioned very well in windy conditions. The main drawback with this system is the high cost of Helium. A wooden platform was designed that carries the image acquisition system.

2.2 Radio Controlled Model Airplanes

Considering payload requirements, radio controlled airplanes were designed. Kadet and Super Squire type aircraft were chosen due to their good stability and simple flight control. Three aircraft were constructed with different types of engines. Two Kadet planes had a wingspan of 2.4m and fuselage of 1.8m, the third airplane is a Super Squire with a wingspan of 1.97m and fuselage length of 1.30m. One of the planes uses an internal combustion “Glow” engine as shown in figure 2.

The other two planes are equipped with brushless electric motors. Test flights confirmed the convenience of electric motors over combustion engines due to quality of images affected by engine vibration and oil in the exhaust. Total flight duration is approximately 25 minutes and can attain elevations of up to 500 meters. Range is limited to the operator visibility or the transmitter capability which can be up to 800 meters depending on the size of the aircraft. Brushless electrical based airplanes are equipped with Lithium Polymer batteries which have proven to be very convenient.



Figure 2: Combustion Glow engine based airplane and T-Rex600 Helicopter

2.3 Radio Controlled Model Helicopter

Helicopters can provide a more precise measurement especially if a specific location is required. For this reason an Align T-REX 600 GF KIT Helicopter was assembled and operated using a payload of 1 Kg. It has a brushless motor and its length is 1.2m and height is 0.4m. Its main drawback is that its flying range is smaller than the airplanes because of its relatively smaller size. The operator’s visibility is reduced and is unable to travel large distances, it is also more difficult to maneuver than the airplanes.

3 Embedded Multispectral Image Acquisition System

Monitoring crops involves the use of imaging sensors with sensitivity in the visible red (650nm) and near infrared (850nm) light wavelengths. CCD and CMOS sensors are arranged in cameras that are able to capture images in wavelengths in the visible region of the spectrum (400 to 700nm). But can also have sensitivity in longer wavelengths up to 950nm which is convenient for this application. Cameras such as the JAI Monochrome CV-A50 IR and Point Grey Firefly MV have these capabilities and are well suited for this purpose.

Digital images are required for each bandwidth, cameras equipped with optical filters that capture light reflected from crops in the selected wavelengths. The two cameras are part of the embedded multispectral image acquisition system that is transported within the platforms described above.

3.1 Image Acquisition System Characteristics

1. Two cameras equipped with appropriate lenses and optical filters with appropriate electronic interfaces.
2. GPS device with USB 2.0 interface.
3. Energy efficient Intel Atom based computer with SSD.
4. Linux operating system. Xubuntu distribution.
5. Software was developed to manage the hardware in order to acquire multiple georeferenced images. The system is able to capture 100 images per second. OpenCV was used for preliminary processing of images.

Two computer systems were used, one based on the Intel D945GCLF Mini-ITX form factor. The other is the FITPC2 which is a smaller and lighter version of Atom system. Both systems worked adequately.

4 Digital Image Processing

In order to extract useful information from the airborne images, several aspects involving digital image processing were considered. Software was developed using Matlab and the Open Computer Vision Library (Bradski, G., 2008) in the C programming language. Images were processed using a desktop computer on land.

4.1 Image registration

It is necessary to reconstruct the crop scenes on a single high resolution image. This requires joining many images in order to assemble a mosaic. This is done using the following algorithms:

1. Harris-Laplace and Speeded Up Robustness Features (SURF) for the detection of control points.
2. Scale-invariant feature transform (SIFT) for the control points description.
3. NNDR for determination of correspondence of control points.
4. Estimation of the Transformation function.
5. Transformation of the test image to the base image.

This procedure was done for the visible red and near infrared images acquired from an altitude of 200m above the crops. The system is capable of merging as many images as necessary. Figure 3 shows the registration of ten VGA (640x480 pixels) infrared images over a scene that covers an area of approximately 12760 m². The image obtained in this example has an 1879x1213 pixels resolution. As can be seen the system is capable of assembling high resolution imagery for precision farming.

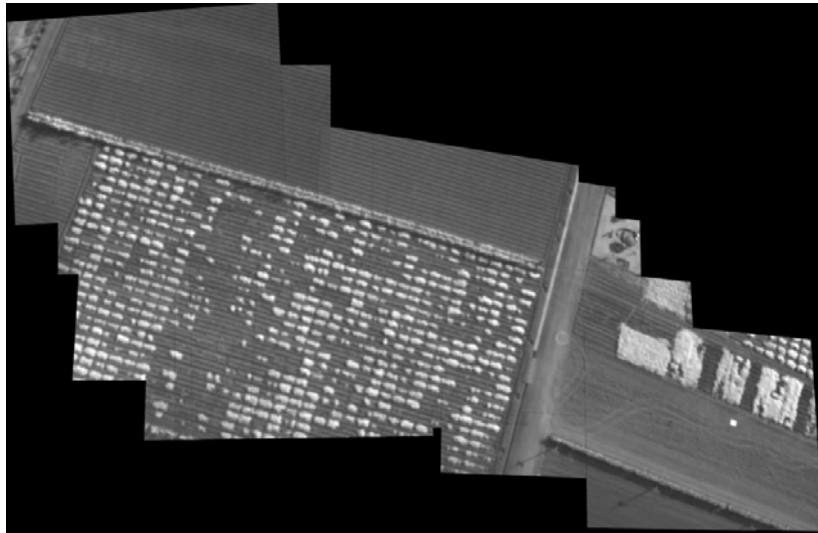


Figure 3: Mosaic obtained by registering several Infra Red potato images.

4.2 NDVI Determination

The Normalized Difference Vegetation Index (NDVI) is a simple numerical indicator that can be used to analyze crop vigor and can help to determine its status (Landgrebe, D., 2003). It requires proper spatial alignment of the red and infrared images. Figure 4 shows a spatial aligned image in which Red and Infrared images were merged using registration techniques. It is a false color image that clearly depicts vegetation and soil characteristics. We call this a multispectral image that can be used as an indicator of presence of vegetation.

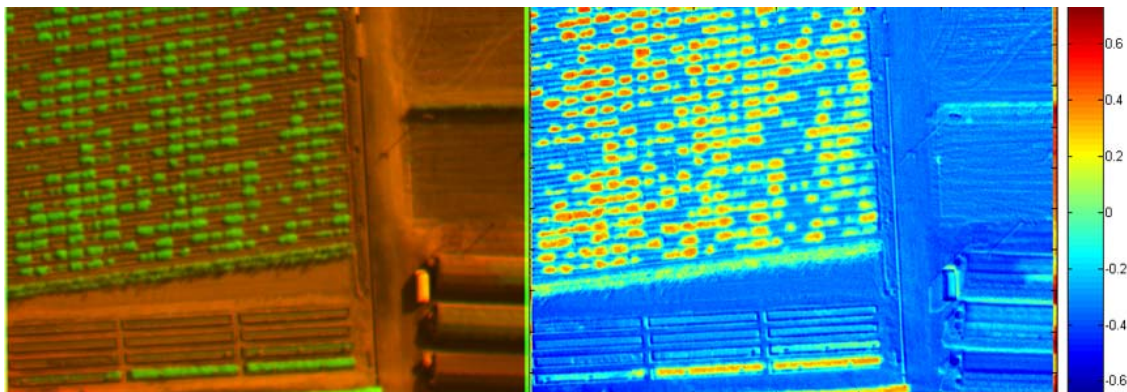


Figure 4: False color Image and NDVI image

Equation (1) describes how NDVI is obtained for a particular scene.

$$NDVI = (NIR - RED) / (NIR + RED) \quad (1)$$

Figure 4 shows a false color image depicting NDVI for a potato crop in early stage of development. As can be seen high values of NDVI can indicate that particular areas have been properly used, while a low NDVI value indicates that the corresponding portion of land is not used properly and correct decisions must be made if the goal is to achieve high productivity. Other types of vegetation indices such as PVI and SAVI can be calculated with these data, and are currently under development.

5 Conclusions

A first step towards a remote sensing system was developed using low-cost airborne platforms containing hi-res sensors and hardware. Preliminary tests show promising results that prove the feasibility of the system. Georeferenced multispectral images are easily obtained in an ordered manner using image processing techniques. Information regarding the status of the crop is yet to be determined and will be the next step of this

project. Sensor calibration, a necessary step, is at an advanced stage of development using a spectroradiometer. This project will be a starting point to improve resource-poor farmer's yield through precision agriculture and thus contributing to their well being.

References

- Bradski, G., 2008. *Learning OpenCV: Computer Vision with the OpenCV Library*, O'Reilly Media.
- Espinosa, J, Mite, F, Cedeño, S, Barriga, S, Andino, J., 2006. *GIS-Based Site-Specific Management of Cocoa, Better Crops*.
- Fedorov, D, Fonseca, L, Kenney, C, Manjunath, B., *Automatic Registration and Mosaicking System for Remotely Sensed Imagery*, Dept. of Electrical and Computer Engineering, UC Santa Barbara, CA.
- Herwitz, S.R., Johnson, L.F., Dunagan, S.E., Higgins, R.G., Sullivan, D.V., Zheng, J., Lobitz, B.M., Leung, J.G., Gallmeyer, B., Aoyagi, M., Slye, R.E. and Brass, J.. 2003. *Coffee field ripeness detection using high resolution imaging systems of a solar-powered UAV*, Proceedings of the 30th International Symposium on Remote Sensing of Environment, Honolulu, HI.
- Landgrebe, D., 2003. *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley and Sons, New York.
- Lillesand, T, Kiefer, R and Chipman, J , 2007. *Remote Sensing and Image Interpretation*, 6th ed. Wiley.
- Mikolajczyk, K., 2006. *A Comparison of Affine Region Detectors*, International Journal of Computer Vision. Springer Science + Business Media, Inc.
- Richards, J, Jia, X, 2006. *Remote Sensing Digital Image Analysis*, 4th Edition, Springer-Verlag Berlin Heidelberg.
- Sabins, F., 1997. *Remote Sensing: principles and interpretation* 3rd ed., Freeman and Company.
- Simard, R, Nolin, M and Cambouris, N, 1998. *Application of Precision Farming to Potato Production in Québec*, Better Crops.
- The Aerial Photography Field Office, <http://www.fsa.usda.gov/FSA/apfoapp?area=about&subject=landing&topic=landing> [accessed 07.19.2010]

Euler Deconvolution of Potential Field Magnetometer and Its Use in Delineation of Regional Faults with Data from Northern Peru

C.J. Solano Salinas^{1,2} and E. Borda³

¹ Centro de Tecnologías de Información y Comunicaciones

² Computer Science School, University of National Engineering

Av. Tupac Amaru, 210, Lima (Peru), - jsolano@uni.edu.pe

³ PETRO-TECH PERUANA S.A.

Av. Los Incas, 460, Lima (Peru), - eborda@gmail.com

KEYWORDS:

Euler Deconvolution, geomagnetic field, Northern Peru

ABSTRACT:

The geomagnetic field measurements can be used to determine the structure of the earth, because the rocks often contain magnetic minerals. The interpretation of data of this nature will determine in some cases, geological characteristics that would help contribute to the success of mining or oil exploration.

The methodology to calculate the depth and location of magnetic sources (dam and failure), through the Euler Deconvolution method and application is made for three magnetic profiles extracted from a data grid in northern Peru. It then lists a series of conclusions about the potential of this methodology.

1 Introduction

Traditional methods in interpreting gravitational magnetometer require, in addition to experts in the field, knowledge to perform geological and geophysical modeling (Nabighian *et al.* 2005).

There are other methods, called semi-empirical, requiring no other knowledge or other specialists. These methods are applied directly to the data, gridded and mapped. They are particularly successful in detecting contacts, with differences in susceptibility, structures similar to dikes, spheres, horizontal cylinders, vertical pipes, etc.

One method widely used to estimate the depth and location of the sources that cause the magnetic anomalies, is the method of Euler Deconvolution.

2 Euler Deconvolution for Magnetic Dam and Failure

Thompson in 1982 proposed a technique for analyzing magnetic profiles based on Euler's relationship for homogeneous functions. The Euler deconvolution technique uses first-order derivatives of x , y , and z to determine the location and depth of various idealized sources that produce them. Among the most common are the sphere, cylinder, thin dike and the failure, the latter two being the case studied in this paper. Each of these sources are characterized by an structural index, the method is in principle applicable for all types of bodies. Reid (1990) extended the technique to 3D data by applying the Euler operator data windows on a grid. Silva and Barbosa (2003) and Introcaso *et al.* (2008) among others helped in the understanding and applicability of the technique.

2.1 Theory

Consider any function of three Cartesian coordinates x , y , and z denoted by $f(x, y, z)$. The observation plane is taken as $z = 0$ and positive down and based on that choose the direction of x is north and y is east.

2.1.1 Euler equation

The function $f(x, y, z)$, is said to be homogeneous of degree n if:

$$f(tx, ty, tz) = t^n f(x, y, z) \quad (2.1)$$

In addition, we show that if $f(x, y, z)$ is homogeneous of degree n , the following equation is satisfied.

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} = nf \quad (2.2)$$

This partial differential equation is known as Euler's equation of homogeneity or simply Euler's equation. Suppose that $f(x, y, z)$ has the following general form:

$$f(x, y, z) = \frac{G}{r^N} \quad (2.3)$$

where $r^2 = x^2 + y^2 + z^2$

$N=1, 1.5, 2, 2.5, 3, \dots$

G : constant independent of x, y, z

Equation 2.3 is clearly homogeneous of degree $n = -N$. Many single point magnetic sources are in the form of this equation.

2.1.2 The structural index

Consider a point source (point mass, magnetic dipole, etc) located at the point x_0, y_0, z_0 on the measurement plane. The total magnetic intensity is of the form:

$$\Delta F(x, y) = f[(x - x_0), (y - y_0), z_0] \quad (2.4)$$

Reid et al. (1990) proposed the relation:

$$(x - x_0) \frac{\partial \Delta F}{\partial x} + (y - y_0) \frac{\partial \Delta F}{\partial y} - z_0 \frac{\partial \Delta F}{\partial z} = -N \Delta F(x, y) \quad (2.5)$$

The gradients in the three Cartesian directions can be calculated using standard potential theory in the domain of space or wave number. In some cases the vertical gradient could have been measured and used directly in Equation 2.5. This equation could be used to analyze mapped magnetic data (Cook 1950).

Let us assume that the transverse gradient, the second term of equation 2.5, is zero. This would be a 2D anomaly in which the equation reduces to the expression:

$$(x - x_0) \frac{\partial \Delta F}{\partial x} - z_0 \frac{\partial \Delta F}{\partial z} = -N \Delta F(x, y) \quad (2.6)$$

The derivatives or gradients in Equation 2.6 can be measured or more commonly calculated from the data, and it would therefore only be sufficient to know the x_0, z_0 and N . The coordinates (x_0, z_0) represent the depth and location along the profile of an equivalent point source and N correspond to the kind of font that best represents the type of anomaly. Several simple models have prescribed values of N (Smellie 1956).

Many geological features have structural indices that best describe its depth and location, so much so that for example a thin 2D dam has a structural index $N = 1$ at the magnetic pole, while a contact has a structural index lower than 0.5, as we checked with this method.

Equation 2.6 can be solved exactly for the unknown values x, z and N evaluating the derivatives and the total field values in the three different coordinates along the profile. This result in three linear equations and three unknowns which can, in principle, be solved if the determinant of coefficients is not zero.

The implementation of Equation 2.6 on directly observable data is not useful for three reasons:

- Many anomalies, even at the magnetic pole, prefer high structural indices, i.e., they are more bipolar in nature. However, the low structural indices are better estimates of depth (Thompson 1982).
- The absolute level of the field is rarely known. The regional fields due to local anomalies are almost always present.
- On real data, the anomalies are rarely represented exactly by point sources.

These factors make the exact solution of Equation 2.6 very suspicious and erratic. To overcome these problems, some methods have been developed that will be described below:

The problem of forcing the method to lower depths for structural indices is determined by an analysis for a number of prescribed structural indices. Then one checks which one produces better solutions by observing the cluster of solutions located in a region of space.

The problem of removing the "bias" of the observed magnetic field is to solve it by assuming that the anomalous field is disturbed by a constant amount B in a window in which Equation 2.6 is being evaluated. The observed number is:

$$F(x) = \Delta F(x) + B \quad (2.7)$$

where B is constant in the x coordinate of the portion of the profile where the analysis is being performed. Finally solving equation 2.7 for ΔF and substituting in equation 2.6 and arranging the terms give:

$$x_o \frac{\partial \Delta F}{\partial x} + z_o \frac{\partial \Delta F}{\partial z} + NB = x \frac{\partial \Delta F}{\partial x} + NF \quad (2.8)$$

Since the real anomalies are only approximations of the simple models, the third main problem is solved by creating an overdetermined set of linear equations. If one evaluates the equation 2.8 in four or more points within a window of a profile, this results in an overdetermined set of linear equations. For this analysis, we used a total of 7 points, i.e. the creation of a total of seven equations for three unknown variables x_o , z_o and B , which will be solved by using the method of least squares.

1.1.1 Application of Euler deconvolution to the case of a dam and a failure

The solution by the method of least squares of an overdetermined set of equations also produces estimates of the standard deviation of the parameter z_o , σ_z . This amount is treated as an error bar on the estimated depths and forms the basis for an algorithm that determines whether a solution is accepted or not.

The magnetic anomaly caused by a magnetic body can be exactly duplicated by placing a suitable distribution of magnetic poles on the surface of the body that causes it. A thin magnetic intrusive body would have induced positive pole on its upper surface and negative poles induced in the subsurface. In the distance, such magnetic sources have bipolar features. On the other hand, the magnetic bodies are intrusive into the deep crust, would have induced poles on its upper surface. But their negative poles for many would be even deeper, so they do not contribute to the total magnetic field. According to this simple reasoning these bodies show polar behavior.

We created an algorithm in MAPLE to perform this study and to calculate the fields and derivatives in the x and z directions in order to form the Euler equations and its corresponding solutions, through the least squares method. From this analysis we obtain the values of $N=0.5$ for the dam and $N=1.0$ for the failure.

We should clarify that the formulation of programs developed for both the dike failure are unrestricted with the magnetic inclination I , however, the best estimate of the depths will be given for those data that are reduced to the magnetic pole.

3 Application to Data of Northern Peruvian Coast

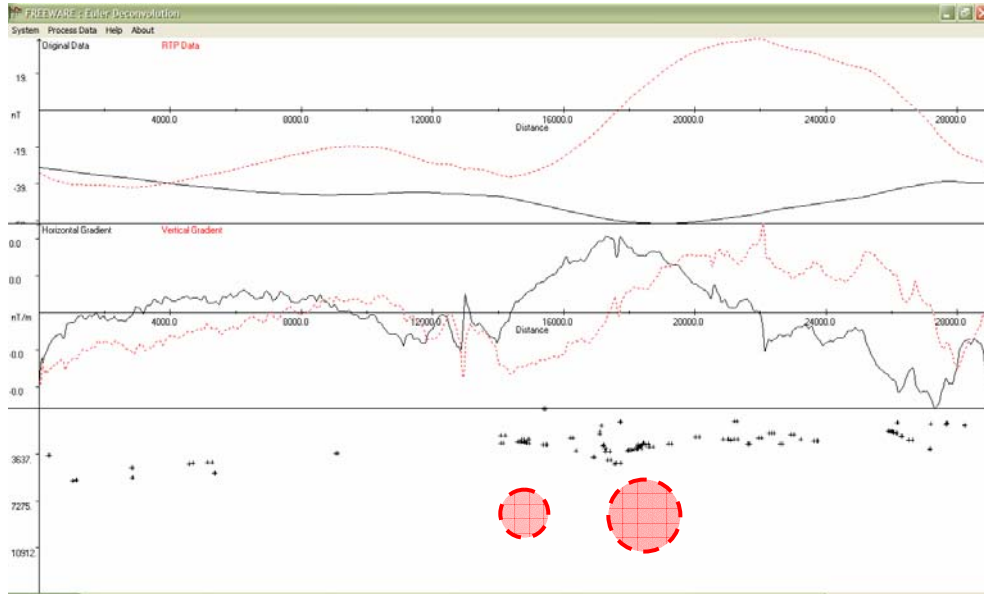
We used high-resolution magnetometer data of northern peruvian coast, planning for this flight lines 500 m apart in the north-south direction and 1500 m for mooring lines in the east-west direction. The flight altitude was 150m above the surface.

In order to extract two-dimensional profiles three lines were generated whose location is shown in a magnetometer map of the studied area, Figure 2.

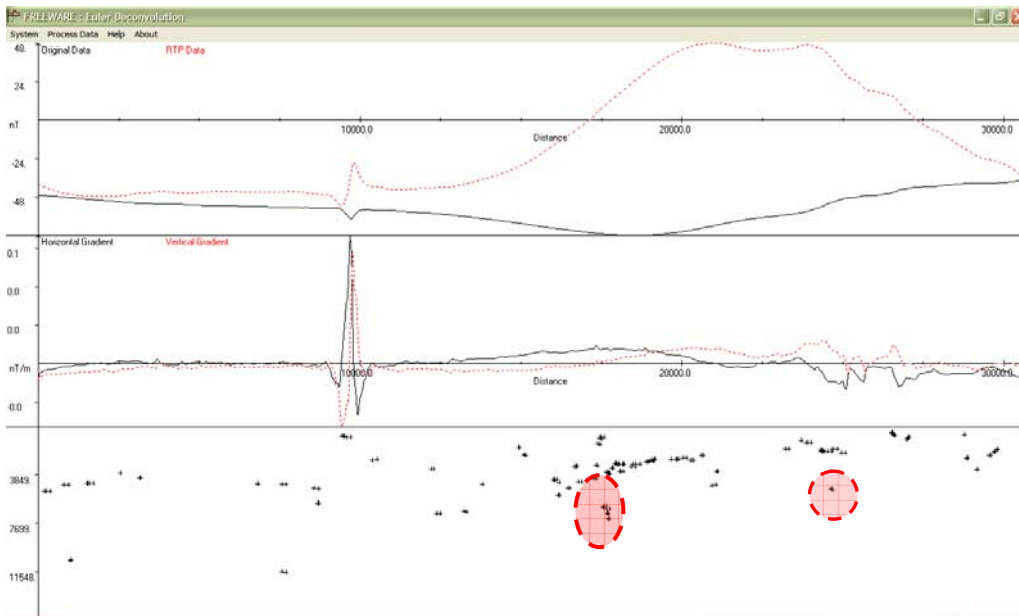
After obtaining the magnetic profiles we proceeded to introduce the values of magnetic field and geographic coordinates to a free software called "**Euler Deconvolution**", in which the data of magnetic declination for the area of -14.2° and slope of -2.2° , was entered in order to reduce the data to the magnetic pole. In Figure 1a, b, c this correction can be seen at the top in lighter color. Besides that you can appreciate the vertical and horizontal derivatives of the field, as was necessary to solve the Euler equations.

In order to observe clusters of solutions in certain areas and to associate them with the existing faults in the area put some limits on these solutions with depth. For example it is known that the Paleozoic rocks in the area is at a variable depth of no more than 5 km so, those solutions greater than this value are simply discarded.

With the results obtained by modeling the faults, we find that the optimal value to determine the depth and location for the failures was using a structural index $N = 1.0$. Thus introducing all these values within this program were obtained the solutions shown in Figure 1. For example in Figure 1a, b, c, which corresponds to the profile 1, 2 and 3, you can see a clustering of solutions in the area marked in circles, which gives us an idea that this area could be a rambling a failure at the level of the Paleozoic rocks.

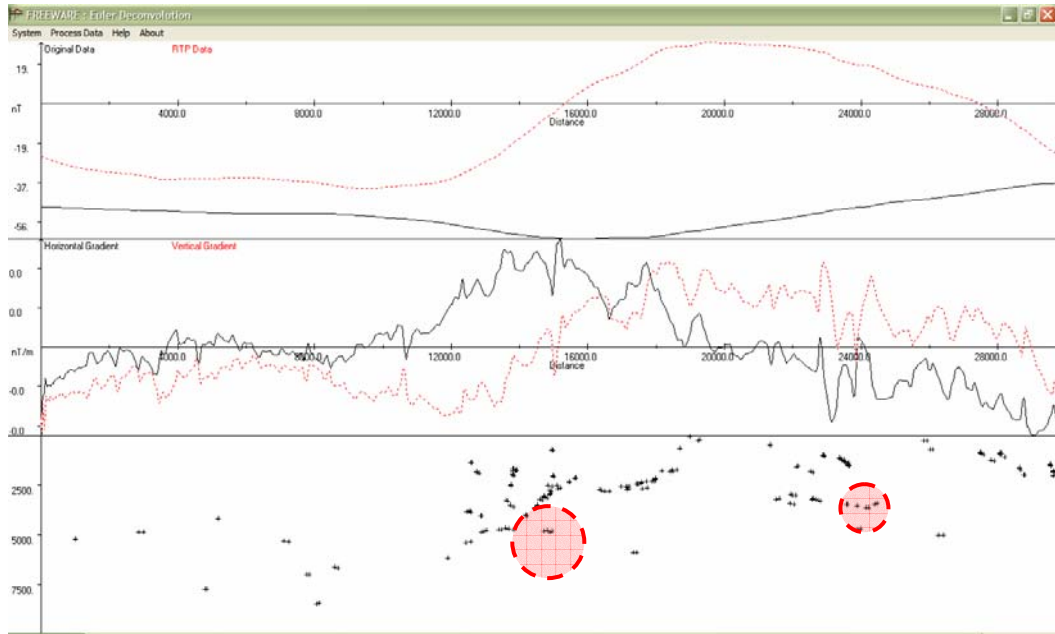


(a)



(b)

*Euler Deconvolution of Potential Field Magnetometer and
Its Use in Delineation of Regional Faults with Data from Northern Peru*



(c)

Figure 1: Magnetometer profiles over the three profiles extracted with the respective estimates of the sources giving rise to the magnetic anomaly. (a) Profile 1, (b) Profile 2 and (c) Profile 3. We used structural index $N = 1.0$, which according to previous tests corresponds to better position values to determine faults

In Figure 2 the locations of the three possible failures are plotted according to the results obtained in the interpretation. What we noticed is that the fault deepens more in the West to the South, because the solutions, as we go from profile 1 to 3, are expanded. The other failure, closer to the sea, correspond to a failure shallow (depth less than 2.5 km). This interpretation was corroborated with the ones obtained by in exploration companies through other methods such as seismic.

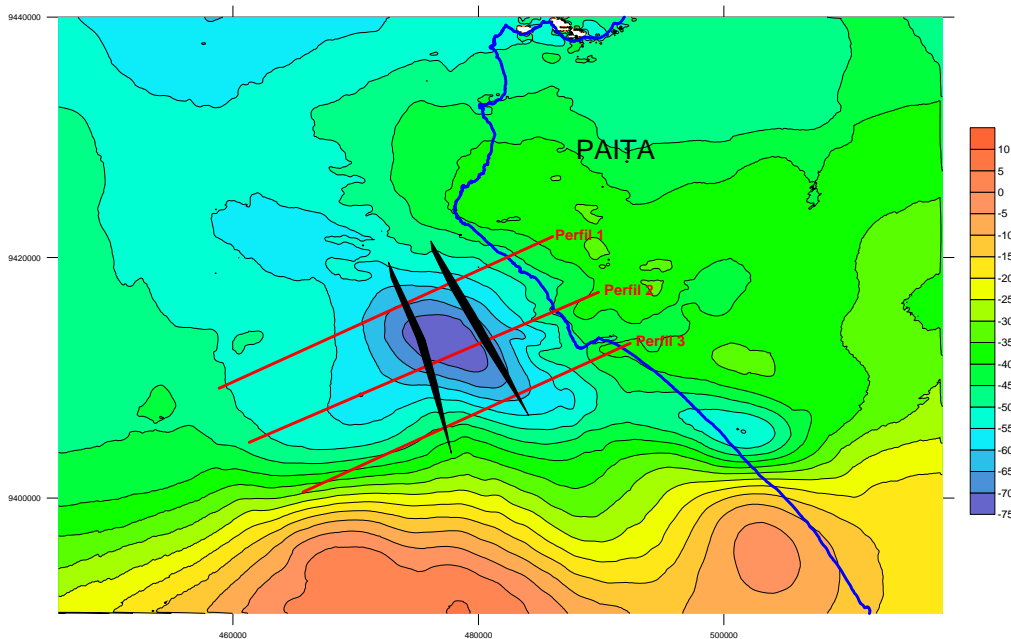


Figure 2: Solutions to the Euler equations along three profiles, where we denote the presence of two major faults

4 Conclusions

We could verify compliance of the homogeneous Euler equation for simple cases of a vertical prism of infinite length, as well as in the case of a failure.

By applying Euler deconvolution of the two simple cases it was noticed that the best values to predict both the depth and position of these simple sources were $N = 0.5$ for the case of infinite vertical prism, while for the fault the position was best predicted with $N = 1.0$. From these values we can conclude that in both cases, the character of the magnetic field strength decreases with distance.

Applying this method to real cases, we see that there is good correspondence between the solutions obtained and those interpreted through other methods such as seismic.

References

- Cook K.L., 1950. Quantitative interpretation of vertical magnetic anomalies over veins. *Geophysics* Vol. 15, pp. 667-686.
- Introcaso A., Ghidella M.E., Ruiz F., Crovetto C.B., Introcaso B. and Paterlini C.M., 2008. Metodos gravimagnetometricos modernos para analizar las características estructurales de la plataforma continental argentina. *Geoacta* 33, pp. 1-20.
- Koulomzine T.H., Lamontagne Y. and Nadeau A., 1970. New methods for the direct interpretation of magnetic anomalies caused by inclined dikes of infinite length. *Geophysics* Vol. 35 N° 5, pp. 812-830.
- Nabighian M.N, Grauch V.J.S., Hansen R.O., LaFehr T.R., Peirce J.W., Phillips J.D. and Ruder M.E., 2005. The historical development of the magnetic method in exploration. *Geophysics* Vol. 70 N° 6, pp. 33ND-61ND.
- Qureshi I.R. and Nayale A.M., 1977. A method for direct interpretation of magnetic anomalies caused by two-dimensional vertical faults. *Geophysics* Vol. 43 N° 1, pp. 179-188.
- Reid A.B., Allsop J.M. Granser H. Mollet A.J. and Somerton I.W., 2000. Magnetic interpretation in three dimensions using Euler Deconvolution. *Geophysics* Vol. 55 N° 1, pp. 80-91.
- Silva J.B.C. and Barbosa V.C.F., 2003. 3D Euler deconvolution: Theretical basis for an automatically selecting good solutions. *Geophysics* Vol. 68 N° 6, pp. 1962-1968.
- Smellie D.W., 1956. Elementary aproximations in aeromagnetic interpretation. *Geophysics* Vol. 21 N° 4, pp. 1021-1040.
- Thompson D.T., 1982. EULDPH: A new technique for making computer-assisted depth estimates form magnetic data. *Geophysics* Vol. 47 N° 1, pp. 31-37.

Acknowledgements

One of the authors, C.J. Solano Salinas, would like to thank the support received by Centro de Tecnologías de Información y Comunicaciones and also by the Research Institute of Sciences Faculty at National University of Engineering.

Eliciting Conflict in Expert's Decision for Land Use Classification

V.K Panchal ¹, Sonakshi Gupta ², Nitish Gupta ³ and Mandira Monga⁴

¹ Defence Terrain Research Lab, Metcalfe House, Delhi, India
Tel: +91-9868328550; Fax:+91-11-23841209;Email: vkpans@ieee.org

²Jaypee Institute of Information Technology University, Sec-64 Noida, India
Tel: +91-9873281211;Fax:+91-11-23841209;Email: sonakshi_gupta89@yahoo.co.in

³Bhagwan Parshuram Institute of Technology, GGSIP University, Delhi, India
Tel: +91-9891467316;Fax:+91-11-23841209;Email: ntshgpt@gmail.com

⁴Netaji Subhash Institute of Technology, University of Delhi, India
Tel: +91-9711313103; Fax:+91-11-23841209;Email: monga.mandira@gmail.com

KEYWORDS:

Satellite Application, Decision Support, Process Modelling, Artificial Intelligence, Information Extraction.

ABSTRACT

Remote sensing data is widely used for the classification of types of land cover such as vegetation, water body etc. Conflicts are one of the most characteristic attributes in satellite remote sensing multilayer imagery. Many models for conflict resolution and analysis have been proposed and studied till time. Conflict occurs in tagging class label to mixed pixels that encompass spectral response of different land cover on the ground element. In this paper we attempted to present a new approach to conflict analysis due to mixed pixels using rough set-based model of machine learning. The paper deals with the idea of identifying the set of spectral bands contributing to conflict among opinions of experts with the help of training data sets.

1 Introduction

Remote sensing data can be put to use in classifying the features in an image into distinct categories[Kiefer 2006]. The categorized images can then be used in different ways – a farmer may use thematic maps to monitor the health of his crops without going out to the field. A geologist may use the images to study the types of minerals or rock structure found in a certain area. A biologist may want to study the variety of plants in a certain location. Remote sensing with multi-spectral satellite imagery works on the concept that different features of the land cover reflect electromagnetic radiation over a range of wavelengths in their own characteristic way according to their chemical composition and physical state. The problem arises when the terrain features do not correspond to the pure object or feature and may not have a prior spectral signature library. Pixels which are in the interface region of two classes e.g. vegetation & water body, can not be clubbed together to a single category, say vegetation, because these pixels do not follow any one particular class's spectral signature. These types of pixels are generally known as mixed pixels. Assigning a class tag to the mixed pixels is cause of conflict in the expert's mind and it is the focus point of the study.

The motivation for this paper is to observe this mixed pixel scenario and attempt to identify the conflicting spectral bands with the help of Rough Set Theory, a new paradigm of Artificial Intelligence.

The paper is organized into five sections. The section following the introduction illustrates the concepts of Rough Set, dependency of the attributes and mixed pixels. The third section describes the methodology. The fourth section reports the results and the analysis of the case study undertaken. The last section summarizes the important findings.

1 Rough Set Theory, Dependency of Attributes and Mixed Pixels

1.1 Rough Set Theory

The Rough set concept is a new mathematical approach to vagueness and uncertainty [Pawlak 1998,2007]. The rough set philosophy is founded on the assumptions that with every object of the universe of discourse we

associate some information. Any set of indiscernible (similar) objects is called elementary set, and form a basic granule of knowledge about the universe.

Any union of some elementary sets is referred to as crisp (precise) set, otherwise it is rough (imprecise, vague). Each rough set has boundary line cases, i.e. objects which cannot be with certainty classified as members of the set or of its complement. Vague concepts, in contrast to precise concepts, cannot be characterized in terms of information about their elements. So we assume that any vague concept is replaced by a pair of precise concepts- called the lower and the upper approximation of the vague concept. The lower approximation consists of all objects which surely belong to the concept and the upper approximation constitutes the boundary region of the vague concept (Pawlak, 1998).

The basic concept of the RST is the notion of approximation space, which is an ordered pair $A=(U,R)$, where

- U : nonempty set of objects, called universe
- R : equivalence relation (sometimes represented by IND) on U , called indiscernibility relation. If $x, y \in U$ and xRy then x and y are indistinguishable in A

Each equivalence class induced by R , ie, each element of the quotient set $R\sim = U/R$, is called an elementary set in A . An approximation space can be alternatively noted by $A=(U,R\sim)$. It is assumed that the empty set is also elementary for every approximation space A . A definable set in A is any finite union of elementary sets in A . For $x \in U$ let $[x]_R$ denote the equivalence class of R , containing x . For each $X \subseteq U$, X is characterized in A by a pair of sets – its lower and upper approximation in A , defined respectively as:

$$A_{low}(X) = \{x \in U \mid [x]_R \subseteq X\}$$

$$A_{upp}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$$

A rough set in A is the family of all subsets of U having the same lower and upper approximations.

1.2 Dependency of Attributes

Another important issue in data analysis is discovering dependencies between attributes. Intuitively, a set of attributes D depends totally on a set of attributes C , denoted $C \xrightarrow{k} D$, if all values of attributes from D are uniquely determined by values of attributes from C . In other words, D depends totally on C , if there exists a functional dependency between values of D and C .

Formally dependency can be defined in the following way. Let D and C be the decision and conditional subsets of A . We will say that D depends on C in a degree k ($0 \leq k \leq 1$), denoted $C \xrightarrow{k} D$, if

$$k = \gamma(C, D) = \frac{|POS_C(D)|}{|U|},$$

$$POS_C(D) = \bigcup_{X \in U/D} \underline{C}(X),$$

$POS_C(D)$ called a positive region of the partition U/D with respect to C , is the set of all elements of U that can be uniquely classified to blocks of the partition $U=D$, by means of C : Obviously

$$\gamma(C, D) = \sum_{X \in U/D} \frac{|\underline{C}(X)|}{|U|}.$$

If $k = 1$ we say that D depends totally on C , and if $k < 1$, we say that D depends partially (in a degree k) on C .

The coefficient k expresses the ratio of all elements of the universe, which can be properly classified to blocks of the partition U/D , employing attributes C and will be called the degree of the dependency. It can be

easily seen that if D depends totally on C then $IND(C) \subseteq IND(D)$. This means that the partition generated by C is finer than the partition generated by D. Let us notice that the concept of dependency discussed above corresponds to that considered in relational databases. Summing up: D is totally (partially) dependent on C, if employing C all (possibly some) elements of the universe U may be uniquely classified to blocks of the partition U/D (Pawlak 1998,2007).

1.3 Mixed Pixels

The basic building block of the study is the concept of mixed pixel, and what way the concept representing vital knowledge is elaborated using rough sets. A fundamental assumption that is commonly made in remote sensing is that each pixel in the image represents the area on the Earth's surface that contains a single class. This is often not the case, with the mixed pixels containing presence of more than one class (Fisher,1997).

Origin of the mixed pixels may be attributed to the following:

- Mixed caused by the presence of small, sub-pixel targets within the area it represents.
- Mixing as a result of the pixel straddling the boundary of discrete thematic classes.
- Mixing due to gradual transition observed between continuous thematic classes.
- Mixing problem due to the contribution of a target (black spot) outside the area represented by a pure but influenced by its point spread function. Also the pixel appears superficially be pure.

2 Methodology

Our objective is to find out a set of spectral bands which are causing experts to take differing view on the feature class of the terrain, of mixed pixels. We have taken a multi-spectral multi-resolution & multi-sensor image of Alwar area in Rajasthan. The area is selected since it contains a good variety of land use features like urban, water body, rocky, vegetation & open areas (Figure 01).

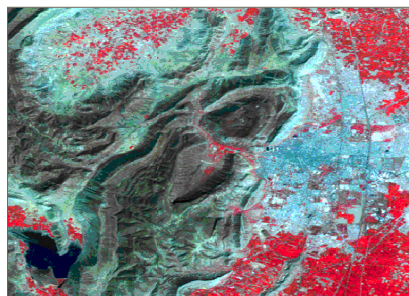
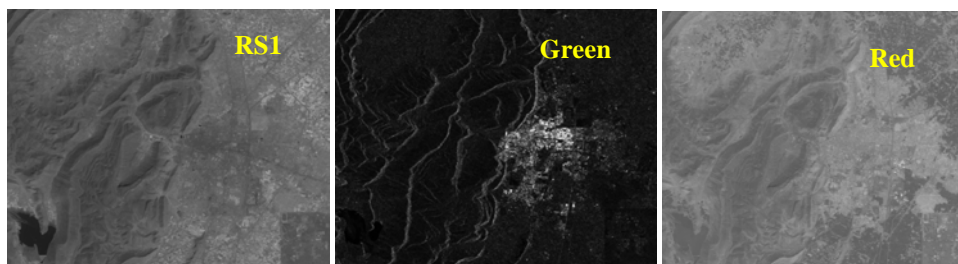


Figure 1: Landuse classification

The multi-spectral geo-referenced image-set consists of satellite images of dimension 472 X 546 pixels. The 4-spectral bands are in the visible bands namely: red, green, near- infrared (NIR) and middle infra-red (MIR) from LISS-III sensor of Indian Remote Sensing sat satellite Resourcesat-1. Also, two SAR images namely: low incidence S1 beam -20° - 27° (RS1) and other is High incidence S7 beam 45° - 49° (RS2) of the same area taken from Canadian Radarsat-1 satellite. The seventh image is digital elevation model (DEM) of the area. The ground resolution of these images from LISS-III, and Radarsat-1 images are 23.5m and 10m respectively. The DEM dataset is also generated from SAR interferometry using RS1 and RS2 and have 25 meter resolution.

We are having spectral signatures from seven bands namely Red, Green, NIR, MIR, RS1, RS2 (Figure 2) and DEM (Figure 3) of the mixed classes as training data set provided by the expert.



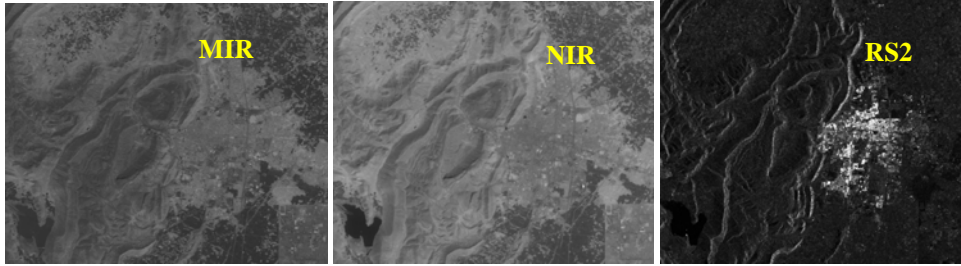


Figure 2: The 6-bands images.

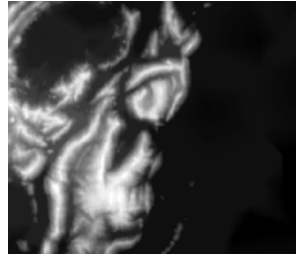


Figure 3: The DEM-band image

The training data set (249 pixels) can be represented in a tabular form similar to that of a relational database tables (Table 1). Rows of the table represent the training pixels and the DN values in the columns related to the 7-bands viz., Red, Green, NIR and MIR, RS1, RS2 & DEM. The table has, therefore, 7 attributes.

	RED	Green	NIR	MIR	RS1	RS2	DEM	Decision
11	76	66	100	104	2	0	30	W-V
12	52	56	65	80	2	0	30	W-V
13	62	56	100	107	22	43	30	W-V
14	54	48	93	107	19	38	30	W-V
15	70	62	115	126	36	33	30	W-V
16	68	56	95	111	40	38	30	W-V
17	52	53	98	115	24	27	30	W-V
18	66	59	111	122	25	27	30	W-V
19	64	57	107	126	16	34	30	W-V
20	68	59	109	111	9	56	34	W-V
21	44	41	95	93	18	18	65	W-V
22	13	22	64	107	21	15	50	W-V
23	25	30	73	87	13	24	54	W-V
24	25	30	73	87	13	12	54	W-V
25	27	30	74	84	10	18	54	W-V
26	27	30	74	84	13	18	54	W-V
27	66	56	135	122	10	19	42	R-V
28	62	56	131	115	21	30	42	R-V
29	62	56	131	115	23	43	57	R-V
30	56	56	133	126	12	32	48	R-V
31	60	54	133	120	13	24	48	R-V
32	60	54	133	120	13	21	60	R-V
33	68	57	138	118	20	32	48	R-V
34	46	48	117	102	21	24	53	R-V
35	42	45	115	102	18	27	53	R-V
36	48	45	117	104	13	25	43	R-V
37	52	51	126	96	18	26	43	R-V
38	50	46	124	107	15	26	39	R-V

Table 1: Spectral Signature in Digital Numbers

2.1 Problem Definition

We may redefine our problem as:

- The expert’s perception in assigning a class tag to a pixel is based on the kind of the images from space-borne sensor, the proximity of the pixel to a major class and extent of the domain knowledge of the expert(Fig.4).
- The representative mixed pixel (of class water body-vegetation) is composed of 7 bands DN Values.
- The dependency ‘k’ of the expert's decision viz water-vegetation mixed pixels on the conditional attributes in the form of 7-bands images is calculated using the positive region.

$$k = \gamma(C, D) = \frac{|POS_C(D)|}{|U|},$$

$$POS_C(D) = \bigcup_{X \in U/D} \underline{C}(X),$$

$$\gamma(C, D) = \sum_{X \in U/D} \frac{|\underline{C}(X)|}{|U|}.$$

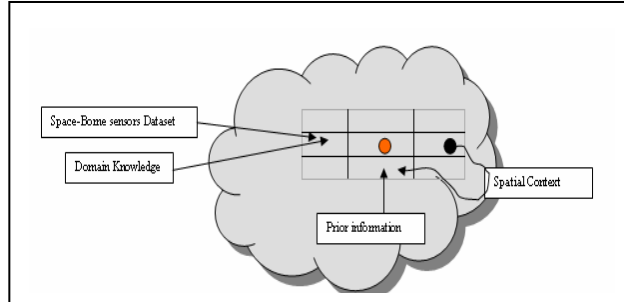


Figure 4: Expert's perception and deriving most appropriate label for a pixel.

3 Results and Analysis

We have calculated the lower approximations for every conflicting classes (Water body-Vegetation, Rocky area - Vegetation, and Urban –Vegetation) calculated the cardinality of positive region and then plotted the graphs to see which spectral bands are causing the conflict in experts opinions.

For the mixed pixels data, the graph is plotted in which the ordinate represents the dependency, k , and *abscissa* represents the band (water-vegetation, Figure 5; urban-vegetaion, Figure 6; rocky-vegetation, Figure 7).

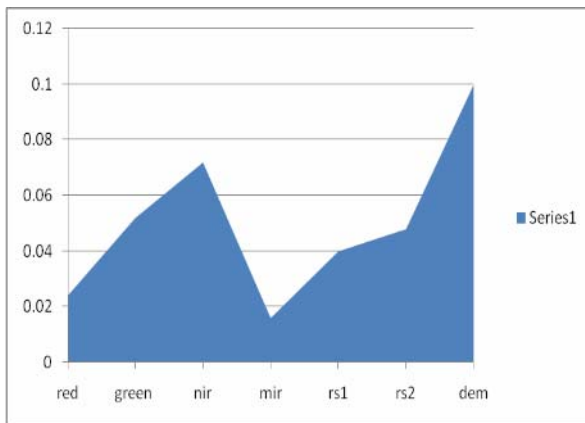


Figure 5: Water-vegetation

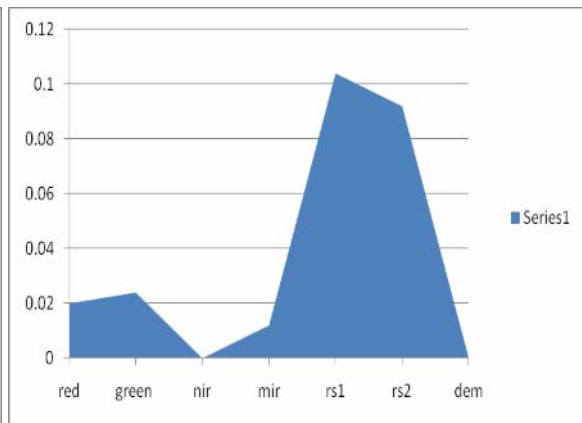


Figure 6: Urban-vegetation

As a representative case , refer to graph Fig.6:urban-vegetation, NIR band & RS1 band plays very crucial role. NIR band is good for delineating vegetation Landover whereas RS1 is good for identification of urban area. & these are the band responsible for creating class tag conflicts in the experts' opinion.

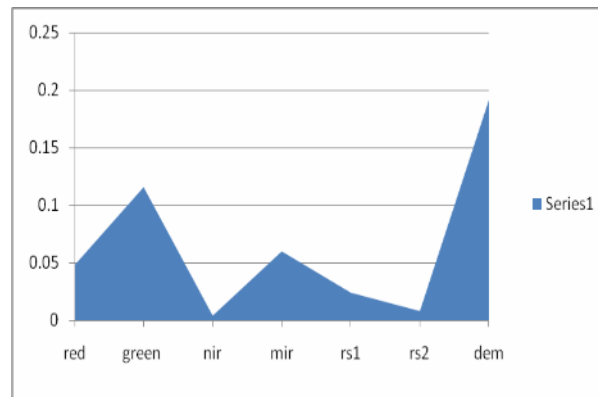


Figure 7: Rocky-vegetation

4 Conclusion

From the study being carried out, we derive the following inferences:

- The dependency of spectral bands based on the lower approximation space of the Rough Set theory facilitates in recognizing the cause of the conflict in assigning mixed pixels a definite class tag / label.
- Spectral bands have a definitive contribution towards the conflict arising in an expert's mind in the process of categorizing boundary regions.

References

- Fisher, P. "The Pixel: a Snare or a Delusion", International Journal of Remote Sensing, Vol.18: pp. 679-685, 1997.
- Pawlak, Z. (1998). Rough Set Elements. In Polkowski L. Skowron A. Rough Sets in Knowledge Discovery 1 Physica- Verlag, 10-30
- Ralph W.Kiefer & Thomes M. Lillesand: : "Principles of Remote Sensing", 2006.
- Zdzislaw Pawlak & Andrzej Skowron. Rudiments of rough sets. Inf. Sci. 177(1): 3-27 (2007).
- Zdzislaw Pawlak. Rough Set Theory and its Applications to Data Analysis. Cybernetics and Systems 29(7): 661-688 (1998).

Designing Airborne LiDAR Data Interoperability Standards for Multiple Organizations

– Canada's first experience in New Brunswick

P. W. Adda^{a,*}, D. Coleman^a, D. B. Finley^b

^aDepartment of Geodesy and Geomatics Engineering, University of New Brunswick, Head Hall, Rm. C-28, 15 Dineen Drive, Fredericton E3B 5A3, NB, Canada, - m9b35@unb.ca

^bService New Brunswick Building, E3B 5G4, Fredericton, NB, Canada, - david.finley@snb.ca

KEYWORDS:

LiDAR, Requirements, Multiple Uses, Interoperability, Common Standards

ABSTRACT:

Considering the huge cost of acquiring LiDAR data for government organizations, where possible, it makes sense to share both the data and the cost of acquisition among multiple organizations. However, firstly, differences in accuracy requirements, data formats and preferred project-specific deliverables requested by multiple organizations, impede the sharing of LiDAR products among multiple users. Secondly, no commonly-shared Quality Assurance specifications and methods have yet been adopted to verify the reliability of the delivered LiDAR products. Thirdly, most nations do not have organized documentation to ease access to the data and information relating to its quality. This paper reports a method used to investigate the strengths and gaps in some existing standards and conduct interviews to determine common user requirements among multiple provincial organizations, which led to the design of minimum LiDAR acquisition and processing specifications to support data interoperability standards.

1 Introduction

LiDAR (Light Detection And Ranging) data and its derived products, have become a major source for digital spatial data acquisition (Raber et al., 2007; Hodgson and Bresnahan, 2004; Sithole and Vosselman, 2003). This is as a result of the advancement of the technology over the past decade (Habib et al., 2005). LiDAR technology consists of the integration of INS (Inertial Navigation System) for orientation, LASER (remote sensing laser range finder and a scanner) for point cloud collection, and GPS (Global Positioning System), to determine the positioning aspects of the collected data. These three units are time-synchronized in a process called registration to achieve instantaneous (x,y,z) positioning and orientation of laser points reflected off features on the ground (Pfeifer and Briese, 2007).

LiDAR survey companies perform surveys with various specifications which produce deliverables that may meet particular project needs, but may not be suitable for others due to the different accuracies and formats employed. In New Brunswick, Canada, at least four departments have acquired LiDAR data, or are interested in using LiDAR for mapping and analysis. Since 2007, three different contractors have already been involved in the LiDAR surveys of selected corridors within the province. After the data has been captured, other departments may be interested in that data because of the period of capture or the area covered regardless of time, or both. However, the different standards make it difficult to integrate the point cloud and its processed products into the traditional datasets of the other departments. Also, the absence of metadata files explaining the data set and its associated specifications does not support data interoperability. In attempts to remedy these difficulties, in December 2009, Service New Brunswick (SNB) contracted the University of New Brunswick (UNB) to do a user-needs study among provincial organizations with interest in LiDAR data and related products and produce a minimum specifications document that can serve as a guide for LiDAR data acquisition and processing projects. The specifications were to allow for data operability across platforms and organizations. Through both research into existing standards and interviews with provincial users, minimum standards for LiDAR data acquisition and quality control were developed. A web enabled inventory was created to provide a quick overview of the extent of LiDAR data coverage in the province.

2 Objectives and Methodology

Recurring questions asked by participating stakeholders during the initial research process to determine user requirements all related to who owned LiDAR data and how it could be utilized. These questions fit into the set project objectives, which was to document metadata concerning existing and available LiDAR data coverage in the province and prepare prototype specifications for LiDAR data acquisition and processing.

To achieve this objective, a five member Project Steering Committee was created comprising of main users and stakeholders of airborne LiDAR data. From the deliberations of the bi-weekly meetings a decision was taken to undertake a Project Initiation and Preliminary Data Collection – involving meetings, interviews and collection of requirements for LiDAR data. The next stage was to deliver mid-project presentations and briefings by preparing and holding workshops with key DTM users and local suppliers to determine their priority characteristics for the specifications. The feedback was used to design specifications. This initial draft was presented to the stakeholders and upon feedback the final specification was forwarded to the lead client (SNB).

3 User needs research

The user needs research was to develop LiDAR specifications that satisfy the requirements of multiple organizations with various mandates as shown in Table 1.

Table 1: Participating stakeholders and their area of interest related to LiDAR

Organization/Department	Area of Interest related to LiDAR
Service New Brunswick (SNB)	Administration
Department of Public Safety (DPS)	Safety advice and emergency response
Department of Natural Resources (DNR)	Protect and manage natural resources
Department of Transportation (DOT)	Safe infrastructure and travel advice
Cities of Fredericton, Moncton and Saint John	Information about the city
Department of Environment (ENV)	Manage land use and waste
Department of Agriculture & Aquaculture (DAA)	Manage agriculture & aquaculture industries
Department of Wellness, Culture and Sports (WCS)	Citizen's health and well being
University of New Brunswick (UNB)	Research and development
NB Power	Crown Corporation – Services
JD Irving (JDI)	Private Business – Services
Leading Edge Geomatics (LEGEO) and CARIS	Private Business – Geomatics
DataQC	Private Business – Quality Control

User requirements of the ten government organizations were considered in the development of the specifications, since the other departments depended on these for their spatial data needs. The UNB project team held discussions with the selected public stakeholders regarding the application of LiDAR data to their mandate. During these meetings, it was realized that most of the departments required Digital Elevation Models (DEMs), contours and 3D extractions for infrastructure, tree heights and bare ground terrain models. The specific requirement of each organization was tabulated and analyzed to determine common accuracy and product needs.

The following tables attempt to summarize commonly used products, various data formats and their required accuracies for provincial organizations.

Table 2: Formats and accuracies (RMSE_z at 68% confidence level) of contours commonly used.

Department	.DGN	NTX	.DWG	x y z	.DXF	.SHP	KML	RMSE _z (m)
Cities								0.20
DPS								0.15
DNR								0.20
DOT								0.02
SNB								0.20
ENV								0.15
DAA								0.20
Total	3	8	6	5	3	8	4	

RMSE_z: Min 0.02 0.17 0.20

Key for Most Used



Table 3: Formats and accuracies of Mass points and TINs commonly used.

2. Mass points and TIN					
Department	ASCI		TIN	Confidence Level (%)	RMSE _z
	Xyzi				
Fredericton & Saint John				68	0.40
DPS and ENV				95	0.15
DNR				95	0.10
DOT				95	0.02
SNB, DAA				68	0.40
Moncton				95	0.30

Min 0.02 Mean 0.24 Max 0.40

At an average cost, the RMSE_z (Root Mean Square Error in vertical) of 15cm met 78% percent of the requirements of participating organizations. Horizontal accuracy of 50 cm was satisfactory. Formats desired were ESRI.shp, CARIS.ntx and ASCII.xyz (i). DNR and DOT requirements exceeded the set 15 cm RMSE_z (Figure 1).

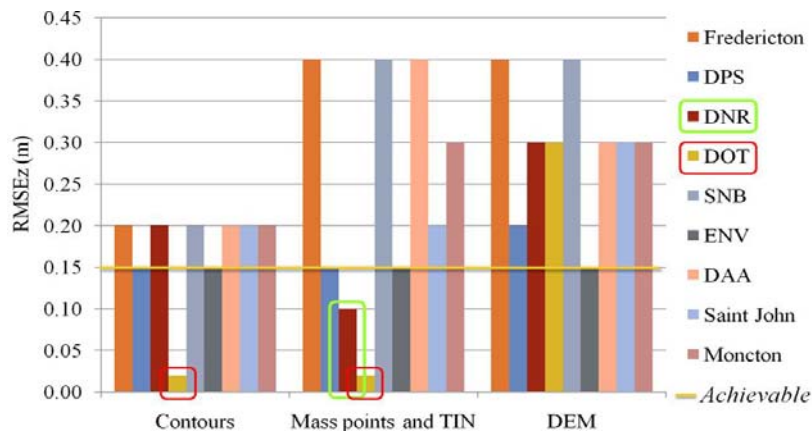


Figure 1: Summary of user product and accuracy needs

3.1 Common needs

Common requirements among participating provincial departments include the need for an inventory of available provincial LiDAR data and supporting documentation, in the form of metadata, the raw LiDAR point cloud, containing hydrographic features, vegetation and infrastructure (including buildings, roads and utilities),

filtered classified point cloud to show bare earth, vegetation, hydrographic features, utilities and infrastructure layers, intensity values, GPS and IMU data for post analysis, raster DEM of bare earth surface, generated features, including contours, lines, points and polygons and 3D models of buildings .

3.2 Differences

The Departments of Transportation and Natural Resources (Lands, Minerals & Petroleum Division) suggested their organizations had exceptionally high vertical accuracy requirements for all elevation data collected ($\pm 2\text{cm}$). This fell outside that required by other departments and by the current capability of LiDAR data technology. Since each department contain several user groups, each with their own applications and requirements, more research is necessary to determine the respective range(s) of requirements in each of these organizations. Alternative methods may be integrated with airborne LiDAR – for instance Aerial Photogrammetry to help meet their requirements. While some organizations are better positioned to afford the acquisition of LiDAR data, others are not. Also, there were questions regarding whether or not the LiDAR data acquired should be freely available to everybody. If data is to be freely available, who should pay for it? Who should be its custodian and under what conditions may the data be accessed?

3.3 Initial compilation of available provincial LiDAR data

The Departments of Public Safety and Environment, and the Cities of Saint John and Moncton have acquired LiDAR data. Table 4 summarizes the data as reported during consultation.

Table 4: Overview of available LiDAR data and products in New Brunswick

Department	Vendor	Area	Date	Accuracy	Format	Applications	Derived Products
DPS/MSP	LEGEO	Several ¹	2008+	$\pm 15\text{cm}$	Several ²	Impending	DEM ³
Saint John	COGS / GeoNET	Marsh Creek	Unknown	$\pm 20\text{cm}$	x, y, z	Research	Elevation data
Moncton	COGS / GeoNET	Turtle Creek	Unknown (2000 + -)	$\pm 30\text{cm}$	Several ²	Several ⁴	DEM, Road lines, Buildings
ENV	COGS	Coastal NB	2003	$\pm 30\text{cm}$	unknown	unknown	unknown

1. From the Mactaquac dam down to Grand Lake including St John River
 2. LAS, ASCII, xyzi, xyz, ESRI Feature Class
 3. DEM by LEGEO/CARIS. 3D flythrough to be produced for flood prone areas
 4. Analyzing effects of tree heights on City infrastructure, 3D modelling, DEM creation

4 Main Components of specified minimum requirements

The minimum specifications document present details of required minimum specifications to be used as a guide by provincial departments when contracting LiDAR contracts. A summary is as follows:

- 1. General responsibilities** – The vendor provides professional, technical and material supplies while the Client provides the quality control checks and information needed by the Vendor to understand, execute, and present the requested data and products.
- 2. Formats and Sizes for data and products** - Formats for raw LiDAR data shall be “.las”. Digital Elevation Models (DEMs) shall be in “CARIS” or “.shp” formats Contours derived from processing the LiDAR data shall be in ESRI .shp, CARIS or xyz formats. Mass Points shall be expressed as ASCII xyz or xyzi (FEMA, 2002).
- 3. Rights of Clients to Delivered Data and Products** – Shared rights for Clients who share cost.
- 4. Ground Control** – GPS ground control (base stations) shall be established to control the project and shall be located at regular intervals within the trajectory to achieve set accuracies.

5. **Datum and Reference systems** – LiDAR data can be collected in Spherical Coordinates (Latitude and Longitude), and GPS Ellipsoidal Heights but must be transformed to local datum and reference system.
6. **Horizontal and Vertical Accuracies** – Absolute vertical accuracies shall be $\pm 30\text{cm}$ while horizontal accuracies shall be $\pm 50\text{cm}$ at 95% confidence level relative to the provincial High Precision Network (HPN).
7. **Quality Assurance** – The Vendor, Client or a contracted 3rd party QA team shall perform checks on the processes involved in the survey process and products to ensure they meet or exceed the minimum standards.
8. **Metadata Documentation** – Shall conform to the North American Profile of ISO19115:2003. Naming shall conform to the NB DTDB 1998 naming convention namely, Individual Tile (1km x 1km window), Block (containing two or more Tiles) and Overview (containing two or more Blocks).
9. **Projects requiring higher accuracy standards** – It is acceptable for provincial departments to acquire data at a higher accuracy than the minimum requirement.

5 Conclusion

The set project objectives were achieved and all deliverables namely, the project report, recommended minimum specifications requirements and supporting Google *.kml* file to view the available provincial LiDAR data coverage generated. In practice and in future developments, the minimum specifications may not always meet project-specific requirements. Subsequently, additional recommendations were designed to cover projects requiring higher specifications.

A proposed new tiling system for naming LiDAR data files was developed based on a modification of the New Brunswick Digital Topographic Database (NB DTDB 98) naming convention. The new LiDAR data tiling system will ensure the easy referencing and understanding of data associated with LiDAR and its related products. The metadata proposed for inclusion as part of these specifications conforms to the adopted North American Profile of ISO 19115:2003 – Geographic Information – Metadata version 1.0.1 (NAP, 2007).

6 Future work

Issues relating to data custodianship, sharing and right of access were strongly debated at stakeholders meetings. Although defining such terms was beyond the scope of this project, these sensitive issues will be dealt with in the next phase of the project. A server is required for hosting LiDAR products to support the storage of original raw LiDAR point clouds and to provide services relating to how data and products can be accessed. A LiDAR workshop should be organized in future to equip and update staff with the relevant skills to utilize LiDAR products.

Conflicting LiDAR product standards and data integration difficulties call for the design of minimum international standards for LiDAR data acquisition and processing. The specifications can be built from harmonizing organizational specifications evolving from project-specific requirements. The next step will be merging these organizational standards to form regional, national and finally international standards

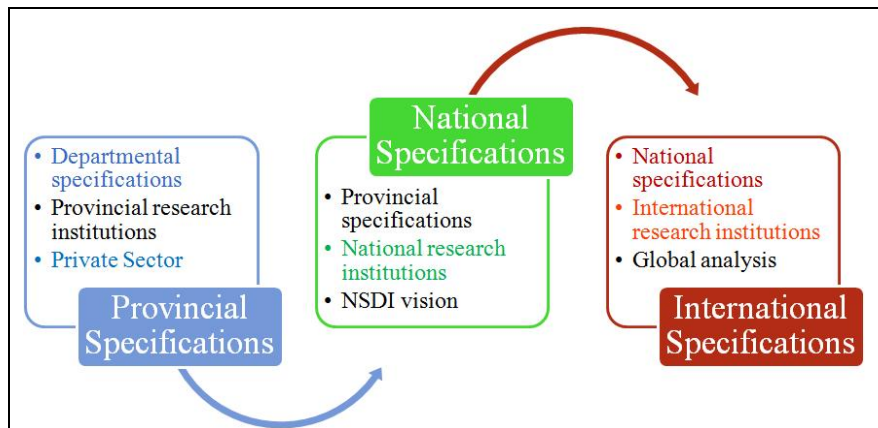


Figure 2: LiDAR data Specifications evolution towards international standards

7. References

- ASPRS (2004) Guidelines Vertical Accuracy Reporting for Lidar Data; Version 1.0 Drafted May 15, 2004; Released May 24, 2004; ASPRS Lidar Committee; Editor: Martin Flood. Last accessed online on March 22 from http://www.asprs.org/society/committees/lidar/Downloads/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf
- FEMA (2003) Guidelines and Specifications for Flood Hazard Mapping Partner, Appendix A: Guidance for Aerial Mapping and Surveying. Accessed from www.fema.gov/library/viewRecord.do?id=2206 on November, 2009
- FGDC (1998) Federal Geographic Data Committee. Geospatial Positioning Accuracy Standards Part 3: National standard for Spatial Data Accuracy. Appendix 3-A (normative): Accuracy Statistics
- Habib, A., Ghanma, M., Morgan, M. and Al-Ruzouq, R. (2005). Photogrammetric and LiDAR data registration using linear features. *Photogrammetric Engineering and Remote Sensing*, 71 (6), pp.699-707
- Hodgson, M. E. and Bresnahan, P. (2004). Accuracy of airborne lidar-derived elevation: Empirical assessment and error budget. *Photogrammetric Engineering and Remote Sensing*, 70 (3), pp.331-339
- NAP (2007) Metadata Working Group. North American Profile of ISO19115:2003 Geographic information – Metadata. Version 1.0.1. Accessed online on March 22 2010
- Pfeifer, N. and Briese, C.(2007) Geometrical aspects of airborne laser scanning and terrestrial laser scanning. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36(part 3/W52), pp 311-319
- Raber, G. T., Jensen, J. R., Hodgson, M. E., Tullis, J. A., Davis, B. A. and Berglind, J. (2007). Impact of LiDAR nominal postspacing on DEM accuracy and flood zone delineation. *Photogrammetric Engineering & Remote Sensing*, 73 (7), pp.793-804.
- Sithole, G. and Vosselman, G. (2003). Report: ISPRS Comparison of Filters. Department of geodesy, Faculty of Civil Engineering and Geosciences, Delft University of Technology. The Netherlands.
- USGS (2010) US Geological Survey national geospatial program base LiDAR Specification Version 13. Was last accessed from <http://lidar.cr.usgs.gov> on March 10, 2010

**Environmental Issues, Sustainable Development,
Millennium Development Goals**

GIS-Based Landuse/ Land Cover Change Characterization in the Humid Tropical Meenachil River Basin, Kerala, South India

Vincy Mary Varghese¹, Brilliant Rajan¹, A.P.Pradeepkumar^{2,*} and Riju Stephen³

¹School of Environmental Sciences, Mahatma Gandhi University, Kottayam 686 560, Kerala, India.

²Disaster Management Division, School of Environmental Sciences, Mahatma Gandhi University, Kottayam 686 560, Kerala, India - geo.pradeep@gmail.com

³Houston Department of Health and Human Services, Houston, Texas, United States

KEYWORDS:

Landuse/ landcover, LULCC, Change Detection, GIS, Meenachil, Kerala, India, Tropical

ABSTRACT:

Land is becoming a scarce commodity the world over due to immense agricultural and demographic pressure. Hence, information on landuse/land cover change (LULCC) and possibilities for their optimal use are essential for the selection, planning, implementation and monitoring of development schemes to meet the increasing demands for basic human needs and welfare while at the same time preserving biodiversity, especially in developing countries. The UN Millennium Development Goals (MDGs) have this key aim of alleviating poverty by half of its 2002 levels by 2015 and maintaining environmental sustainability. Many studies have shown that biodiversity protection enhances the latter and concomitantly reduces poverty. The Meenachil River originates from the Western Ghats, the mountain chain that straddles the eastern border of the state of Kerala in India. Its watershed extends from 9°25' to 9°55'N latitudes and 76°20' to 76°55'E longitudes. The satellite data of IRS-P6-LISS-III (acquired on February 9, 2004) and IRS-IB-LISS II path/row (25/62) of March 9, 1996 were used. Land mismanagement, landuse changes especially the conversion of paddy fields, under-utilization of potential crop lands, increasing population etc. are the major driving forces of LULCC in the Meenachil River Basin. This study is in tune with one of the UN Millennium Development Goals of striving to integrate sustainable development principles into country policies and reversing environment resource loss.

1 Introduction

‘With increasing global challenges, such as population growth, climate change, and overconsumption of ecosystem services, we need further integration of the poverty alleviation and biodiversity conservation agendas’ opined J.D.Sachs and co-authors in their seminal *Science* paper of 2009 (Sachs et al, 2009). UN Millennium Development Goals have been slow in being achieved and the present paper is an attempt at addressing some of the concerns which prompted the adoption of these goals by the UN.

Land is an area of the earth’s surface; human interference and interaction with this natural environment result in landuse. The equilibrium of nature is maintained by maintaining all types of land such as forestland, wetland, wasteland, cultivable land, etc. in a balanced way. According to Konecny (2003) ‘land cover describes the physical appearance of the earth’s surface, while landuse is a land right related category of economically using the land’. Landuse (three dimensional) is the expression of man’s management of ecosystems in order to satisfy some of his needs. 95% of landuse decisions are made at the local level (Kleppel, 1998), a change in landuse pattern involves a shift in landuse such as urbanization, reservoir damming, deforestation, irrigation, agricultural practices, etc. To understand the significance of these changes in landuse pattern, detailed knowledge about the landuse practices, time series landuse pattern changes and its effects on environmental processes and systems are needed. The knowledge gained through the categorization and case studies of landuse change will be crucial in developing regional and global models of landuse/land cover change (LULCC). Landuse analysis has become a pre-requisite for developmental activities of a region. The unsustainable use of land and land resources is considered as one of the key factors in degrading the environmental quality of life on earth. The growing demands on land have resulted in a crisis of land mismanagement. Land resources are the foundation for the socio-economic development at the national, regional and local levels in many of the developing countries such

* Corresponding author.

as India. To prepare for optimal landuse planning, it is necessary to have knowledge of the existing landuse patterns. The impact of landuse on the prevailing surface and subsurface hydrologic conditions is remarkably high. Within a basin, the dynamics of hydrologic processes are governed partially by the temporal and spatial characteristics of inputs and outputs and the landuse conditions (Shih, 1996). Often it is forests which are at risk in the process of LULCC (Munsi et al 2010)

1.1 Landuse in Kerala

The landuse of Kerala shows a unique zonation from the hill ranges (>2500 m) to the coast. The riverbanks and coastal stretches cover nearly 10% of the total area of the state and are potential locations for wetland paddy cultivation. The upper reaches of the rivers, particularly the Periyar plateau to the south and Wayanad plateau to the north, are extensively cultivated with plantation cash crops like tea, coffee and cardamom. Rubber, tapioca, cashew and pepper populate the midland zone or foothills. At the lower elevations of the river valleys, tapioca, coconut and other tree crops are interspersed with paddy cultivation. The lowland coastal area is dominated by coconut and partly by cashew and paddy. The diversity of tree crops is markedly high at the lowland – midland junction. The forest diversity is confined to the highlands mostly above the 100 m contour (>70 m above sea level is highland, 70 – 7 m is termed midland and < 7 m is considered as lowland in this paper). Converting the forest area for a developmental setting accommodates the plantations.

2 Objectives and Study Area

The objectives of the present study include preparation of landuse/land cover maps of the Meenachil River Basin using Survey of India toposheets and satellite imageries, analysis of the LULCC during the last few decades, understanding the influences of human interventions in the basin and the creation of comprehensive and effective mitigation strategies for land conservation in the study area using remote sensing and GIS. The synoptic view of the area allows better monitoring capability, especially when the coverage is repetitive, interval is short, and resolution of the image is high. Remote sensing provides data in several discrete bands, enabling creation of false colour composites (FCC) and the interpretation accuracy is thereby increased visually and digitally. It provides real-time and unbiased base line information (Gupta 2001). Techniques and methods of using satellite imageries as data sources have been developed and successfully applied for landuse classification and change detection in various environments including rural, urban, and urban fringe areas (Shepard, 1964; Robinove et al, 1981; Jensen and Toll, 1982; Fung, 1990).

The Meenachil River originates from the Western Ghats. Its watershed extends from 9°25' to 9°55' N latitudes and 76°20' to 76°55' E longitudes. The major tributaries of the Meenachil River are Kurishumalai, Thrikkovil, Chattappuzha, Kadapuzha, Kulathukadavu, Poonjar, Chittar, Lalm Thodu, Pannagan Thodu and Meenadam Aar. The Meenachil River Basin has 57 panchayats and 3 municipal towns within its ambit (KSLUB, 1996). The length of the river is 78 km and the drainage pattern is generally dendritic and the entire basin is situated in the district of Kottayam, Kerala state, southern India. The major environmental issues in the study area are frequent floods and droughts, high rate of sedimentation and debris slides, salinity intrusion into rivers and ground water aquifers, lack of flushing and concentration of pollutants, especially in the lower stretches of rivers, river bank erosion and land degradation.

3 Materials and Methods

Data from Survey of India (SOI) topographic maps (58 C/6, 58 C/9, 58 C/10, 58 C/11, 58 C/13, and 58 C/14) on a scale of 1:50,000, IRS-IB-LISS II data acquired on March 9, 1996 and IRS-P6-LISS-III data acquired on February 9, 2004 were used for landuse classification. The satellite digital data was rectified using Survey of India (SOI) topographic maps, a reconnaissance survey was carried out to collect the ground information. The GIS database generated from the topographic sheets was further updated with the latest changes in the watershed. The image elements were correlated with ground truthing and the interpretation key was developed. The tonal variation representing the different classes was marked on the hard copy image. Base maps including road, railway, settlement, village location and watershed boundary was extracted from the topographic sheets. AutoCAD Map 2000, Arc GIS 8.3, Erdas 8.6 and ArcMap were used in the preparation of the thematic maps.

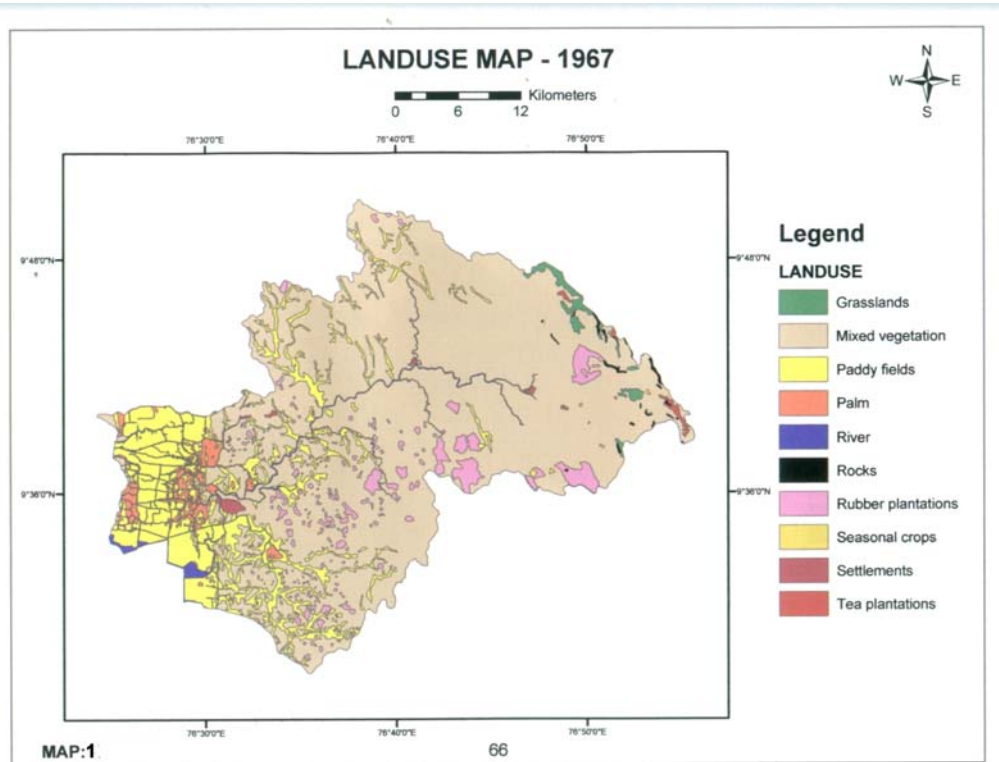


Figure 1: Landuse map of 1967.

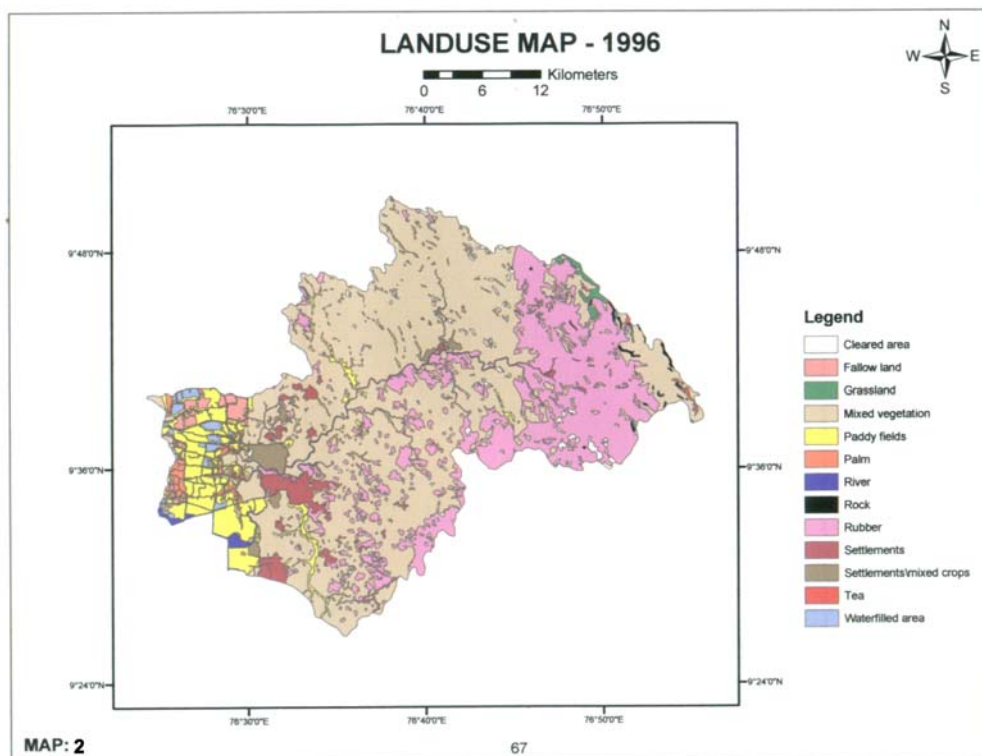


Figure 2: Landuse map of 1996.

4 Landuse/ Land Cover Mapping

Landuse categories such as agriculture, riverbed, water, urban, fallow, wasteland, etc. have been identified and mapped from the SOI topographic sheets. The landuse of 1967 was mapped, classified and calculated accurately from the toposheets and it was compared with those prepared from the satellite imageries (IRS 1B & P6 (LISS II & III)). IRS-P6-LISS-III data was used as the source for the landuse/land cover mapping. The interpretation key formulated during fieldwork has been used. The shadowed areas were put to corresponding classes on the basis of ground knowledge. From the rectified image, the area of interest i.e. the Meenachil River Basin, is extracted using the corresponding GIS layer. Supervised classification was carried out to prepare the landuse map. For this, initial identification of various landuse classes was carried out on the image. The registration and digitization of the watershed was done using AutoCad 2000 to create landuse coverage. Six landuse categories i.e. agriculture, urban, fallow, water, riverbed and wasteland were identified. Landuse/Land Cover map of 1967 was prepared from toposheets while those of 1996 and 2009 were prepared from the satellite imageries based on ground observations.

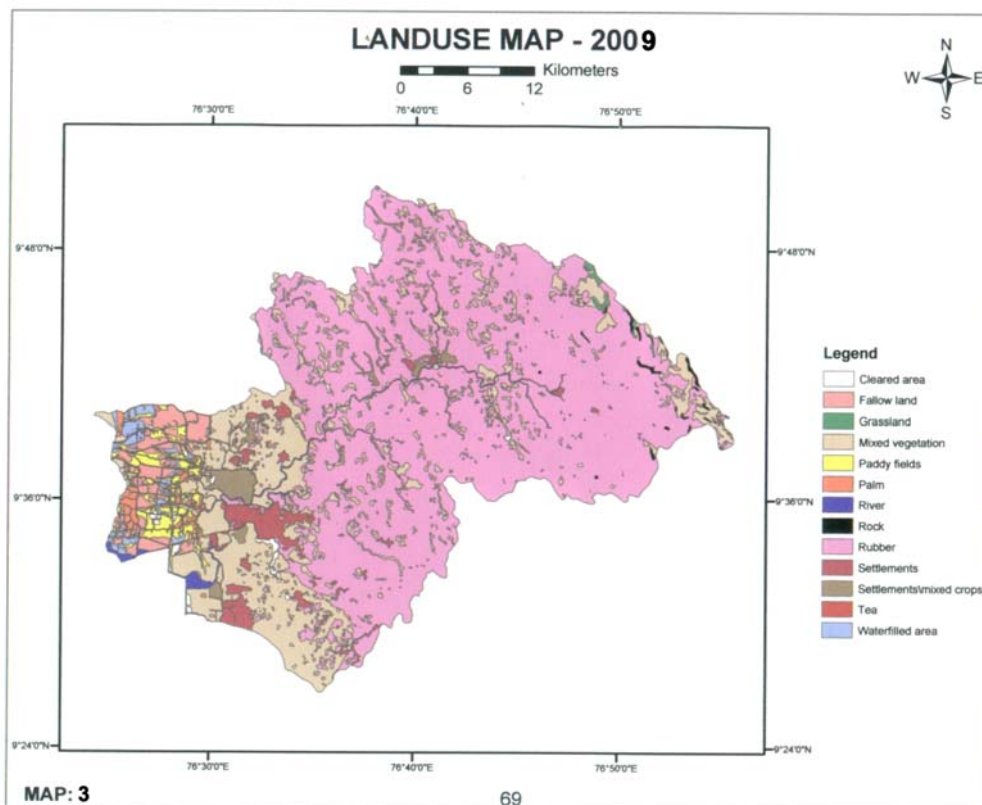


Figure 3: Landuse map of 2009.

5 Results and Discussion

The change detection maps were prepared by overlay analysis of the landuse maps of the years 1967 (Figure 1), 1996 (Figure 2) and 2009 (Figure 3). The landuse categories and its areal extent in various years have been tabulated in Table 1. The total study area is 120820.28 hectares. The most striking change in landuse has been the sharp increase in area under rubber plantation from 4.15% in 1967 (Figure 1) to 64.13% of the total area in 2009 (Figure 3), with a concomitant decrease in paddy cultivation from 15.71% of the area in 1967 to just 2.04% in 2009. There was an increase in built-up area from 439.9 hectares in 1967 to 3635.52 hectares in 2009, possibly due to increased affluence due to the good market prices for rubber produce. Considerable reduction was also noticed in grassland area, palm cultivation and tea plantations; which means reduction in total agricultural land. The mixed vegetation level decreased from 74.13% to 20.7% of the total area. There was little change in the rock area. The built-up land was 0.37% in 1967, 2.64% in 1996 and 3.01% in 2009. Urbanization process is in progress in the area. The geographical conditions of the district are suitable for rubber cultivation. The conversion of paddy fields, which are sinks for excess water, to other purposes increases the flood hazard.

Because of the reduction in water holding areas (wetlands), the district is also facing the threat of drought. Thus mismanagement of land, especially the conversion of paddy fields to building sites/other crops, under-utilization of potential croplands, coupled with growth in population; drive most of the landuse changes in the Meenachil River Basin.

Landuse of the study area							
	Category	1967		1996		2009	
		Area (ha.)	%	Area (ha.)	%	Area (ha.)	%
1	River	2394.89	1.98	2394.89	1.98	2394.89	1.98
2	Built up land	439.9	0.37	3193.86	2.64	3635.52	3.01
3	Paddy fields	18985.32	15.71	8851.74	7.33	2462.48	2.04
4	Rubber plantation	5044.95	4.18	28726.69	23.77	77484.29	64.13
5	Tea plantations	267.23	0.22	155.38	0.13	20.39	0.02
6	Mixed vegetation	89576.3	74.13	69071.08	57.17	25054.34	20.7
7	Seasonal crops	93.55	0.08				
8	Palm	2666.36	2.21	1546.33	1.28	1466.69	1.21
9	Waste land						
10	Rock	357.88	0.3	357.88	0.3	357.88	0.3
11	Grass land	993.91	0.82	608.48	0.5	282.47	0.23
12	Cleared area			2024.64	1.68	1127.01	0.93
13	Water filled area			1016.97	0.84	1293.08	1.07
14	Fallow land			1147.97	0.95	3173.82	2.63
15	Settlement with mixed crops			1724.37	1.43	2067.43	1.71
	Total	120820.28	100%	120820.28	100%	120820.28	100%
	ha. = hectare						

Table 1: Landuse/ land cover of the study area

6 Conclusion

Consideration of the existing socio-economic scenario is necessary before implementing any sort of landuse practices in the study area in the future. It is expected that the findings of the investigation will undoubtedly be of use to planners and local bodies to implement suitable landuse plans in the Meenachil River Basin, thereby achieving *eco-preservation* and also *enabling the restoration of degraded land units* to the maximum possible extent. The major finding of this study is that rubber plantations are smothering out other vegetation in the area, and the cost of this monoculture to the biodiversity of the area needs investigation. Recommendations on the basis of this study include (1) conversion and reclamation of paddy fields should be avoided, (2) sustainable utilization of land resources should be practiced (3) crop rotation should be implemented to improve soil fertility (4) crops should be selected based on soil properties. Local people must be made aware of the consequences of conversion of paddy fields, land and water management activities must be conducted only after detailed landuse planning, sand mining from rivers should be regulated, further expansion of rubber plantation at the expense of other crops and large-scale cultivation in hilly areas should be discouraged. UN Millennium Development Goals for the earth as a whole may be achieved by such small scale steps that cumulatively would enable us to live a better, environmentally sustainable life in cooperation with the flora and fauna of the earth.

References

- Fung, T., 1990. An assessment of TM imagery for land-cover change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4), pp. 681-84.
- Gupta, H. S., 2001. Remote sensing techniques for evaluating landuse/land cover: a case study. *Indian Forester*, July 2001, pp.755-761.
- Jensen, J. R., and Toll, D. L., 1982. Detecting residential land-use development at the urban fringe. *Photogrammetric Engineering & Remote Sensing*, 48(4), pp. 629-43.
- Kleppel, G., 1998. The landuse coastal ecosystem study (LU-CES). Presentation at the Environmental Policy Forum. Clemson, SC.

- Konecny, G., 2003. *Geoinformation: Remote sensing, Photogrammetry and Geographic Information Systems* 248p.
- KSLUB - Kerala State Landuse Board, 1996. *Water Atlas*. Government of Kerala, 1996.
- Munsi. M., Malaviya, S., Oinam, G. and Joshi, P. K. A., 2010. Landscape approach for quantifying land-use and land-cover change (1976–2006) in middle Himalaya. *Regional Environmental Change*, 10(2), pp.147.
- Robinson, C. J., Chavez, P. S., Jr., Gehring, D. and Holmgren, R., 1981. Arid land monitoring using Landsat albedo difference images. *Remote Sensing of Environment*, 11, pp. 133-56.
- Sacks, J.D. et al., 2009. Biodiversity conservation and millennium development goals. *Science*, 325(5947), pp.1502 – 03
- Shepard, J. R., 1964. A concept of change detection. In *Proceedings 30th Annual Meeting of the American Society of Photogrammetry*. Washington, D. C., 17-20 March, pp. 648-51.
- Shih, S. F., 1996. *Integration of Remote Sensing and GIS for Hydrologic Studies, Geographical Information System in Hydrology*. Kluwer Academic Publishers, Netherlands. **Acknowledgements**

VMV and BR wish to express their sincere thanks to late Dr. R. Satheesh, Reader, John C. Mathew and H. Vijith, School of Environmental Sciences, Mahatma Gandhi University, Kottayam for their guidance. They are grateful to both Dr.A.P.Thomas and the Director, School of Environmental Sciences, Mahatma Gandhi University, Kottayam for providing all the necessary institutional support during the period of study.

Forest Fire Risk Analysis and Management System Using Geoinformation Technology

Subin K. Jose¹, Santhosh Kumar², M.M. Jesnamol³, N.K. Sreeraj³, G.Madhu⁴ and Babu Ambat⁵

¹ School of Environmental Sciences, Mahatma Gandhi University, Kottayam 686 560, India
josesubin@gmail.com

² Centre for Earth Research and Environment Management, c/o Nansen Environmental Research Centre (India), Gopal Residency II Floor, Thottakkat Road, Cochin 682011, India

³ Disaster Management Division, School of Environmental Sciences, Mahatma Gandhi University, Kottayam 686 560, India

⁴ School of Engineering, Cochin University of Science and Technology, Kalamassery, Ernakulam, India

⁵ Centre for Environment and Development, Thozhuvankode, Trivandrum, India

KEYWORDS:

GIS, Remote sensing, Forest fire, Kerala, India, tropics, millennium development goals

ABSTRACT:

A forest fire hazard is influenced by ecological, human and climatic conditions. Forest fire directly causes biodiversity loss, forest degradation and climate change. It has direct ecological, economic and social impacts. Fire-prone area modelling is done to have the most effective system of fire management and at the same time to reduce its deleterious effect on ecosystems, communities and landscapes. Geospatial techniques are proving to be powerful tools to assess the forest fire risks and their management. The present study deals with identification of fire-prone areas and their management in Periyar tiger reserve, the mega biodiversity hot spot in Western Ghats mountainous region of Kerala in southern India, using GIS and remote sensing techniques. Large areas of forest are affected by fire in every year in this region. Two types of data are used in the study i.e. spatial and nonspatial data. Spatial data mainly includes remote sensing data, forest block, GPS field data, settlement boundaries, road/trek path network and most importantly the existing water bodies in that area. The nonspatial data pertains to meteorological data on temperature and rainfall. The parameters in consideration for the analysis are vegetation type, forest disturbance, vicinity to settlement, distance from road, slope, aspect, temperature, amount of litter and rainfall. All these parameters have direct/indirect influence on the occurrence of fire and were integrated using GIS. A multiparametric weighted index model has been adopted to derive the 'fire-risk' zone/fire-prone area map. The final output shows forest fire-risk area map of Periyar tiger reserve in four categories such as very high risk, high risk, moderate risk and low risk. Based on this final output, the study attempts to give insight into the use of remote sensing and GIS for sustainable fire management. For that, various spatial analysis have been carried out like identification of suitable sites for construction of check dams to ensure water availability in combating fires and for locating fire watch towers for better fire management. The present study effectively proved that a system for sustainable forest fire management could be established, and such a system would be in tune with UNESCO's Millenium Development Goal 7 of 'Ensuring environmental sustainability'.

1 Introduction

Western Ghats, a large biodiversity hotspot of India is located in the southern part of peninsular India. At present the entire area of Western Ghats is under severe threat due to increasing human intervention and over exploitation of natural resources. The increasing biotic pressure has led to the occurrence of frequent forest fires, which has resulted in fragmentation and degradation of the forests, many of these fragmented forest landscapes are in highly endangered status and show alarming signals of accelerated biodiversity loss. Statistics show that between 1991 and 2003, more than 25,000 hectares of forests in Kerala have been destroyed by fires (KFD, 2002). A recent study by the French Institute of Pondicherry shows that forest fire is a direct threat in 97% of forest divisions of Kerala. The causes of forest fire can be classified into three main categories (1) Natural, (2) Intentional/deliberate due to man, and (3) Unintentional/accidental due to man (Pandey and Barik, 2006). A

forest fire event is influenced by climatic conditions and is aided by the accumulation of fuel on forest floor (Subrat and Rikhari, 1997).

The prominent factors leading to fires are: 1. Vegetation type and density - dry and dense vegetation is obviously more susceptible to fire than the moist and sparse one; 2. Climatic factors - the climatic regime determines the vegetation in a region and hence plays a dominant role in ascertaining the fire-prone sites. Drier the climate, more prone would be the site (Flannigan and Harrington, 1988, Robinson, 1991). 3. Physiographic factors - viz., altitude, aspect and topography influence micro-climatic conditions and therefore, indirectly affect the vegetation. Aspect plays a vital role in the spreading of the fire. Southern and south-western slopes, exposed to direct rays of sun are more prone to catch fire rather than northern and north-eastern slope aspects. In the case of topography, fire travels most rapidly up-slopes and least rapidly down slopes (Rothermel and Richard, 1972). 4. Distance from roads – human and vehicular movements and such activities on the road provide enough scope for accidental/man-made fire. Near the roads, more would be the chance of fire occurrence. 5. Vicinity to settlements – the area near to the habitats/settlements is more prone to fire because the habitats/cultural practices of the inhabitants can also lead to incidental/accidental fire (Butler et al., 1991). Internet-based forest fire management system is used for the forest fire monitoring and prevention (Burke et al., 1997).

2 Materials and Methods

Since the study adopted a GIS-based methodology, both spatial and attribute data of various thematic layers, available in the form of maps and published reports were collected from various sources. Primary data were collected using various techniques such as field visits, discussion with locals to fill any data gaps.

Spatial data were collected from Survey of India (SOI) topographic map sheets of 1:50,000 scale. The SOI maps became the source for a number of basic thematic layers like sanctuary boundary, contour, drainage, settlements, roads/trek paths and water body (Figure 1).

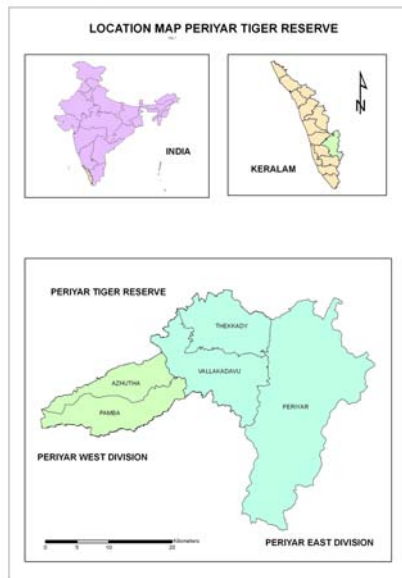


Figure 1: Location map

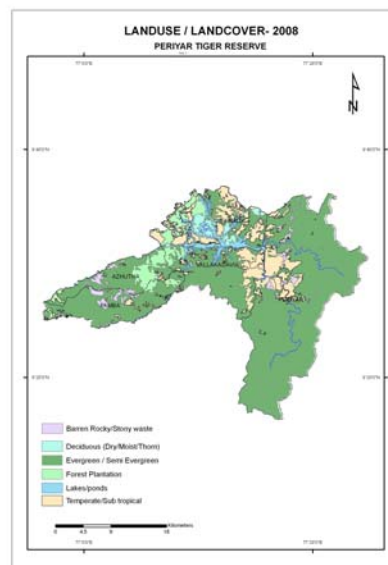
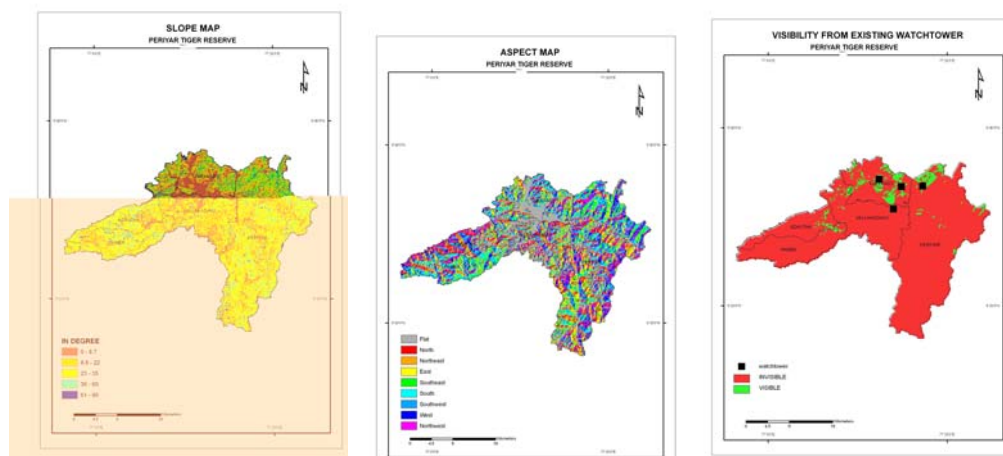


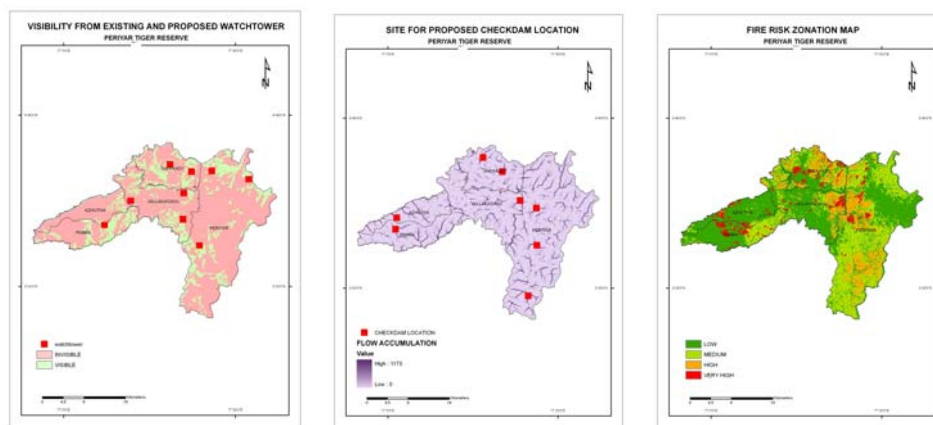
Figure 2: Landuse/landcover map of 2008

IRS-1D LISS-III digital image acquired on 12 April 2008 was used to prepare the landcover map (Fig. 2). Visual interpretation of the standard False Colour Composite (FCC) with 4-3-2 RGB combination as well as digital classification of the image using ERDAS 8.6 image processing software was used for the preparation of the landcover map (Kennedy, 1994; Elvidge et al., 1997). Field level data collection and mapping was carried out using toposheet as the base map (Song and Fan, 2001). The real world coordinates were collected using Magellan Spot Track Map GPS receiver. A hard copy of the vegetation map prepared from the satellite image was taken to the field and the vegetation polygons in the map were verified. Different reasons for the occurrence of fire in Periyar tiger reserve were collected by formal and informal discussions with locals, tribal people and forest officers. Details regarding trek path and fire lines inside the sanctuary were collected during field visit. The various data collected were processed and put into ArcGIS.

Different factors that contribute to forest fire were identified and evaluated on the basis of historic data and ground truth. For the delineation of forest fire-prone areas, thematic layers were generated. Forest disturbance zone of the study area was identified. The occurrence of fire in the disturbed forest is high compared to the undisturbed area (Chand and Madhavilatha, 2004). Meteorological data like rainfall, humidity, temperature were analyzed and these parameters were found to have direct correlation to the forest fire. Increase in temperature increases the chance of fire occurrences and whereas rainfall and humidity have the opposite relation with fire (Valdiya, 2006).



Figures. 3,4,5: Slope map, aspect map and visibility from watch tower maps



Figures. 6,7,8: Visibility from existing and proposed watch tower, checkdam location and fire risk zonation map

The ranks of the classes of each theme were assigned in a scale of 1-10 by reclassifying themes in the Spatial Analyst module of ArcGIS software in such a way that higher ranking is given to the class which has high positive relation with fire proneness than the class which shows lesser positive relation with it (Jo et al., 2002). The areas closer to the settlement or trek paths are high fire risk areas and therefore are given higher rankings. The rank decreased with the increase in distance from the settlement or trek paths. Similarly slope (Fig.3) and aspect (Fig.4) of the terrain also have indirect influence on fire risk as they control the vegetation density and composition. Steeper slopes lead to less fuel moisture and less air humidity (Pandey and Barik, 2006). The vegetation in such slopes is mainly dry and catches fire easily (Lawrence et al., 2006). Moreover, once the fire starts, it is likely to spread faster up-slope than down-slope and along steeper slope (70–90) than gentle ones (0–10). The slope from 70–90° has high fire risk when compared to flatter terrains. Similarly, the west and southwest facing aspects get more sunshine than the eastern and northeastern aspects in summer and tend to make the vegetation drier, which increases the vulnerability to fire (Hochkirch and Adorf, 2007). In the vegetation type, higher ranking is given to deciduous and temperate/subtropical forest as the fuel content of these vegetation is higher and they become dry in summer, igniting easily. The visibility from the existing watchtower was analysed (Fig. 5). The line of sight does not cover the entire or the more fire vulnerable areas in the sanctuary. For the better fire fighting strategy new fire watch towers are needed.

The identification of suitable sites for new watch towers was carried out using visibility analysis by giving points in such a way that maximum terrain area is visible from each points which is near to fire prone sites. Suitable sites for the construction of new watch towers (Fig.6) were located using the digital elevation model (DEM) generated from the contours and spot heights of the area derived from the SOI toposheet.

A suitable site for the construction of check dam, that can help prevent the spread of forest fire is identified with the help of hydrological analysis tools in GIS. The site for check dam (Fig.7) is identified by locating the sites with high flow accumulation and sites close to the fire-prone area.

3 Results and Discussion

The fire-prone/fire-risk zonation map shows that the study area is divided into four zones based on fire proneness, viz. very high risk, high risk, moderate risk, low risk areas / zones (Fig.8). The areas near to the settlements are delineated as very high fire risk zones by considering a combination of factors such as the high human disturbance, favorable vegetation types (mostly deciduous), steep slopes and southwest and western aspects. In moderate risk zone, the human disturbance is minimal and the vegetation type is semi-evergreen. The slope varies from 15–30° and the aspect is mainly east and southeast. Areas that are away from the settlements and trek paths, where the vegetation is mainly evergreen and semi-evergreen, in the east and northeast aspects and with gentle slopes come under low risk zone.

This fire-prone area map can serve as the base information for developing a fire management system for the sanctuary as suggested by Close and Kelly (1993). During fire-occurring months, more fire watchers need be appointed in the fire-prone areas to control the fire (Ani 2002). In the Periyar tiger reserve, the visibility from existing watch towers was analysed (Fig.6) and it was found to be insufficient to cover the whole sanctuary area. Optimal sites for six new watch towers were hence identified based on the visibility analysis. Round-the-clock fire watching in identified very high risk fire-prone area will help to reduce the risk of fire spread (Klaver 2004). Construction of check dam near to fire-prone areas will help to prevent the spread of fire (Ciesla 1993). The eight new sites identified for the construction of check dams across the streams/rivers near the fire-prone areas will not only store water but also retain the moisture content in the area. The impounded water can also be used to extinguish fire.

4 Conclusion

The result of the study shows that GIS could be successfully employed in identification of fire-prone areas and their management in the Periyar tiger reserve. GIS analysis has taken into consideration a wide range of suitability parameters in identifying fire-prone area, suitable sites for the construction of check dams and construction of fire watch towers. The landuse map prepared from the remote sensing data has shown that the area constitutes deciduous forest, evergreen/semi evergreen, forest plantation, temperate/subtropical forest, barren rocky area and lake. Temperate/subtropical forests are more vulnerable to fire because it provides dry fuel for the outbreak of forest fire. The current data acquisition with the help of remote sensing technology plays a key role in the identification of fire risk area. The early identification of fire risk area will increase the participation of local people in the prevention of fire outbreak. The fire-prone area map and management plan can be used by the forest officials to develop and improve the forest fire fighting strategies within the sanctuary. Implementation of GIS, GPS and remote sensing techniques and well-trained officials are needed for efficient fire management and this will reduce the biodiversity loss in terms of flora and fauna of the reserve. The GIS-based methodological frame work developed as part of this study can be effectively used in similar situations elsewhere for identification of fire-prone areas and their management. The prevention and early identification forest fire with the help of geoinformation technology is very useful in the world-wide move towards the control of global warming and climate change and is in tune with the UNESCO's Millenium Development Goal 7 of 'Ensuring environmental sustainability'.

References

- Ani, J.R., 2002. How to check a forest fire. *Indian Forester*. 126 (7). pp. 766-771.
- Burke, T.E., Cameron, P. J., Lime, S. D. and Lane S., 1997. An Internet-based forest fire information system. *Unasylva*. 48 (2), pp.32 - 38.

- Butler, D. R., Stephen, J. W. and George, P. M., 1991. GIS Applications to the indirect effects of forest fires in Mountainous Terrain. Fire and the environment: Ecological and cultural perspectives. *Proceedings of an International Symposium*, pp.212 - 223.
- Chand, K.T.R. and Madhavalatha, K., 2004. Forest fire monitoring using ENVISAT-AATSR Data *Journal of the Indian Society of remote sensing*, 32(4), pp.317-322.
- Ciesla, W. M., 1993. Remote Sensing, GIS and Wild land Fire Management. A global Perspective. *Proceedings of the International Workshop on Satellite Technology and GIS for Mediterranean Forest Mapping and Fire Management, Thessalonik*, pp. 21 - 35.
- Close, P and Kelly, R., 1993. GIS Applications in wild land/urban interface fire planning: The Missoula County (Montana) Project. *Proceedings: symposium on fire in wilderness and park management. USDA Forest service*, pp.180 - 186.
- Elvidge, C.D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R. and Davis, C. W., 1997. Relation between satellite observed visible near infrared emissions: population, economic activity and electric power consumption. *International Journal of Remote Sensing*, 18(6), pp.1373 - 1379.
- Flannigan, M. D. and Harrington, J. B., 1988. A study of the relation of meteorological variables to monthly provincial area burned by wild fire in Canada 1953 – 80, *Applied Meteorology*, 27, pp. 441 - 452.
- Hochkirch, A. and Adorf, F., 2007. Effects of prescribed burning and wildfires on Orthoptera in Central European peat bogs. *Environmental Conservation*, 34(3), pp.225-235.
- Jo, M., Le, M., Lee, S., Jo, Y and Baek, S., 2002. The development of forest fire forecasting system using Internet GIS and satellite remote sensing. <http://www.gisdevelopment.net> (accessed 27 June 2010)
- Kennedy, P.J., 1994. An improved approach to fire monitoring in West Africa using AVHRR data. *Int. J. Remote Sensing*, 15(11), pp. 2235-2255.
- KFD., 2002. *Neyyar Wild Life Sanctuary Management Plan 2002-2012*, Kerala Forest Department, Trivandrum, pp.1-78.
- Klaver, R.W., 2004. Global forest fire watch: wildfire potential, detection, monitoring and assessment <http://www.rss.dola.wa.gov.au> (accessed 27 June 2009)
- Lawrence, A., Paudel, K., Barner, R. and Malla, Y., 2006. Adaptive value of participatory biodiversity monitoring in community forestry, Nepal. *Environmental Conservation*, 33(4), pp.325-334.
- Pandey, H.N and Barik, S.K., 2006. *Ecology, Diversity and Conservation of Plants and Ecosystems in India*. Regency Publications, New Delhi.
- Robinson, J.M., 1991. Fire from space – global fire evaluation using infrared remote sensing. *Int. Jnl of Remote Sensing*, 12(1), pp. 3-24.
- Rothermel, P and Richard, C., 1972. A mathematical model for predicting fire spread in wild land fires. *USDA Forest Service Research Paper INT - 115*, Ogden, Utah, USA.
- Song, W. and Fan, W., 2001. Self-organized criticality of forest fires in China. *Chinese Science Bulletin*. 46(13).
- Subrat, S. and Rikhari, H.C., 1997. Forest fire in the central Himalaya: climate and recovery of trees. *Int Jnl. Biometeorol* . 40, pp.63–70.
- Valdiya, K.S., 2006. *Coping with Natural Hazards: Indian Context*, Longman, Hyderabad. pp.250-256.

Forest Disturbance Analysis Using Geoinformatics in Pamba Range of Periyar Tiger Reserve, Kerala, India

Abin Varghese^{1,*}, John C. M.², Punnen Kurian³ and Thomas A. P⁴

¹ Tropical Institute of Ecological Sciences, P.B. No.24, Kottayam, India, - abin.udumbanattu@gmail.com

² Dr. R. Satheesh Centre for Remote Sensing and GIS, School of Environmental Sciences, Mahatma Gandhi University, Kottayam, India

³ St. Mary's College, Manarcadu, Kottayam, India

⁴ Advanced Centre for Environmental Studies and Sustainable Development, Mahatma Gandhi University, Kottayam 686 560, India

KEYWORDS:

Sabarimala pilgrimage, PTR, NDVI, Landscape analysis, Disturbance map

ABSTRACT:

Periyar Tiger Reserve (PTR), the largest protected area in the state of Kerala, in India with an area of 925 sq.km. is under severe anthropogenic pressure due to the annual pilgrimage to the Sastha temple (Ayyappa temple), located at Sabarimala, in the south-western Pamba range of the reserve. The present study explores the extent of forest degradation caused by the pilgrimage activity using the tools of geoinformatics. The land use/ land cover change analysis over the years, observed drastic conversion of dense forest area into other land use like permanent and temporary constructions, open forest, grassland and barren area. The study revealed that Sabarimala pilgrimage and related activities have caused significant loss and changes in the floristic pattern and cover density of the area, but it is confined to areas with direct human interactions only. The NDVI analysis showed the loss in vegetation vigour of the study area. The fragmentation (31.12%), patchiness (34.37%) and porosity (61%) measures showed high values indicating loss of forest cover, density and diversity in the trekking path, temple and Pamba Triveni areas. The Disturbance Index map showed heavy disturbance along the trekking path from Pamba Triveni to Sabarimala and the temple premises and it is evidently due to the pilgrimage activity. If this trend prevails the damage to the remaining forests would be irreparable. Controlled pilgrimage along with proper conservation programmes is recommended.

1 Introduction

Indiscriminate anthropological activities result in devastating impacts on forests, especially in hotspots of biodiversity like the Western Ghats (Roberts *et al.*, 1998). Various levels of applications of geoinformatics on biodiversity conservation and management have been reported from all over the world (Riitters *et al.*, 1997). Periyar Tiger Reserve (PTR) is the largest protected area in the state of Kerala with an area of 925 km². Sabarimala Sastha Temple, a famous Hindu pilgrimage centre, is situated in the deep dense forests in the Pamba range in the south west portion of PTR, at an elevation of 461 m above mean sea level. It records an inflow of more than 500,000 pilgrims every year (KFD, 1999). The Pamba range of PTR was composed of forest types which included tropical evergreen, semi-evergreen, moist deciduous and grassland and the area supports rich biodiversity, amongst which are a large number of endemics (KFD, 2007). The temple and the major trekking paths come under this range. There are only few studies on the pilgrimage activity in relation to forest disturbance in PTR. The present study was conducted in this background to determine the extent of forest disturbance in Sabarimala and Karimala sections of the Pamba range of PTR using geoinformatics tools.

2 Methodology

The impact of Sabarimala pilgrimage on forest cover was estimated through remote sensing data and GIS analysis coupled with ground truthing. The study area covers 28 km² of reserve forest area coming under Pamba administrative range, covering Karimala and Sabarimala sections (77°10'- 77°50'E and 9°24'- 9°28'N). It is

* Corresponding author

situated within the Survey of India toposheet 58G/3 (1:50000 scale) of 1967. This toposheet is used as the base map to create various thematic layers of the surveyed area. For different raster based analysis different thematic layers developed from satellite image IRS P6 LISS III (2004) were used. The vegetation vigour analysis of the area was done through Normalized Difference Vegetation Index (NDVI) analysis with the aid of ERDAS Imagine 8.5 version. Landscape analysis, including derivation of parameters such as porosity, fragmentation and patchiness was also performed. Along with these parameters the road distance and pilgrimage disturbance analysis were done. Importance Value Index (IVI) of vegetation obtained through field study was also taken as a layer for disturbance analysis of the area.

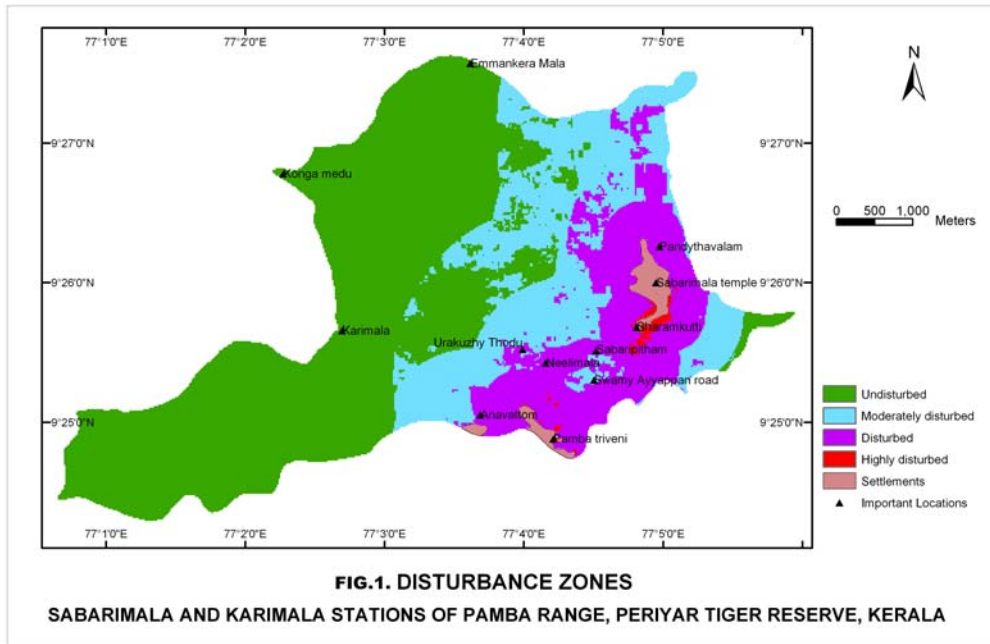


Figure.1: Disturbance zones around Sabarimala and Karimala stations, Periyar Tiger Reserve, Kerala

Raster analysis was done using Raster Calculator by which the ranks and weightages of each individual thematic layer is taken into consideration in determining the forest disturbance zones (Fig.1). The forest disturbance was computed by adopting a linear combination of the defined parameters on the basis of probabilistic weightage. The weightages were based on the importance of each layer in degrading the forest. Pilgrimage disturbance includes the disturbances caused due to the pilgrims’ rituals in the forest area, permanent and temporary constructions and related commercial activities. The IVI raised from quadrat study in the field was also taken as a raster layer by performing Inverse Distance Weighted (IDW) analysis and classified by giving ranks. The thematic layers were subjected to raster calculation in the Spatial Analyst extension of ArcGIS 8.3 using the formula:

$$\text{Disturbance Index (DI)} = R \text{ Land use/ land cover} \times Wt_1 + R \text{ NDVI} \times Wt_2 + R \text{ Pilgrimage disturbance} \times Wt_3 + R \text{ Road disturbance} \times Wt_4 + R \text{ IVI classified} \times Wt_5 + R \text{ Porosity} \times Wt_6 + R \text{ Fragmentation} \times Wt_7 + R \text{ Patchiness} \times Wt_8$$

where R = Reclassified and Ranked .

3 Results and Discussion

The land use/ land cover analysis of the study area using remote sensing data revealed that significant extent of the forest area was converted to non-forest land uses during the last few decades. The forest area decreased from 98.58% in 1967 to 54.43% in 2004 due to the increasing pilgrimage activity in the area (Table.1). The forest area has been converted to other land uses like construction of temporary and permanent buildings, open forests with cleared undergrowth to provide resting place for pilgrims, grasslands and barren area. The total removal of trees led to the increase in the barren area and grasslands. The increasing number of pilgrims in the recent years (Sathyapalan, 2002) shows a direct correlation with decreasing natural forest cover.

Land Use Type	AREA (km ²)	
	1967	2004
Forest	26.66	14.72
Open forest	0	2.01
Grassland	0	9.24
Barren area	0.24	0.64
Settlements	0.15	0.44
Total	27.045	27.045

Table.1. Land use/land cover pattern of 2004 in Pamba range (Sabarimala- Karimala stations) in PTR, Kerala

SI No:	Landscape Indices	Level of Disturbance	Area (km ²)	Area (%)
1	Vegetation vigour (NDVI)	Very low	0.44	1.57
		Low	2.89	10.32
		High	16.89	60.32
		Moderately high	7.08	25.29
2	Fragmentation	Low	18.63	68.89
		Moderately high	7.7	28.47
		High	0.68	2.51
		Very high	0.04	0.14
3	Patchiness	Low	17.76	65.67
		Moderately high	8.2	30.32
		High	1.01	3.72
		Very high	0.09	0.33
4	Porosity	Low	10.55	39.01
		Moderately high	16.1	59.53
		High	0.39	1.43
		Very high	0.01	0.04

Table.2: NDVI and Landscape analysis of 2004 in Pamba range (Sabarimala- Karimala stations) in PTR, Kerala

The landscape analysis of the study area using indices like fragmentation, patchiness and porosity, the degree of forest degradation and loss in vegetation vigour based on Normalized Difference Vegetation Index (NDVI) of the year 2004 is given in Table 2. The Sabarimala and Karimala stations showed considerably low vegetation vigour with respect to the erstwhile evergreen patches of the area. The low vegetation vigour denotes extensive deforestation and is mostly confined to Sabarimala temple and Pamba Triveni region, which shows a direct correlation with pilgrimage activity. Anthropogenic activities in forest areas such as massive pilgrimage is reported as a principal cause of forest degradation (Chauhan and Nayak, 2005; Joshi *et al.*, 2006). It is found that

SI No:	Level of disturbance	Area (km ²)	Area (%)
1	Undisturbed	15.09	55.80
2	Moderately disturbed	6.5	24.03
3	Disturbed	4.83	17.86
4	Highly disturbed	0.12	0.44
5	Settlements	0.46	1.70

Table.3: Disturbance zone analysis of 2004 in Pamba range (Sabarimala- Karimala stations) in PTR, Kerala

8.71 sq. km. (31.11 %) of the study area was fragmented and the effects are pronounced along the trek path to temple area and around the temple shrine, which is evidently due to the pilgrimage-related activities. Forest fragmentation can be considered as an index of human disturbances (Buckely, 2000; Krishnasagar, 2003). 9.29 sq. km (34.28%) of the study area has patchiness at various levels (Table 3). As patchiness is a measure of density of patches of all types in a unit area (DOS and DOB, 2004), higher more patchiness in an area indicates lesser floral density (Mathew *et al.*, 1999). The current data also is in agreement with this notion and considerable loss in density has been noted. Similar observations on forest patchiness are made in the Western Ghats region of Tamil Nadu (Araújo and Miguel, 2002).

We found 16.5 sq. km (61.01%) of the area facing porosity at various levels, which indicates higher fragmentation and habitat change. The higher porosity value than the fragmentation and patchiness indices indicate the higher interaction among landscape elements, heterogeneous nature and highly fragmented habitat (Turner *et al.*, 2003; Singh, 2004).

The Disturbance Index map showed heavy disturbance at trekking path from Pamba Triveni to Sabarimala and the temple premises (Fig.1) and 46.13% of the forest was found disturbed (Table 3). The disturbance observed in the study area is evidently due to the pilgrimage activity and if the trend prevails the damage to the remaining forests would be irreparable. Stringent measures to curb the uncontrolled pilgrimage activity along with the implementation of proper management and conservation programmes are recommended.

4 Conclusions

The current pilgrim process in the Pamba range of PTR has caused significant disturbance that has led to the loss of almost half of the forest area. Conservation strategies for this region should include limiting pilgrimage activities to the disturbed area through controlled pilgrim tourism programmes. The moderately disturbed and disturbed zones need more attention for the conservation of biodiversity. Disturbed and highly disturbed zones in the study area should have special management programme, and measures are to be taken to improve the vegetation cover. The forest disturbance map of the area can be used for the identification of the areas for conservation and management.

REFERENCES

- Araújo, A. and Miguel, B 2002. Biodiversity hotspots and zones of ecological transition. *Conservation biology*. 16(6): 1662–1663.
- Buckley, L.B.2000. The Dynamics of Temperate Forest Fragmentation: an Examination of Susceptibility to Woody Invasive Species. Unpublished bachelors' Degree of Arts with Honors in Biology.
- Chauhan, H.B and Nayak, S. 2005. Land use/ Land cover changes near Hazira region Gujarat using remote sensing satellite data. *Journal of Indian Society of Remote Sensing*. 33(3), 411- 20.
- DOS (Department of Space) and Department of Biotechnology Government (DOB) of India. 2002. Biodiversity Characterization at Landscape Level in Western Ghats India, using Satellite Remote Sensing and Geographic Information System.
- Joshi, P.K., Kumar, M., Midha, N., Vijayanand, M and Paliwal, A. 2006. Assessing areas deforested by coal mining activities through satellite remote sensing images and GIS in parts of Kobra, Chattisgarh. *Journal of Indian Society of Remote sensing*.34(4): 415- 21.
- KFD (Kerala Forest Department). 1999. Report on the impact of Sabarimala pilgrimage and recommendations. A report prepared by the Forest Department.
- KFD (Kerala Forest Department). 2007. Forest reviews of year 2007. A report prepared by the Forest Department.
- Krishnasagar, M. 2003. Landscape characterisation of Jhabua and Ratlam district (Madhya Pradesh) using satellite remote sensing data and geographic information system. Forestry & ecology division Indian institute of remote sensing (NRSA) Department of space. Govt. Of India.
- Mathew, G., Rugmini, P and Sudheendrakumar, V.V.1998. Insect biodiversity in disturbed and undisturbed forests in the Kerala part of Western Ghats. *KFRI Research Report* 135.
- Riitters, K.H., Neill, R.V.O., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B., Jackson, B.L.1995. A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*. 10: 23–39.
- Roberts, S. A., Hall, G. B and Calamai, P. H.1998. Analyzing Forest Fragmentation Using Spatial Autocorrelation, Graphs and GIS. Region of Peel Planning Department publishing, USA.
- Roy, P.S and Behera, M.D. 2002. Biodiversity assessment at landscape level. *Tropical Ecology* 43(1): 151-171.
- Sathyapalan, C.K.2002. Carrying capacity based development plan of Sabarimala- A preliminary approach. *Indian cartographer*. 03:257-61.

- Singh, S. 2004. A geospatial approach landscape characterization and phytosociological analysis in Umaria District, Madhya Pradesh. FED,IIRS, Dehradun.
- Turner,W., Sacha, S., Gardiner, N., Fladeland, M., Sterling, E and Steininger, M. 2003. Remote sensing for biodiversity science and conservation. Trends in Ecology and Evolution. 18(6):1- 9.

Options for Biodiversity Monitoring within REDD+

Steffen Entenmann

M.Sc. Integrated Natural Resource Management
University of Freiburg
Institute for Landscape Management
Tennenbacher Straße 4
D-79106 Freiburg im Breisgau

KEYWORDS:

REDD+, Biodiversity, Monitoring

ABSTRACT:

The increasing deforestation and forest degradation in developing countries accounts for about 20% of all anthropogenic greenhouse gas emissions (IPCC 2007). The development of incentive schemes for Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+) is therefore currently one of the most important issues of the post-Kyoto negotiations of the United Nations Framework Convention on Climate Change. The idea behind REDD+ is to compensate developing countries for avoiding deforestation, thereby reducing CO₂ emissions. REDD+ has the potential ability to stimulate synergies between the climate protection goals of the UNFCCC and the objectives for biodiversity conservation set by the Convention on Biological Diversity.

The presentation introduces the research project “The protection of forests under global biodiversity and climate policy”, which aims to identify the institutional and technical preconditions necessary to obtain biodiversity benefits from REDD+.

Remote Sensing and Geographical Information Systems (GIS) play hereby an important role. They can be powerful instruments to monitor impacts of the activities that are likely to be carried out under REDD+ and to collect data for management decisions. An overview on the existing tools and direct and indirect approaches to monitor and assess biodiversity using Remote Sensing and GIS is given.

The research project is funded by the German Federal Agency for Nature Conservation with funds from the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety. For more information please refer to <http://www.landespflege-freiburg.de/forschung/redd.en.html>

The ArtWET tool: a Georeferenced Approach Assessing Runoff Related Pesticide Pollution in Surface Waters across Europe

David Elsaesser and Ralf Schulz

Institut für Umweltwissenschaften, Universität Koblenz-Landau, Fortstr. 7, 76829 Landau, Tel: +49-6341-280-31330, Fax: +49-6341-280-31326, E-Mail: elsaesser@uni-landau.de

KEYWORDS:

Pesticide, model builder, exposure, GIS, rainfall runoff

ABSTRACT

As an integrated part of the EU Life project ArtWET, a tool was built to model pesticide pollution in surface waters following rainfall runoff events at the European scale. The geodata used for simulation is taken from freely available sources. The OECD-REXTOX and USDA Curve number models were combined to calculate predicted concentrations in stream. The potential effects are modeled with the toxic units approach. First runs with several commonly used pesticides show a realistic range of runoff pollution and potential toxicity values. The whole approach was realized using ArcGIS model builder.

1 Introduction

Pesticides are widely used in agriculture, but adverse effects may be observed when the substances are transferred to natural ecosystems (Schulz, 2004). Nonpoint-source pollution through runoff, drainage and spray drift accounts for a majority of all surface water pollution (Zaring, 1996).

During peak application of pesticides in a watershed, a mixture of numerous substances may be transported to the waterbodies (Schulz, 2004; Battaglin and Goolsby, 1999; Thomas et al., 2001).

As an integrated part of the EU Life project ArtWET (Gregoire 2009) we focus on the role of vegetation in optimizing the potential of agricultural ditches and detention ponds for pesticide mitigation. In the present study, the focus was set on the development of a simple and user-friendly tool to assess the risk of pesticide pollution in small streams following a rainfall-runoff event.

2 Data and Methods

2.1 Geodata

The basis of the simulation is a geodata layer, which contains all agricultural area within a buffer zone of 50 m around European streams. Most of the source data was taken from freely accessible data portals of the European Commission Joint research center. The attributes are described in Table 1.

Attribute	Unit	Conversion	Source	Reference
Type of agriculture		-	Corine Land cover	EC 2006
Hydrological soil type		+	Soil database	Panagos 2006
Slope	%	+	SRTM	Werner 2001
Discharge	L/s	-	Hydrosheds	Lehner 2008
OC in topsoil	%	-	Soil Database	Panagos 2006
Length of Riversegments	m	+	EC-JRC, IES	Vogt 2007
Curve Number		+	USDA	Zhan 2004
Plant interception	%	-		Linders 2000

Table 1: Geodata

2.2 Pesticide and Event Data

To run the simulation further parameters of the substance and rain event need to be defined (Table 2).

Parameter	unit	Source
season	-	selection
Width of buffer strips	m	-
Precipitation amount	mm	Weather or climate data
Precipitation duration	minutes	Weather or climate data
Pesticide applied amount	g/ha	Pesticide registration
Pesticide: DT ₅₀	days	PPDB (2009)
Pesticide: K _{OC}	mL/g	PPDB (2009)
Pesticide: toxicity	µg/L	PPDB (2009)

Table 2: Parameters for simulation

2.3 Model Development

The model was programed in ESRI ArcGIS Model builder. It consists of ten main modules (figure 1). Module *a* is the input Dataset and a selected area of simulation. After calculating fields of the parameter precipitation and season, the plant interception based on Linders (2000) is calculated with the information of crop type and season in module *b*. The discharge in the waterbody is calculated in module *c*, based on season and data of hydrosheds (Lehner 2008). After calculation of the fields for the parameters “bufferstrips”, “toxicity”, DT₅₀, and K_{OC}, the K_d is calculated in module *d* by dividing the K_{OC} by the Content of organic carbon in topsoil. The discharge in of surface runoff (Q_{Runoff}) is calculated in module *e* using the precipitation amount and the Runoff Curve number model (Zhan 2004):

$$Q_{Runoff} = \frac{25.4 * (CN * (\frac{P}{2.54} + 2) - 200)^4}{CN * (CN * (\frac{P}{2.54} - 8) + 800)} \quad (1)$$

Q_{Runoff} : Peak flow in stream during rain event in L/s

CN : Curve Number

P : precipitation in mm

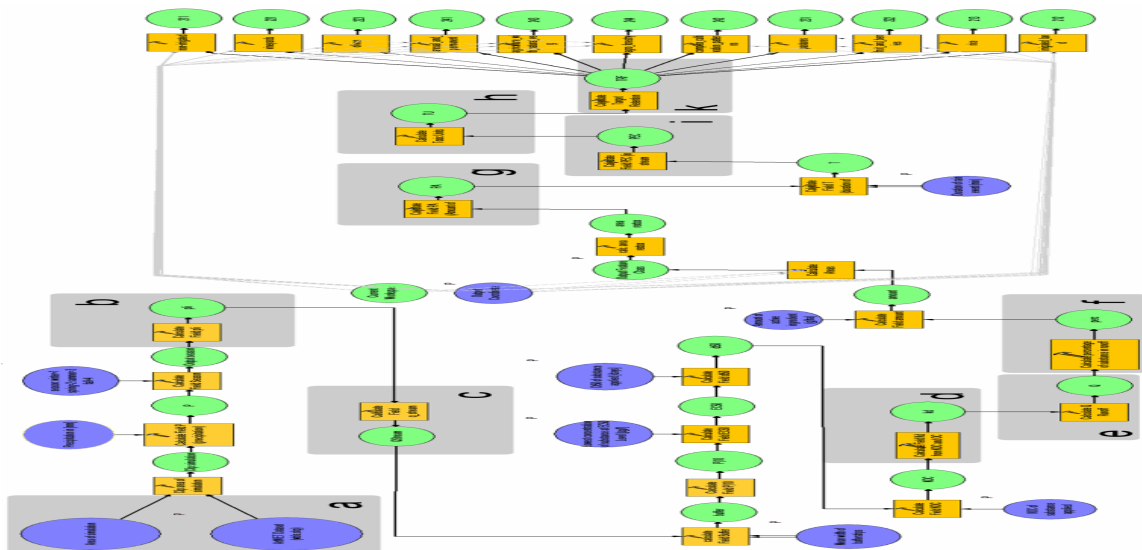


Figure 1: Overview of the model with modules a-k.

The percentage of applied amount within the surface runoff is calculated in module *f* using the modified REXTOX model (Probst 2005), that was proposed by the OECD:

$$L_{Runoff} = \left(\frac{Q_{Runoff}}{(P * 10)} \right) * e^{-3 * \frac{\ln 2}{Dt50}} * \frac{1}{1 + Kd} * \left(1 - \frac{Pli}{100} \right) * slope * 0.83^{Buffer} * 100 \quad (2)$$

L_{Runoff} : Percentage of applied substance in runoff

$Dt50$: Half life of applied substance in soil (days)

Kd : soil-water partitioning coefficient

Pli : Plant interception

$Slope$: the slope factor is calculated using the methods of Probst (2005)

$Buffer$: mean width of densely vegetated buffer strips

With use of the parameter “amount of active ingredient” and calculated areas of the single polygons of the dataset, the amount of substance applied on field is calculated in module *g*.

In module *h* the concentration of the substance in stream is calculated with the second part of the REXTOX model:

$$PEC = L_{Runoff} * PA * \frac{1}{Q_{Stream}} * T * 60 \quad (3)$$

PEC : Predicted in stream peak concentration in $\mu\text{g/L}$

PA : amount of substance applied in the simulation area in μg

T : duration of rain event in minutes

Acute toxicity data of the substances for fish, algae and aquatic invertebrates can be used to assess potential toxicity of the substance based on toxic units. Toxic units are calculated in module *i* for each peak concentration of the substance. Specific LC_{50} or EC_{50} values for acute toxicity to *Oncorhynchus mykiss* (fish LC_{50} 96 hours), *Daphnia magna* (aquatic invertebrate EC_{50} 48 hours), algae (EC_{50} growth 72 hours) can be found in the Footprint Pesticide Properties database (PPDB, 2009). The TUs are calculated using formula 4 (Peterson, 1994; Junghans et al., 2006):

$$PTU = \frac{PEC}{EC_{50}} \quad (4)$$

PTU : Potential toxicity in toxic units

EC_{50} : lowest concentration

The PTU value is used in module *k* to derive a target retention factor (TRF). The TRF describes a percentage of reduction of peak concentration, which is needed to mitigate the simulated pollution.

After the last calculation the result is split into polygon layers each containing the areas of different type of agriculture.

2.4 Graphical User Interface

The stylesheet of the user interface and the documentation was modified in order to make the tool easy to use. The documentation and help files were translated in English, German, French and Italian (figure 2).

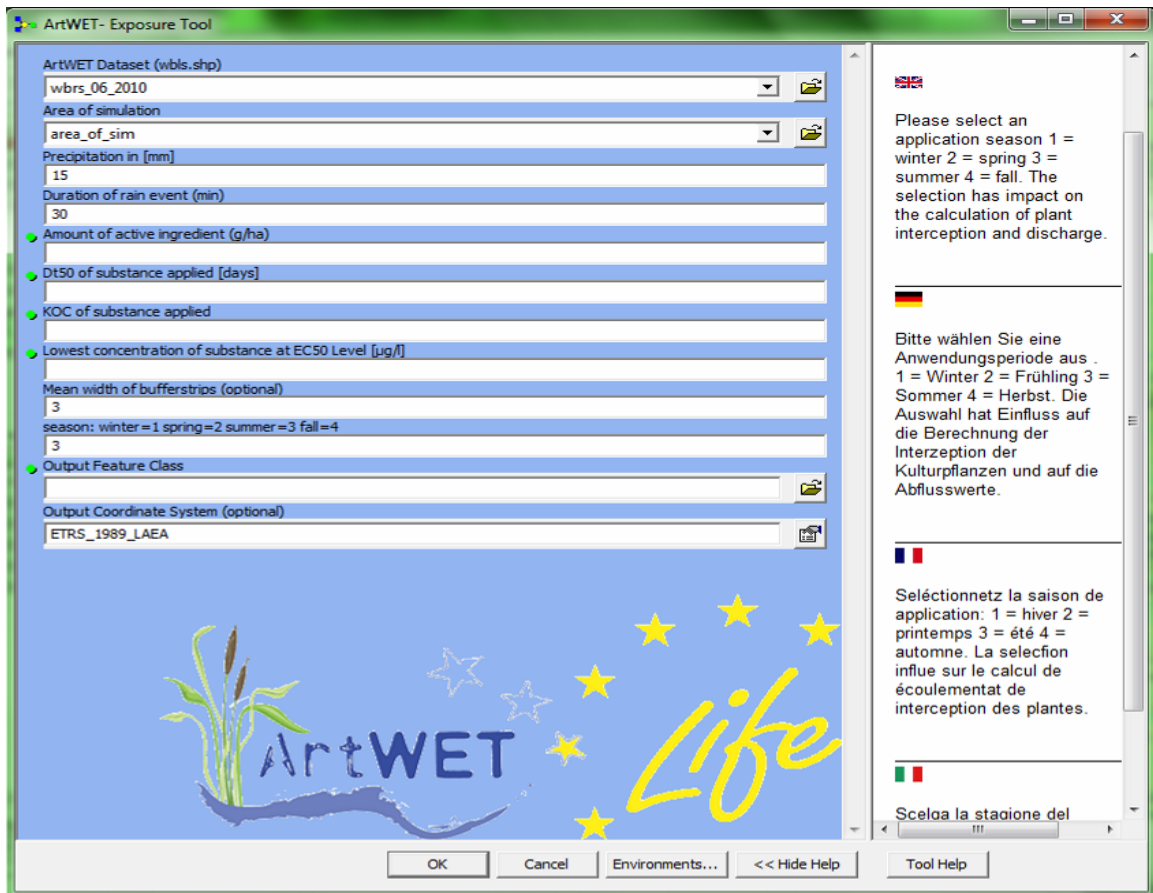


Figure 2: Grapical user interface and helpfile of the ArtWET model.

3 Results and Discussion

The exposure model was built using ArcGIS tools and Visual basic expressions. Verification of all modules was done. First runs with several commonly used pesticides show a realistic range of runoff pollution and potential toxicity values. Particularly the small streams with lower discharge show a high risk of adverse effects within the waterbody (figure 3). The predicted concentrations are within the same order of magnitude as measured peak concentrations in small streams. To perform a validation further monitoring data for all regions of the European Union is needed. Nevertheless the risk of pollution after runoff events is clearly displayed. Limitations of the model are mainly given by the quality of the source data. Especially the low resolved and relatively old landuse data (approx. 90 m) and values for the discharge in small streams during rain events could lead to inaccurate results. To deal with these uncertainties the database is updated regularly and ground truthing for selected regions has to be performed.

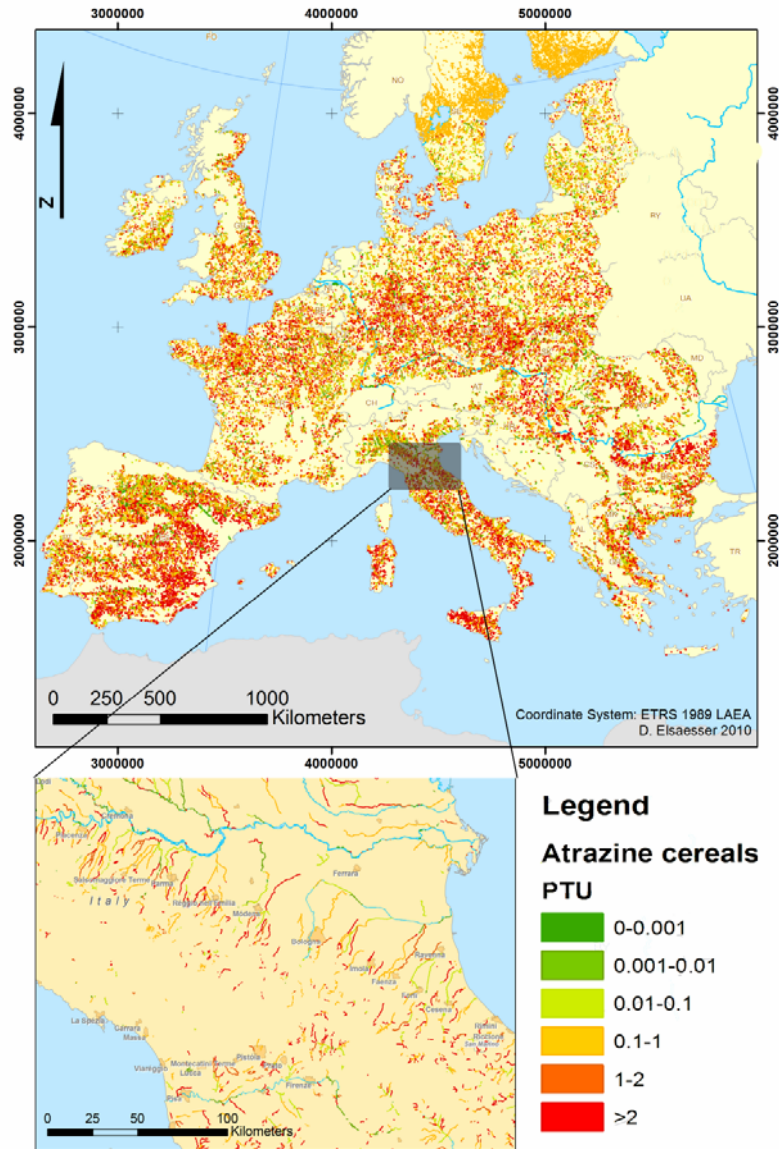


Figure 3: Result of running the model with the substance atrazine applied in summer on cereals with the maximum amount of 1000 g/ha

4 Conclusions and Outlook

The ArtWET risk of runoff exposure model is a promising tool to assess runoff derived pesticide pollution in regional scale. At this stage the model gives a clear output of runoff risk. To validate the quantitative results of simulations with the model on European level an extensive literature study is performed with data from runoff monitoring. Additional modules are developed to implement mitigation measurements into the model.

References

Battaglin, W.A., Goolsby, D.A., 1999. Are shifts in herbicide use reflected in concentration changes in midwestern rivers? *Environmental Science and Technology* 33, 2917–2925.

European Commission (2006): Corine land cover update 2000: Technical guidelines. Office for Official Publications of the European Communities, Luxembourg. ISSN 1725-2237

Gregoire, C., Elsaesser, D., Huguenot, D., Lange, J., Lebeau, T., Merli, A., Mose, R., Passport, E., Payraudeau, S., Schuetz, T., Schulz, R., Tapia-Padilla, G., Tournebize, 370 J., Trevisan, M., Wanko, A., 2009. Mitigation

- of agricultural nonpoint-source pesticide pollution in artificial wetland ecosystems. *Environmental Chemistry Letters* 7, 205–231.
- Lance McKee with Marc Van Liedekerke and Panos Panagos (2005):-21020 Ispra (VA) - Italy . OGC (Open Geospatial Consortium) - OGC User December 2005.
- Lehner, B., Verdin, K., Jarvis, A. (2008): New global hydrography derived from spaceborne elevation data. *Eos, Transactions, AGU*, 89(10): 93-94
- Panagos, P. (2006): The European soil database. *Geo: International* , July/Aug 2006 Volume 5 · Issue 7: pp. 32-33
- Peterson, Dennis R. (1994): Calculating the aquatic toxicity of hydrocarbon mixtures, *Chemosphere*, Volume 29, Issue 12, December 1994, Pages 2493-2506, ISSN 0045-6535
- PPDB (2009). The Pesticide Properties Database (PPDB) developed by the Agriculture & Environment Research Unit (AERU), University of Hertfordshire, funded by UK national sources and the EU-funded FOOTPRINT project (FP6-SSP-022704).
- Probst, M., Berenzen, N., Lentzen-Godding, A., Schulz, R. (2005): Scenario based simulation of runoff-related pesticide entries into small streams on a landscape level. *Ecotoxicology and Environmental Safety* 62/2 (2005) 145-159.
- Schulz, R., 2004. Field Studies on Exposure, Effects, and Risk Mitigation of Aquatic Nonpoint-Source Insecticide Pollution: A Review. *J Environ Qual* 33, 419–448.
- Thomas, K.V., Hurst, M.R., Matthiessen, P., Sheahan, D., Williams, R.J., 2001. Toxicity characterisation of organic contaminants in stormwaters from an agricultural headwater stream in south east England. *Water Research* 35, 2411 – 2416.
- Vogt, J.V. et al. (2007). A pan-European River and Catchment Database. EC-JRC (Report EUR 22920 EN) Luxembourg
- Werner, M. (2001): Shuttle Radar Topography Mission (SRTM), Mission overview, *J. Telecom. (Frequenz)*, v. 55, p. 75-79.
- Xiaoyong Zhan, Min-Lang Huang (2004): ArcCN-Runoff: an ArcGIS tool for generating curve number and runoff maps, *Environmental Modelling & Software*, Volume 19, Issue 10, Pages 875-879, ISSN 1364-8152
- Zaring, D., 1996. Agriculture, nonpoint source pollution, and regulatory control: The clean water acts bleak present and future. *Harvard Environmental Law Review* 20.

Geographic Cataloging of Ethnobotanical Information – A Review of Ethno medicinal uses of *Saraca asoca* (Roxb.) De wilde in India

R. Jaishanker, Arathyram R.S, V. Saroj Kumar, and C.S.P Iyer

School of Ecological Informatics
Indian Institute of Information Technology and Management – Kerala
Technopark Campus, Kariyavattom (PO)
Trivandrum-695 581, Kerala, India

KEY WORDS:

Ethnobotany, *Saraca indica*, GIS.

EXTENDED ABSTRACT

India is one of the richest floristic regions of the world and has been a source of plant products since antiquity. 35,000-70,000 species of plants are reported to be used for medicinal purposes¹. Medicinal plants occupy an important position in the socio-cultural, spiritual and therapeutic arena of people of India. 80% of the population relies on traditional medicine, mostly plant drugs, for their primary health care needs².

The ethnic groups in different regions of India use medicinal plants to cure various ailments. This indigenous knowledge is based on their necessities, instinct, observation, trial and error and long experience. Some of this knowledge got widely tested and accepted over period of time to become part of the recognized Indian systems of medicine (ISM). However, accelerating erosion of indigenous communities leads to irreplaceable loss in tacit knowledge within them. Where indigenous knowledge is lost through acculturation people will be drawn into unsustainable ways of using natural resources. Ethnobotanists are racing against time to assimilate indigenous knowledge.

Over the past couple of decades there has been a resurgent interest in traditional medicine. This is closely linked to both the conservation of species and the instance of knowledge, respectively. Ethnomedicine (folk medical claims) are an important constituent of this knowledge. The multi-locational or multi ethnic use, i.e. application of same indigenous herbal cures in distant, disconnected regions, and among unrelated ethnic groups is a good criterion for validation and credibility of such claims.

The strong links between biological and cultural diversity uniquely position ethnobotany to help craft effective local solutions to many of the global issues that confront mankind. It is of central importance for understanding the collective experience of humankind in a series of exceedingly diverse environments and using those experiences to meet the challenges that we face. Ethnobotany is vital to preserving the diversity of plants as well as to understand and interpret the knowledge by which we can address the challenge effectively and sustainably.

This communication is a specific case of Asoka [*Saraca asoca* (Roxb.)De wilde] from the bigger framework where the authors attempt geographic cataloging of ethno medicinal uses of medicinal plants in India. Asoka, belonging to family Caesalpiniaceae is a small evergreen tree. It has high medicinal value and is used in many ayurvedic drugs. Bark contains tannins, glycosides and alkaloids (leucopelargonidin and leucocyanidin). Bark of the tree is used in medicines for women's genitor-urinary socaproblems, dyspepsia, fever, burning sensation, visceromegaly, piles, ulcers, menorrhagia, metropathy, leucorrhoea and pimples. Due to overexploitation of this species, IUCN has listed this species as globally vulnerable³. In the present communication, an attempt is made to present an overview of the ethnobotanical uses of *Saraca asoca* by various indigenous communities from the existing published literature by using GIS tools. Such spatial compilation also serves the purpose of establishing a net work among the researchers and community involved in the conservation and documentation of Rare, Endangered and Threatened (RET) plants. We compiled information from 15 peer-reviewed papers from different journals on the indigenous uses of *Saraca asoca*. The study discovered the ethnobotanical potential of *Saraca*, with 15 different medicinal uses, across more than 50 ethnic groups in India was spatially depicted using a Geographic Information System (GIS). Figure 1 is a screenshot of various ethno medicinal uses of *Saraca asoca* across different ethnic groups of India.

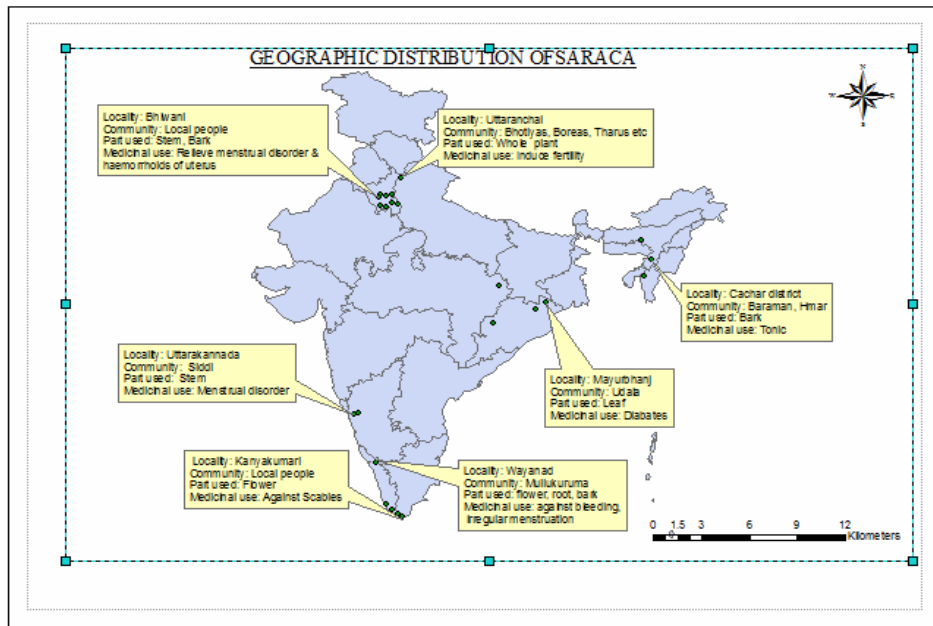


Figure 1: Spatial depiction of ethnomedical uses of *Saraca asoca* by various ethnic groups in India.

References

- Maiti, S. Air layering –A new propagation method of *Asoca*, *Newsletter*, Vol.9: 1, DMAPR, 2008
- Ponnu, S, Santhi, D.K Jacob, N & Suresh B; Safety measures with herbs, *Indian pharmacist*;2:9-12.2003
- Tewari.D.N. Report of the Task Force on *Conservation & Sustainable use of Medicinal Plants*, Planning Commission, Government of India.2000.

Morphometric Analysis and Prioritization of Microwatersheds in Valiyathodu Sub-Watershed of Manimala River, Kerala, India

V.B. Rekha^{1,4}, A.V. George², K.R.Vineeth¹ and U. Johnson¹

¹ School of Environmental Sciences, Mahatma Gandhi University, Kottayam 686 560, India-
rekhavb1@gmail.com

² Christ College, Irinjalakkuda, Kerala 680125, India

KEY WORDS:

Morphometry, Bifurcation ratio, Stream frequency, Texture ratio, overland flow, Kerala, India, GIS

ABSTRACT:

In this study, morphometric analysis and prioritization of eight microwatersheds of Valiyathodu subwatershed of Manimala River, Kottayam district, Kerala, India is carried out using GIS. The morphometric parameters are stream length, bifurcation ratio (Rb), (linear parameters), drainage density (Dd) (relief parameters), texture ratio (T), stream frequency (Fs), circulatory Ratio (Rc), constant channel maintenance (Ccm) and length overland flow (Lof) (aerial parameters). The highest bifurcation ratio among all microwatersheds is 3.33 which indicates the strong structural control on the drainage. The maximum value of drainage density is 3.24 for the microwatershed named V8, which has weak or permeable subsurface material. High value of stream frequency 12.24 indicates the high relief and low infiltration capacity of microwatershed V2. The higher texture ratio value of 3.15 of V2 can be attributed to the high relief and low infiltration capacity of the area. Specific scores and weightage were given to calculated values of morphometric parameters based on the water holding capacity and five classes of prioritized zones were demarcated. Here it has been found that the very high zone (12.6 km²) shows the maximum water holding capacity.

1 Introduction

Our survival on earth essentially depends on water, nature's valuable gift to mankind. Water resources are extremely limited but renewable and exhibits diversity in its quality and quantity. Today the world is facing a water crisis and the available surface and ground water resources are insufficient to meet mankind's requirements. The optimal and sustainable development of the resources is a pre-requisite so that it is assessed reliably to avoid any future problems regarding its qualitative and quantitative availability (Schumm, 1956). Assessment of various morphometric parameters is essential while exploring the water potential zones and this study has been framed with the objective of clearly calculating the morphometric parameters of eight microwatersheds of Valiyathodu subwatershed of Manimala River and demarcating the prioritized zones based on water holding capacity.

Valiyathodu is the midland subwatershed of Manimala River and it lies between 9°30'00" to 9°36'40" N latitude

and 76°40'50" to 76°45'50" E longitude (Figure.1) and it covers an area of 54.83 km². Geologically the

formations are Precambian in origin and the Valiyathodu subwatershed is characterised by charnockite, gabbro, dolerite and quartzite. Land use pattern of Valiyathodu sub-watershed is rubber plantations, mixed crop, tapioca, pineapple, urban area and settlements.

Morphometry is the measurement and mathematical analysis of the configuration of the Earth's surface, shape and the dimensions of its landforms (Clarke, 1996; Horton, 1945). This analysis can be achieved through

⁴ Corresponding author.

measurement of linear, aerial and relief aspects of basins and slope contributions (Nag and Chakrabarty, 2003). Earlier, morphometric analysis using remote sensing techniques has been attempted by Srivastava and Mithra (1995), Srivastava (1997), Nag (1998) and all have concluded that remote sensing has emerged as a powerful tool in morphometric analysis. Geographical Information System (GIS) techniques have already been used for assessing various terrain and morphometric parameters of the drainage basins and water sheds as they provide a flexible environment and a powerful tool for the manipulation and analysis of the spatial information particularly for better understanding of the future scenario (Vijith, 2006).

2 Methodology

Valiyathodu subwatershed and associated drainage networks were delineated from Survey Of India topographical maps (No.58 C/10, of 1:50000 scale) and digitized using Arc GIS software and the morphometric parameters calculated. The methodology adopted is given in table 1 and the results are summarized in table 2. Morphometric analysis for stream length, bifurcation ratio (Rb), (linear parameters), drainage density (Dd) (relief parameters), texture ratio (T), stream frequency (Fs), circulatory Ratio (Rc), constant channel maintenance (Ccm) and length overland flow (Lof) (aerial parameters) has been carried out. Morphometric analysis is a significant tool for prioritization of microwatersheds even without considering the soil map. For the prioritization of microwatershed, the subwatershed was separated into eight micro watersheds and the morphometric parameters of the delineated microwatersheds were carried out separately as per the methodology prescribed by Strahler, 1957. After the calculation of morphometric parameters, the calculated values were added as attribute data and the thematic maps were prepared by using the morphometric parameters such as drainage density, bifurcation ratio, stream frequency, texture ratio, constant channel maintenance, and length overland flow. After the preparation of each theme, a specific weightage was given to each theme and specific ranks were given to individual features based on water holding capacity of both subwatersheds. The water holding capacity of an area is controlled by various factors need not be same in the area. Therefore each parameter is assigned a weightage based on water holding capacity. Here the higher weightage is given to drainage density followed by bifurcation ratio. On the other hand the influence of length overland flow and constant channel maintenance being comparatively less, thus the lower weightage is given to this. After the criteria for weightage were finalized for each theme the different units of each theme were assigned a knowledge based scores ranging from 1 to 10 based on their influence in holding water. These are represented in table .2. After assigning the weightage and rank to each theme and individual features all the themes were converted into raster format using spatial analyst extension of Arc GIS software. While converting into raster the scores assigned to the individual features were taken as the value field. Then the individual themes were normalized by dividing the theme weightage by 100. The raster calculator option of spatial analyst was used to prepare the integrated final prioritized maps for both sub-watersheds. The map algebra used in the raster calculator is,

$$\text{Prioritized map} = \text{Drainage density} \times 0.3 + \text{Bifurcation ratio} \times 0.25 + \text{Stream frequency} \times 0.2 + \text{Texture ratio} \times 0.15 + \text{Length overland flow} \times 0.1 + \text{Constant channel maintenance} \times 0.05$$

Morphometric Parameters	Methodology	Where, Lu=Total stream length of Order 'u' Nu=Total number of stream segment of order 'u' Nu+1=Number of segment of next higher order L=Total length of streams A=Area of watershed N=Total number of streams N1=Total number of first order streams P=Perimeter of watershed Lb=Basin length
Linear Parameters		
Stream Length (Lu)	Length of stream	
Bifurcation Ratio (Rb)	$Rb = Nu / Nu+1$	
Aerial Parameters		
Drainage Density (Dd)	$Dd = L/A$	
Stream Frequency (Fs)	$Fs = N/A$	
Texture Ratio (T)	$T = N1/P$	
Length overland flow (Lof)	$Lof = 1/2Dd$	
Constant channel maintenance(Ccm)	Inverse of drainage density	

Table.1.:Methodology adopted for the computation of morphometric parameters.

DRAINAGE-VALIYATHODU SUB WATERSHED

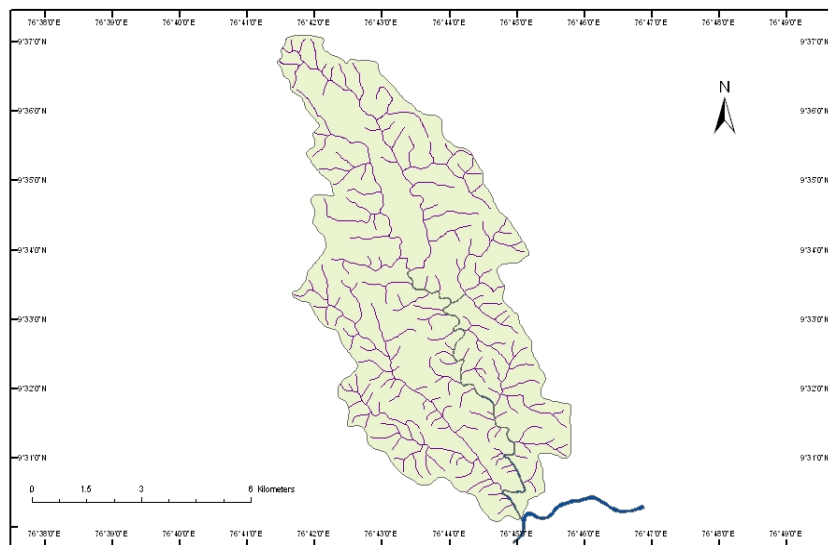


Figure 1: Drainage of Valiyathodu subwatershed

PRIORITIZED MAP OF VALIYATHODU

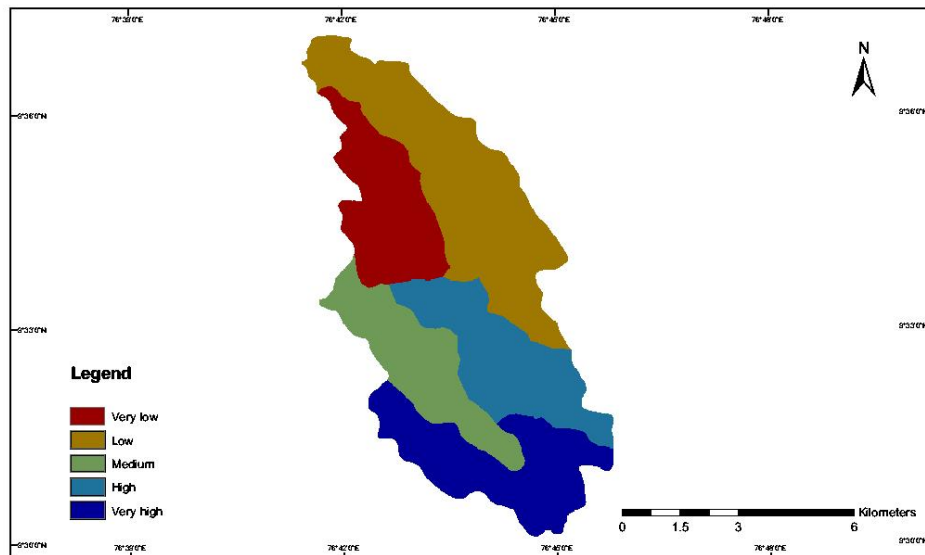


Figure 2. Prioritized map of Valiyathodu

Name of micro-watershed	Area	Perimeter	Lu	Lsm	Rb	Dd	Fs	Text	Lof	Ccm
V1	8.727	15.399	19.33	0.62	1.80	2.21	3.55	1.03	0.22	0.45
V2	6.614	14.587	21.20	0.26	1.71	3.20	12.24	3.15	0.15	0.31
V3	8.360	17.123	23.69	0.31	1.97	2.83	8.97	0.81	0.17	0.35
V4	12.959	20.097	31.39	0.54	2.69	2.42	4.47	1.54	0.20	0.41
V5	4.502	10.595	11.89	0.41	1.60	2.64	6.44	1.41	0.18	0.37
V6	3.076	8.482	9.26	0.40	2.91	3.01	7.47	1.41	0.16	0.33
V7	3.713	9.927	11.44	0.44	3.33	3.08	7.00	1.81	0.16	0.32
V8	6.890	13.685	22.35	0.50	2.37	3.24	6.38	1.82	0.15	0.30

Table 2: Result of morphometric analysis of Valiyathodu micro-watersheds

Themes	Weightage for themes	Layer class	Rank
Drainage density	30	2.20-2.40	1
		2.40-2.60	3
		2.60-2.80	5
		2.80-3.00	7
		3.00-3.20	9
		3.20-3.40	10
Bifurcation ratio	25	1.50-1.80	1
		1.80-2.10	3
		2.10-2.40	5
		2.40-2.70	7
		2.70-3.00	9
		3.00-3.30	9
Stream frequency	20	3.30-3.60	10
		3.50-5.00	10
		5.00-6.50	9
		6.50-8.00	8
		8.00-9.50	5
		9.50-11.00	3
Texture ratio	15	11.00-12.50	1
		0.80-1.20	10
		1.20-1.60	9
		1.60-2.00	9
		2.00-2.40	5
		2.40-2.80	3
Length Over Land Flow	10	2.80-3.20	1
		0.150-0.160	10
		0.160-0.170	9
		0.170-0.180	9
		0.180-0.190	8
		0.190-0.200	6
		0.200-0.210	4
		0.210-0.220	3
0.220-0.230	1		
Constant channel maintenance	5	0.300-0.340	10
		0.340-0.380	8
		0.380-0.420	4
		0.420-0.460	1

Table.3: Weightage and rank given to themes and individual features

3 Results and Discussion

The highest stream order for Valiyathodu subwatershed is 4 and the number of streams is 311. Among the delineated microwatersheds, bifurcation ratio is maximum for V7 and minimum for V5. Lower values of Rb are characteristic of microwatersheds which have suffered less structural disturbances and the drainage pattern has not been distorted and from this it has been found that V7 is geologically more complex and structurally more controlled than V5. Drainage density is one of the important indicators of the fluvial control. Its measurement provides a numerical measurement of landscape dissection and runoff potential. Here the high drainage density of 3.24 of V8 can be attributed to weak or impermeable subsurface material and high resistance to infiltration. Higher value of stream frequency of 12.24 is obtained for V2 and lower value 3.55 for V1. Texture ratio is an important factor which depends on the underlying lithology, infiltration capacity and relief aspect of the terrain. Here it is found that microwatershed V2 has high relief and low infiltration capacity. Constant channel maintenance is the inverse of drainage density and it depends on the rock type, permeability, vegetation cover, as well as duration of erosion. From this it has been concluded that V8 has less infiltration capacity and high runoff and consequently V8 has less water holding capacity. The integrated final prioritized map shows five classes of prioritized zones which are, very high, high, medium, low, very low (Figure 2). Here it has been found that the very high zone (12.6 km²) shows the maximum water holding capacity. In this zone, the bifurcation ratio is maximum, texture ratio and stream frequency is also minimum.

4 Conclusion

The morphometric parameters evaluated using GIS helped understand various terrain parameters such as nature of bedrock, infiltration capacity, surface runoff etc. The present study will help in watershed development and management. Among the five demarcated prioritized zones, the maximum water holding capacity is shown by the very high zone and the maximum importance will be given to this zone during the exploration of water potential zones in the future for watershed management.

References

- Clarke, J.I., 1966. Morphometry from Maps. Essays in geomorphology. Elsevier, New York, pp. 235 – 274.
- Horton, R.E., 1945. Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology. *Geo Soc. America Bulletin*, 56, pp.275-370.
- Nag, S. K., 1998. Morphometric analysis using remote sensing techniques in the Chaka sub-basin, Purulia District, West Bengal. *J. Indian Soc. Rem Sens*, 26 (1& 2), pp 69 - 76.
- Nag, S.K. and Chakraborty S., 2003. Influence of rock types and structures in the development of drainage network in hard rock area. *J. Indian Soc. Rem. Sens.*, 31(1), pp. 25-35
- Schumm, S.A., 1956. Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *Geology*, 67, pp. 597- 646.
- Srivastava, V.K., 1997. Study of d pattern of Jharia coalfield (Bihar), India, through remote sensing technology. *J. Indian Soc. Rem. Sens.*, 25(1), 41-46
- Srivastava, V.K. and Mitra, D., 1995. Study of Drainage pattern of Raniganj coalfield (Burdwan District) as observed on Landsat-TM/IRS LISS II Imagery. *J. Indian Soc.Rem. Sens.*, 23(4), pp. 225-235.
- Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. *Transactions of the Am. Geophys. Union*, 38, pp. 913- 920.
- Vijith, H. and Satheesh, R., 2006. GIS based morphometric analysis of two major upland sub-watersheds of Meenachil river in Kerala. *J. Indian Soc. Rem Sens*, 34 (2), pp.181- 185.

Acknowledgement

The authors gratefully acknowledge the Director, School of Environmental Science, Mahatma Gandhi University, Kottayam for the support extended during the course of study.

Geoinformation for Tourism Development

Mirka R. de Zimmermann^{a,*} and Maria Fernanda Serra^b

^a Eberhard Karls University Tübingen
Rümelinstraße 19-23, 72070 Tübingen, Germany, mirka_rod@gmx.de

^b KIA Lands, Canada

KEYWORDS:

Tourism Potential, Sustainability, Regional Development, Geoinformation

ABSTRACT:

This paper handles the significance of Geoinformation in order to support the analysis and planning of Tourism as contribution to the regional development in Bolivia. The main motivation for this project is to know more about the tourism potential in some regions and how planers and tourists could be supported by and take advantage from Geoinformation availability.

1 Introduction

Bolivia belongs to those countries with worldwide the biggest biodiversity features. Moreover, the country owns very reach ancient cultures. Although all these together mean a big potential for tourism, Bolivia plays not a big role at the International Tourism Market. However, the significance of this sector to the national economy is not to underestimate. The international tourism is for the country between the main foreign currency sources after the Soya export and Oil derivates.

During the last years tourism has been considered in some rural Communes in Bolivia as possible potential development factor to be implemented. According to the principles of sustainable development this process must lead to the population welfare increment, conserving or rather improving the ecological stability. That means for the tourism, that its development must occur according to these principles. The sustainability can then be appreciated when it improves the population quality of life in such a way that still offers to later generations social security and therefore no irreparable resource damage will occur (Vorlaufer, 1996).

Resolution processes of the decision-makers in the context of sustainable development regarding tourism, planning and development will be always a complex process. Nowadays, communes and other organizations must also consider social, economical and environmental aspects and include them into the decisions. The use of Geoinformation based solutions can then positively contribute not only to the impact's estimation of possible decisions, but mainly along the analysis and planning processes as well.

A recent study regarding the trend about nature-based tourism found that in 15 of 20 analyzed countries the number of visitors to their nature reserves increased. According to their findings the visitor numbers in the US and Japan have been falling slightly since the 1990's although in three-quarters of the 20 countries, the visitation to nature reserves is increasing approximately in 7 or 8 percent per year. The increases were on average positive in Africa, Europe, Asia and Latin America (Balmford, 2009).

This paper handles the significance of Geoinformation in order to support the analysis and planning of Tourism as contribution to the regional development. The whole study is still under work, however general preliminary ideas will be presented.

* Corresponding author.

2 Data and Methods

2.1 Data

Like in other South American countries (Netzband, 2007) the institution in charge of the creation and distribution of georeferenced information i.e. topographical maps is the IGM (Military Geographic Institute). Mainly, due to budget limitations most of those maps were not updated at least during the last decade. Nevertheless, topographical maps corresponding the areas of study in scales 1:50000 and 1:100000 were acquired from IGM for this study.

The information regarding population, including aspects like density and socioeconomic measures are available through the National Statistics Institute (INE). Moreover, every four years they publish information concerning the main characteristics of the international Tourism in Bolivia (INE, 2007).

On the other side, although satellite data is being processed through consultancy companies and research institutions their results are not enough accessible or broad disseminated. In general, digital data come from different sources and usually in different projection systems and need format transformations and re-projection.

In order to make some progress towards an efficient management of natural and related resources is fundamental to know where are they located or at least to know how to have access to that information. In this regard, a list including some important nature and culture attractions in different routes of the study area was tracked by means of a Navilock-GPS device.

2.2 Methods

Actually not only GPS technology and remote sensing provide us with the necessary means in order to substantially improve the availability of updated geographic information. As already Goodchild (2007) points out:

“The Global Positioning System (GPS) has revolutionized the processes of surveying, allowing the rapid and accurate determination of absolute position on the Earth’s surface, and remote sensing provides a massive and constant flow of Earth imagery. But many other types of geographic information are not visible from above, or cannot be extracted from imagery by any automated process. They include the names humans attach to features, otherwise known as geonames or gazetteer entries; environmental information, including measures of air quality; cultural information, including information on the use of land and buildings; and population information, including population density and socioeconomic measures.”



Figure 1: Study areas in Cochabamba, Bolivia.

The areas of study are located in Cochabamba (see Fig. 1) and they were regularly visited during a couple of fieldworks, every time taking more than two months duration. Moreover, some empirical research was also conducted, i.e. data gathering under qualitative and quantitative aspects, Interviews from experts, key persons at different local levels. Two main target groups were identified: the stakeholders from different institutions involved in the regional development and the tourism service providers i.e. Tourism experts, etc.

Some questionnaires were designed, which covered different aspects from the tourism offer and demand until the stakeholder expectations regarding regional development plans. In both cases, the intention was to find out about the general aspects regarding the current situation and later on concerning the advantages and limitations they find in their regions themselves. Usually, the tourism related issues within the communes are handled by: the Human Development and the Productive Development and Environment Units.

The questionnaires concerning the available potential focused on aspects like location of main nature and culture attractions up to the available infrastructure. Actually, the fieldwork was also combined with DTM derived from SRTM data and analysis of Google Earth images.

3 Approaching the Tourism Potential

3.1 Points of Interest

Getting GPS-tracks along principal tourism attractions can be not only a time-consuming but also an expensive process. There are some Internet services providing free of charge georeferenced data, though under limited use conditions. Moreover, it was possible to verify that information concerning the study areas although is available through sites like for example OpenStreetMap (<http://wiki.openstreetmap.org/wiki/Bolivia>) or Google Maps (<http://maps.google.com/>), being the data degree of detail different from region to region. During the present study some examples of route tracks were saved, these routes included some of the infrastructure availability on the study area surroundings. Actually, the visualization of the tracks over Google Earth, provides a good insight of the area taking into account that there not exists such kind of maps yet.

3.2 Technical Issues

The use of Geoinformation based systems in the area of tourism planning and management began only at the beginning of the 90's (Bahaire, 1999). However, its application spread not far away until the last years when Google Earth appeared and further tourism oriented applications have been since then implemented.

In order to overcome the current lack of basic updated mapping from some areas in the country low-cost and free-ware solutions (Emde, 2008, Zeil, 2007) could be an appropriate technological alternative. Furthermore, gathering additional geoinformation from main tourism routes within the study areas could be possible by i.e. training the trekking guides on the use of GPS for routing-tracking and mapping purposes.

Therefore, the cartography of the study areas would be complemented and could be integrated as Tile Map Service overlay, for example (<http://www.outdooractive.com>). These maps could also include e.g. a layer for the slopes inclination in order to help the user to be aware of possible danger on some routes. Further functionalities like GPS tracks, altitude profiles, etc could be also incorporated. By having access to the updated digital cartography, the tourists could create their own outdoors-tours by using a routing-planning system. This kind of tool could for example help by designing tours along favorite points of interest, hiking and biking trails, etc. The system could be build based on any open source map technology supporting not only special functionalities for outdoor tours-planning but other tourism regarding activities, e.g. including some tourism potential attractions on the journey as well.

4 Conclusions

Nowadays, most of the information concerning tourism destinations is available on the Internet. Every time more and more travellers start already the trip searching for some detailed information regarding, i.e. destinations, lodging and tour facilities, last minute offers, travelogues, etc. over there (Brözel, 2010). Therefore, prior to the trip the tourist is already somehow informed and knows what he wants and what to expect from his destination. Additionally, it would be useful for tourists having access to some geoinformation in order to identify possible risks and qualities from the destination, between other handy issues.

On the other side, for tourism planners getting to know what and where are the tourism potential supports the regional characterization or differentiation towards strengthen the regional identity. This factor is very important for future regional marketing purposes as well.

Concrete steps to be suggested at this stage would be for example that the units in charge of tourism from the communes should have more interaction with travel agencies and tourism experts or guides, in order to improve the planning and promotion of the tourism concept from their respective regions.

Bibliography

- Bahaire, T. et al. (1999) The application of Geographical Information Systems (GIS) in Sustainable Tourism Planning: A review. *Journal of Sustainable Tourism*, Vol. 7, No.2, U.K. Lincoln.
- Balmford, A. et al. (2009) A Global Perspective on Trends in Nature-Based Tourism. *PloS Biol* 7(6):e1000144, doi:10.1371/journal.pbio100144. USA
- Brözel, C., Wagner, W. (2010) *Tourismus und Internet, Reisen und Reisevorbereitung in der neuen Informationswelt*. Uni-edition GmbH, Berlin.
- Emde, A.; Christl, A. (2008) *Webmapping mit Open Source GIS-Tools*. O'Reilly.
- Goodchild, M. (2007) Citizens as voluntary sensors: Spatial Data infrastructure in the World of Web 2.0. *International Journal of Spatial Data Infrastructures Research*, Vol. 2, p. 24-32.
- Google Maps: <http://maps.google.com/> [last accessed 18.07.2010]
- Instituto Nacional de Estadística (2007) *Encuesta del Gasto del Turismo*, La Paz – Bolivia.
- Open Street Map: <http://www.openstreetmap.de/> [last accessed 18.07.2010]
- Outdoor Active: <http://www.outdooractive.com/> [last accessed 18.07.2010]
- CGIAR Consortium for Spatial Information: <http://srtm.csi.cgiar.org/> [last accessed 18.07.2010]
- Netzband, M. et al. (2007) *Applied Remote Sensing for Urban Planning, Governance and Sustainability*. Springer Verlag Berlin Heidelberg, p. 77-95.
- Vorläufer, K. (1996) *Tourismus in Entwicklungsländern: Möglichkeiten und Grenzen einer nachhaltigen Entwicklung durch Fremdverkehr*, Darmstadt: wissenschaftliche Buchgesellschaft, pp. 123-124.
- Zeil, P., Kienberger, S. (2007) *Geoinformation for Development, bridging the divide through partnerships*. Herbert Wichmann Verlag, Hüthing GmbH & Co. KG, Heidelberg.

WHO's Public Health Mapping and GIS Programme

Imran Hossain^{a*}, Tania Puspita Firdausy^b and Franz-Josef Behr^c

Department of Geomatics, Computer Science and Mathematics
University of Applied Sciences Stuttgart

Schellingstraße 24, D-70174 Stuttgart, Germany

^a rana_soil@yahoo.com, ^b tania.puspita@gmail.com,

^c franz-josef.behr@hft-stuttgart.de

KEYWORDS:

WHO, Health Mapper, GIS, Public Health, Interactive Mapping

ABSTRACT:

Nowadays spatial related data is getting more attention and considered to be very important. Many organizations and even countries now consider spatial data as one of their assets for development. This is due to the reason that it identifies geographic location of features and boundaries on earth. With this information a country or an organization can identify their geographical strengths and weaknesses. The World Health Organization has also realized this potential in spatial data to help mitigate and control the spread of diseases around the globe. In order to cater this notion they developed a simple user friendly GIS tool and services which helps the organization perform their daily GIS routines and spread information to the people who are concerned. These tools and services are The Health Mapper, The Global Health Atlas and The Remote Field Data Collection Tool. These tools and services are available to be downloaded from the WHO website. This paper will describe the main functionalities of the available tools and services developed by WHO.

1 Introduction

The World Health Organization (WHO) is the directing and coordinating authority for health within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries, monitoring and assessing health trends (WHO, 2010).

It is required to provide up to date spatial data to decision makers at all of the levels that are concerned with the public health system. This is due to the reason that up to date spatial data is needed to properly plan, manage and monitor, any public health programme. Geographic Information Systems (GIS) and remote sensing data are necessary for WHO in order to keep the surveillance and control of infectious diseases, the determination of geographical distribution of diseases, analyzing spatial and temporal trends of diseases, mapping populations at risk, stratifying risk factors, assessing resource allocation, planning and targeting interventions and monitoring diseases and interventions over time (WHO, 2010). Therefore, GIS is an essential tool for such an organization. Realizing the necessity of GIS for public health information matter, the WHO has developed a program named the WHO public health and GIS programme in 1993. Thus, the main goal of the present study is to explore the present GIS programme, services and options offered by WHO in their website for public health mapping.

2 History of GIS in WHO, goals and strategic objectives

The public health and GIS programme used by WHO was firstly developed in 1993 with UNICEF to push the attempt to prevent the spread of the guinea worm disease that occurs in isolated, poor rural areas. The main goal of this program was to strengthen the capacities of infectious diseases and public health programmes at local, national and global levels through the use of GIS and mapping technologies (WHO, 2010). The programme has two main objectives:

- To strengthen national surveillance, prevention and control activities through the delivery of user-friendly data collection, management and mapping applications.

- To strengthen the management, analysis and monitoring of priority infectious diseases at global levels, through the development of rapid electronic surveillance systems supported by user-friendly web-based mapping.

3 WHO GIS services

To achieve these objectives the World Health Organization has developed some GIS related tools, applications and services that are intended to help, create and manage critical spatial information related to infectious diseases and public health programmes. The tools and services are explored below

3.1 The Health Mapper

The Health Mapper is a GIS tool developed by WHO to provide critical information concerning with WHO infectious disease programmes at national and global levels. It is a user-friendly data management and mapping system customized specifically for public health users. The system facilitates data standardization, collection and updating of data on epidemiology and on interventions and provides immediate visualization of data in the form of maps, tables and charts. The purposes of the Health Mapper tool are given in the next page (Health Mapper 4.3 Learners Guide part 1, 2008):

- To give public health users a readymade standardized digital database containing information considered essential by most of them including country boundary maps, environmental factors (such as lakes, rivers, elevation) as well as vital information on basic population and basic health, school and water infrastructures.
- Provide the user with user-friendly icon driven functions to automatically create maps, tables and charts of their data.
- Provide a simple data management interface in which the user can easily enter and update public health indicators in a standard geographic format.
- Make the system coherent to operate at both local and global levels.
- Provide the above to the users at low or no cost.

The system used by this tool is based on Map Objects, so this tool is compatible with ArcView. The data formats it supports include shape files, .tif, .jpg, and .bmp. The database it uses is Microsoft Access. As stated in the Health Mapper documentation there are some key features of Health Mapper tool (see fig.1), it is possible to:

- Create thematic maps
- Create interactive graphs/charts (line graphs and bar charts)
- Create interactive reports of data
- Create, save and retrieve a set of favourite maps
- Selection, measurement zoom in and out, and panning capabilities
- Overlay additional layers
- Buffer creation around selected features
- Ability to create interactive graphs from maps to help asses trends over time.

With all the functionality stated above the Health Mapper tool is intended to support and make decision making more efficient, easier, faster, and more reliable, as it will be based on the information or facts gathered and compiled by this tool. The tool will help in the assessment of resource allocation, monitoring geographical distribution of the disease, determining the relationships and correlation between diseases help to determine intervention targets, make the data collection, analysis, and management more reliable, and monitor diseases and intervention over time.

The Health Mapper tool is also equipped with a data management tool. This tool enables users to incorporate program indicators to the geographic database. Also it facilitates users to integrate their own public health data. Besides the data management tool, the Health Mapper tool consists of several extension tools, namely:

1. The shape editor
2. The import geographic data tools

3. The point to polygon tool

The Health Mapper also packages a database of core baseline geographic, demographic and health information, including the location of communities, health care and education facilities, accessibility by road, access to safe water and demography. The system is currently in operation to support a range of infectious diseases in over 60 countries in all regions of WHO. Key infectious disease programmes currently using the system include Roll Back Malaria, Lymphatic filariasis Elimination, Buruli Ulcer, Guinea worm Eradication, Onchocerciasis Control Programme, and Polio Eradication. In 2003, the system supported the implementation surveillance of HIV/AIDS/STIs, Tuberculosis, Control of Communicable Diseases in Complex Emergency Situations, Outbreak Alert and Response and Integrated Management of Childhood Illnesses (WHO, 2010).

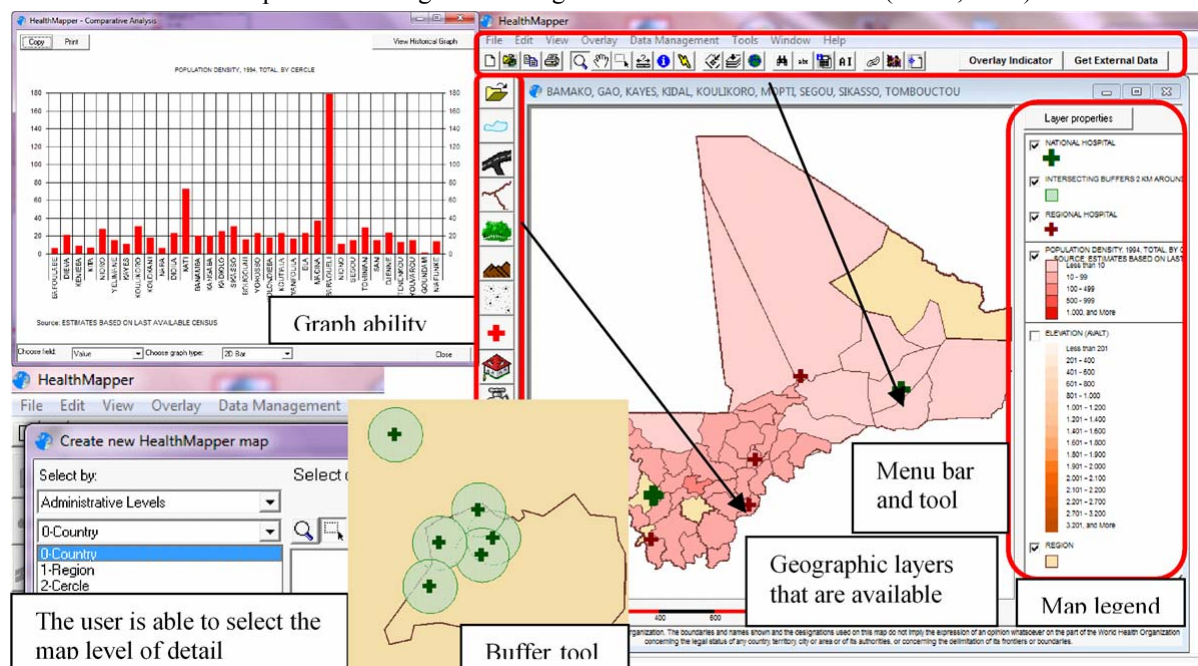


Figure1: The Health Mapper tool interface

After exploring the health mapper tool it was concluded that this tool is a powerful GIS tool. As stated above about its key features, this tool comprises all of the basic GIS functionality that is needed by WHO. This tool is also user friendly; the interface is simple and can be easily understood by a novice user. This tool can achieve its purposes described above.

3.2 Global Health Atlas

WHO has launched the first global online atlas of infectious diseases, a new tool for infectious disease surveillance and control which builds on the features of the Health Mapper. Over 300 indicators for more than 20 infectious diseases of major public health concern are included in the database.

In a single electronic platform, the Atlas is bringing together for analysis and comparison standardized data and statistics for individual diseases and indicators at country, regional, and global levels. The analysis and interpretation of data are further supported through information on demography, socioeconomic conditions, and environmental factors. Maps are used to display data on the prevalence of individual diseases, the location of at-risk populations and vector habitats, and patterns of antimicrobial drug resistance. Such maps offer easy visualization of conditions ranging from the number of villages infected with guinea worm disease, through the monitoring of polio surveillance in a selected country, to the status of resistance to chloroquine throughout the African continent. The global health atlas has three different interfaces for using data, which are explored in the next sections.

3.2.1 Data Query

This interface allows users to browse, view, query, search the contents of the WHO's Communicable Disease global database and output data in reports, charts and maps.

Figure 2: Data query interface

With the help of this web interface one can have the public health data through a list of filter like category (Communicable and non communicable diseases), diseases category, and sometime a subtopic. The user can then retrieved the data according to spatial extent like for the whole world, for a specific continent, WHO region and more specifically for a particular country (fig. 2). After that the user can chose three different output formats i.e. map, chart and a Microsoft excel database file.

3.2.2 Interactive Mapping

Interactive mapping provides a user-friendly mapping interface that allows users to select geographic areas of interest and create maps of diseases, the location of health facilities, schools, roads, geographic features. Interactive features allow users to select layers of additional information to be displayed on maps and the specific areas where interventions have been applied. For the purpose of reports, data can be displayed as charts, graphs, and tables of country-specific statistics. Trends over time can be assessed, and situations in different countries can be compared, either for a single or several diseases.

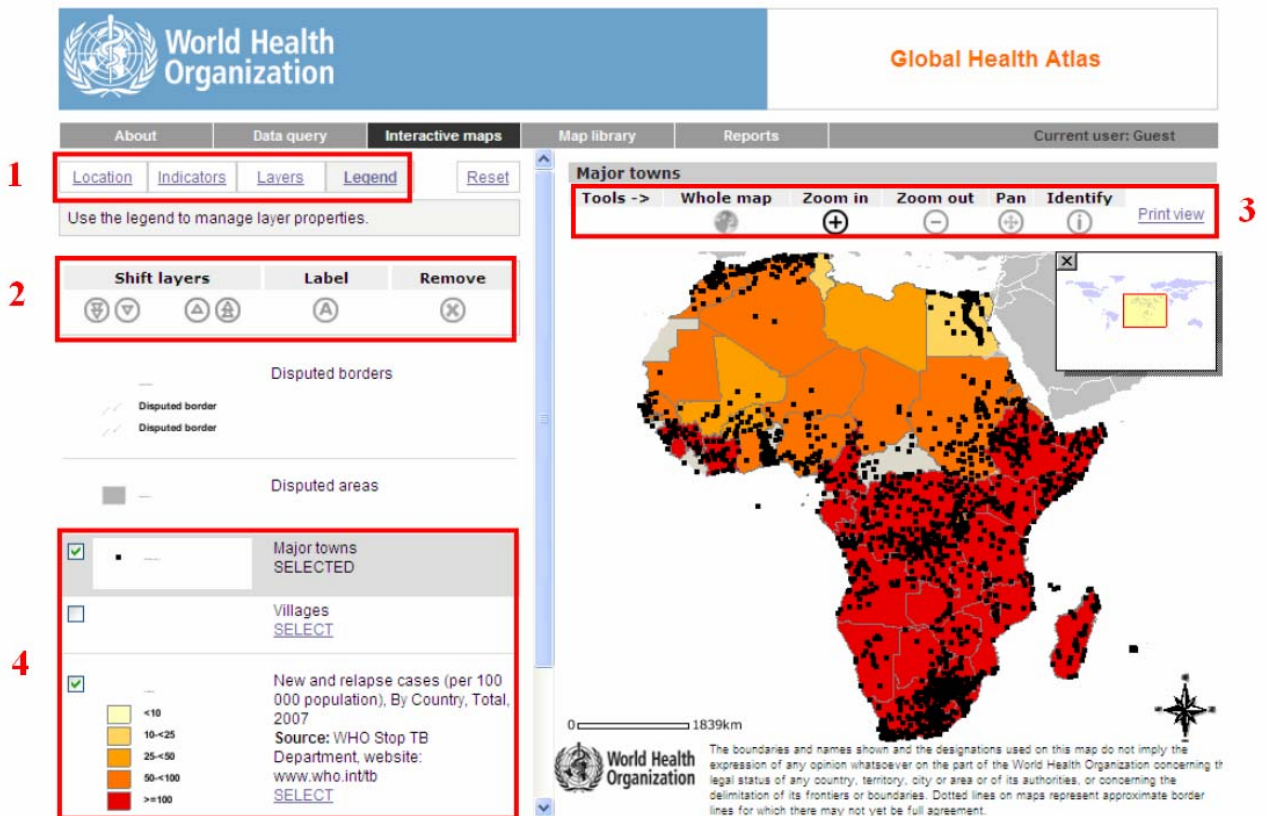


Figure 3: Interactive mapping interface

1. The user can select the location of interest. In the location section, there are three different options a user can choose i.e. the world as whole, location as a continent and location as a country. In the indicator section, a user can choose their maps from a wide range of health indicators. In addition to the health indicator, a user can choose some demography and socio-economic indicators and infrastructure layer for mapping in the layer section. And finally with the legend option the colour and other legend items could be edited.
2. This section of the interface is used to provide labels for the features, removing the labels and moving layers up and down to get an optimum view of the map.
3. This section provides the option for zooming and panning of the map like any other standard GIS application. A user can have the attribute of a feature through the identify option as well.
4. A user can switch on and off the layers with this section of the interface.

3.2.3 Maps and Resources

Maps and resources provide access to the public domain to static maps and related documents, publications and statistics on infectious diseases. The map gallery includes an extensive list of maps on major health topics, maps can be searched by geographic coverage, health topics or keywords.

With all these features, the Global Health Atlas provides a useful new monitoring tool for partners assisting WHO in its efforts to combat infectious diseases. Even those unfamiliar with the technology or new to the field of infectious disease control can readily visualize and analyze the many disease patterns and determinants, trends, and interrelationships displayed in the Atlas.



Figure 4a: Map gallery interface

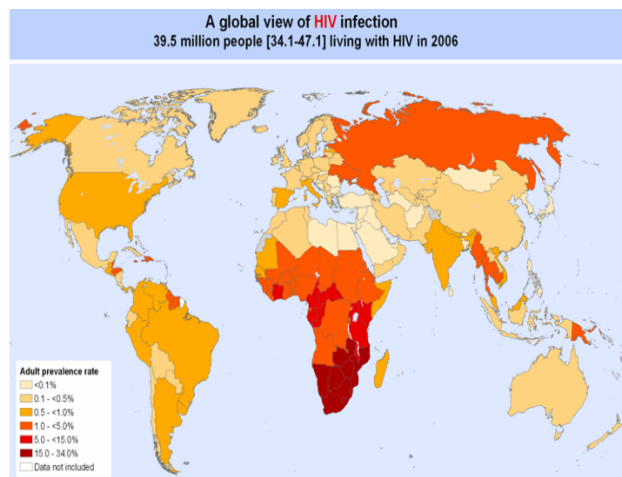


Figure 4b: Output map of the query

3.3 The Remote Field Data Collection Tool

The remote field data collection tool is another tool related to GIS developed by WHO. It is used to support field data collection activities. It uses GPS technology embedded in the hand held devices used to collect field data. The collected GPS data can be downloaded directly into the Health Mapper tool and can be used for further analysis.

WHO has been routinely using Global Positioning Systems (GPS) to map and track infectious diseases at community levels. GPS are now routinely used by village outreach teams for onchocerciasis, guinea worm, African Trypanosomiasis (sleeping sickness), lymphatic filariasis among others. Increasingly such systems are being used during the investigations of disease outbreaks for rapid mapping of cases and deaths and within the context of complex emergency situations they are being used for mapping internally displaced persons, refugee camps and rapid epidemiological assessments.

4 Conclusion

The GIS tools and applications developed by WHO are suitable and even user friendly for non professional users. The tools and applications in the field is to serve not only as support to the surveillance requirements of eradication and control programmes but also as early alert systems for the detection, verification and response to outbreaks of new and previously unknown diseases. WHO's public health and GIS programme has been leading a global partnership in the promotion and implementation of GIS to support decision-making for a wide range of infectious disease and public health programmes.

References

- Al-Shorbaji N. 2006: WHO EMRO's approach for supporting e-health in the Eastern Mediterranean Region. Eastern Mediterranean health journal, 2006, 12(Suppl. 2):S237-52
- Black M et al. 2004: Using GIS to measure physical accessibility to health care. Geneva, World Health Organization, http://www.who.int/kms/initiatives/Ebener_et_al_2004a.pdf [accessed 05.05.2010]
- Cormley EK, McLafferty SL.2002: GIS and public health. New York, Guilford Press, 2002
- World Bank, 2010. The online atlas of the Millennium Development Goals <http://devdata.worldbank.org/atlas-mdg/> [accessed 05.05.2010]
- World Health Organization, 2010. Public Health Mapping and GIS. http://www.who.int/health_mapping/en, [accessed 05.05.2010]

Monitoring of fruit fly (Diptera Tephritidae) and their relation with biotic and abiotic factor in Santa Elena region, Ecuador

Alvaro Cañadas^{1,**}, Viviana Duque², Oscar Calahorrano² and Christian Garófalo²

¹Ecuadoran Agency of Quality Assurance in Agro-quality, AGROCALIDAD, Ed. MAGAP, piso 9, Avd. Eloy Alfaro y Amazonas, Quito (Ecuador). alvaro.canadas@agrocalidad.gov.ec

²Department of Phytosanitary Vigilance, Ecuadorian Agency of Quality Assurance in Agro-quality, AGROCALIDAD, Ed. MAGAP, floor 9, Avd. Eloy Alfaro y Amazonas, Quito (Ecuador), direccion.general@agrocalidad.gov.ec

KEYWORDS:

Economic implications, Invasive species, Geographical Information Systems, Principal Component Analysis, Pest Surveillance

ABSTRACT:

*Monitoring field conditions are the foundation of modern agricultural management in order to improve the efficiency of the data collection procedure. Thus, it is necessary that the Ecuadoran Agency of Quality Assurance in Agro-quality collect environmental data and environmental fluctuations in order to determinate the pattern of Tephritid fruit flies. One of the important areas of fruit production is Santa Elena, located in the western coast. The main objective was to determine the monitoring of fruit flies and their relation with biotic and abiotic factors. The factors used in the methodology adopted for this study was the slope, soil texture and climatological maps as well as 694 traps in a research area of 3466 ha distributed in central, northern, southern and western part of Santa Elena region, and this was subjected to principal component analysis. The results indicated that the minimum–maximum temperature, relative humidity and rainfall were the major climatic factors influencing fly populations. The soil texture did not have any influence on the state of development of caterpillars. Trapping and rearing data indicated that the genus *Ceratitidis* was common during the dry season and reached a peak at the end of the dry season. It attacked both early and mid-season agricultural production. On the other hand, *Anastrepha fraterculus*, *A. serpentina*, *A. punensis*, *A. manihoti*, *A. chichlayae*, *A. pickeli*, *A. obliqua* which were also identified, are native species and were widely distributed. This highlights the high priority of management of fruit flies, which is the main limitation for the export of diverse fruits and vegetables produced in Ecuador.*

** Corresponding author

Fixing Spatial Development Strategies for a City with GIS

Zahid Hasan Siddiquee

GIS & RS Specialist, Institute of Water Modelling, Ministry of Water Resources,
Bangladesh, E-mail: zhs@iwmbd.org

KEYWORDS:

GIS, DSS, MDG, Spatial Strategies, Sustainable Development

ABSTRACT:

Population density and the availability of build able land can be the two most important governing factors in determining the growth direction of a city in spatial context. Land with suitable elevation and areas with low population density can be the preferable options for new development A city can render better services in the long run for people when it is guided by the right kind of spatial development strategies. Here in this study Geographic Information System (GIS) is deployed in order to derive the best possible choices for the city's new development areas. The right spatial planning would help Barisal to meet up the Millennium Development Goals (MDG) in time.

1 Introduction

The study area is Barisal City located in the southern part of Bangladesh. This is small city at the center of the climate change impact zone or Seal Level Rise zone of Bangladesh. In order to make the city livable with special emphasis on sustainability there is no alternative to viable spatial planning. The city under our study is one of the main in the coastal region of Bangladesh. Urban areas are the hubs of economic activities. People come to the city in search of better livelihood and employment, which leads to an increase in the city's population. Gradual increase in the size of population indicates that there is a necessity for extra space to accommodate new growth. The ultimate success of the new growth greatly depends on the spatial strategies. These policies can dictate the spatial extent and direction of the development trend. The following sections will provide an account of these strategies for Barisal City.

2 Input Data

The input data set includes information on projected population in 2030, land levels for deriving digital elevation model and highest flood level with a return period of 80 years. Population data contains the number of people, who will be living on an acre of land in the next twenty years. The average, minimum and maximum population density in the study area is 46, 2 and 158 persons per acre. Land levels are comprised of spot values taken randomly from the mean seal level (z values) in meters. The elevation data are then interpolated into continuous raster surface (DEM). The average, minimum and maximum values are 2.10, 0.14 and 4.77 m. The data on the highest flood levels are also available with a return period of 80 years (2.23 m), this is used to find out the areas without the risk of flooding.

3 Procedures

The diagram below illustrates the processes involved in the study. A number of filtration criteria are applied on the study area with a view to find the desired locations. The goal was to determine the future growth direction of Barisal City. In this regard the first criterion was to find out areas with comparatively low population density. The areas with less people are thought to be the best possible option for ne development. The second filtration criterion was based on the availability of suitable land in terms of favorable elevation. The areas with higher elevation ill require less investment for new infrastructure development as well as provide safe places from natural hazards like flooding. The result of the above two processes are then matched with the flood free zone of the study area which ultimately reveals the potential areas for the future development of Barisal City.

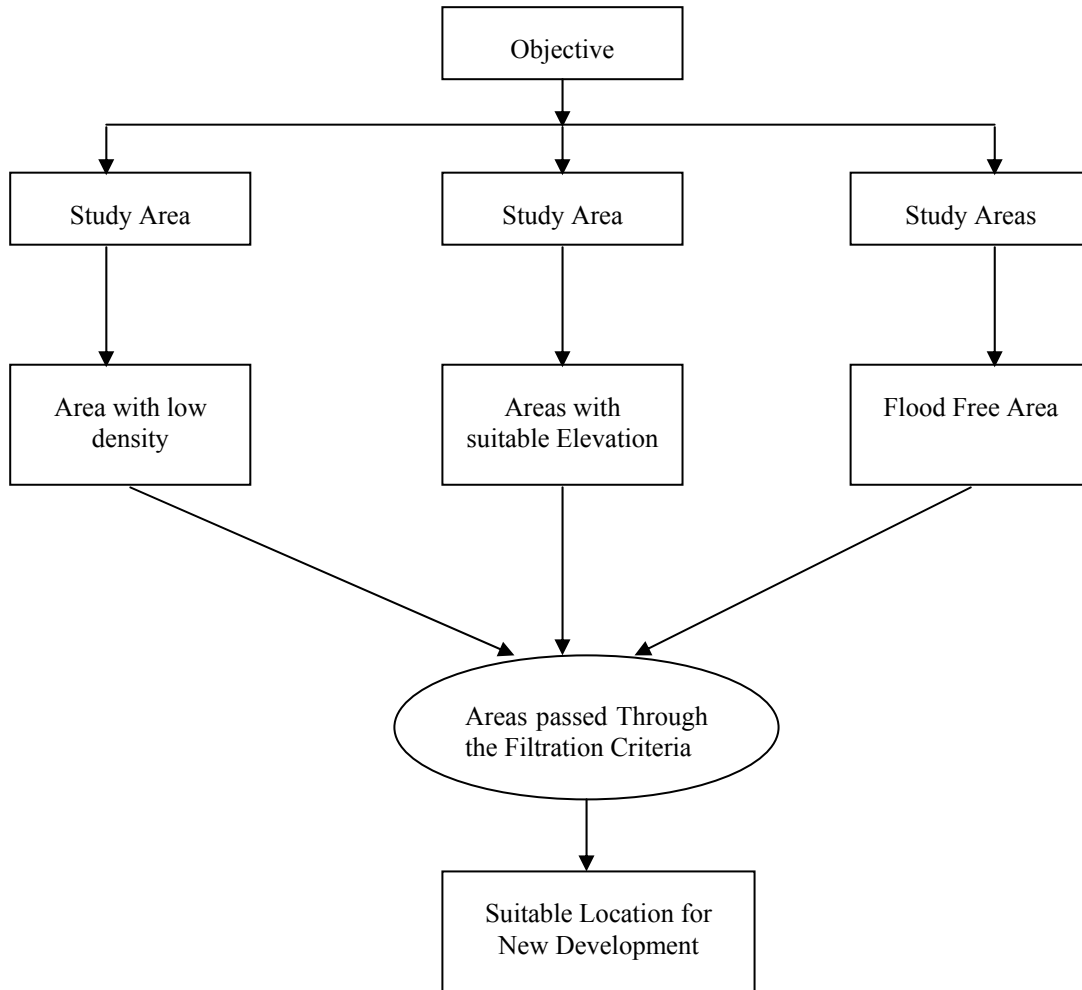


Figure 1: Schematic Diagram of the Procedures followed

3.1.1 Population Density

Objective of this study is to devise a decision support system with a view regarding the future development trend of Barisal City. Areas with low population density are considered as most suitable for new development. The data was in number of population for each administrative unit (wards) in persons per acre.

The vector data set was then converted into raster one for spatial analysis. The rasterized density data was with a pixel size of 20 x 20 m and then reclassified from 1 to 10. Here in this coding 1 means least population density (figure 2) and 10 indicate highest density. Hence the areas having lower recorded values are the more suitable locations.

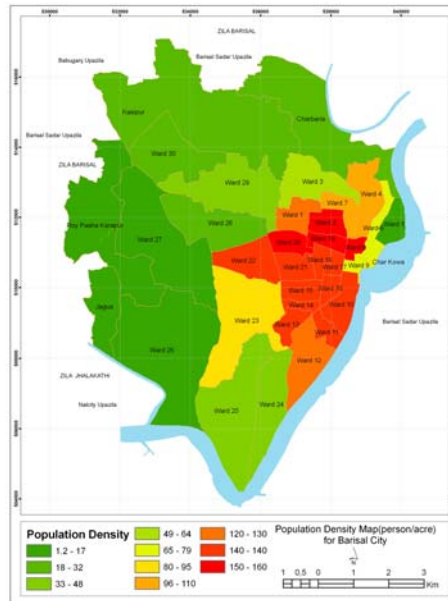


Figure 2: Population density in persons per acre.

3.1.2 DEM

The land levels are collected from the field with an approximate 100 meter interval. The surveyed spot levels are interpolated to derive a raster surface with IDW method. This method is chosen as the undulation of the ground surface is not much higher. This is very clear from range (4.63) of values in the rasterised (20x20m) data that is reclassified from 1 to 10. Here in this classification 1 denotes best location with regards to elevation in color it is represented in red (Fig 3). The preferred area with higher elevation would require less investment to provide new infrastructure services to support the township.

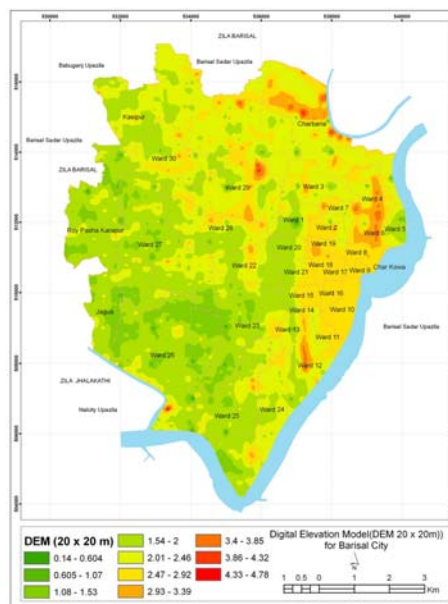


Figure 3: Digital Elevation Model (DEM) of the Study Area

3.1.3 Inundation Map

Flooding is a very common phenomenon in Bangladesh. Flood hazards strike almost every year leaving a devastating impact on both lives and resources. The areas to be selected for expansion of Barisal City must be free from the risk of being flooded. Inundation map (Figure 4) is created considering highest flood levels and

natural drainage systems like rivers and canals. This depicts the areas with the risk of being flooded and flood-free zones. Only the flood-free areas are considered suitable for the new development.

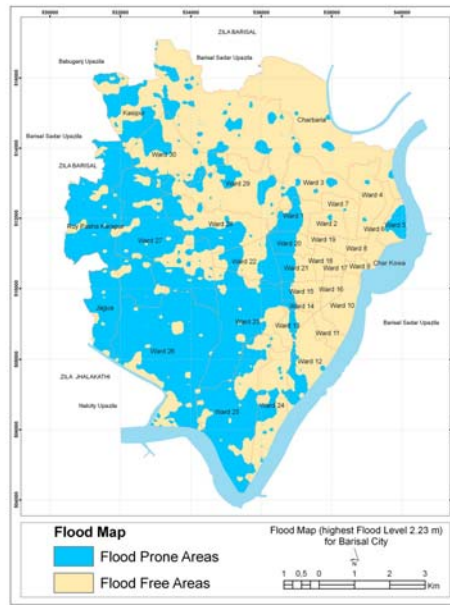


Figure 4. Inundation Map of the Study Area

3.1.4 Composite Map

Here in this section all three criteria are put together in order to find out the most advantageous areas for accommodating the future growth of Barisal City. The first two sets of data were weighted and overlaid in the same scale from 1 to 10 where 1 represents best possible location. This overlaid data was then further matched with flood free zones. This process gives the potential areas for further development of Barisal city in future. This area of interest lies on the upper right corner of the study area which is marked with red circle.

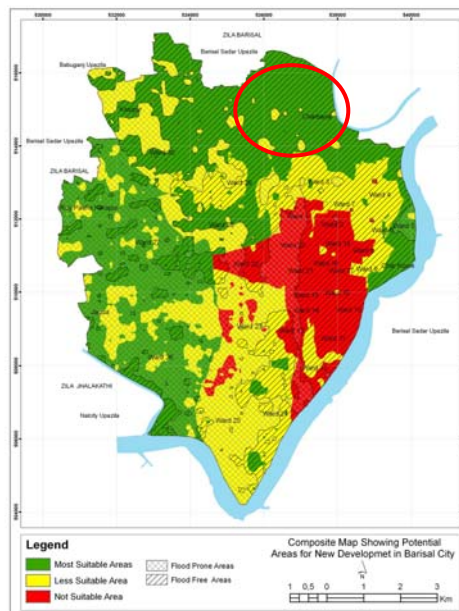


Figure 4: Composite Map for Spatial Development Strategies

4 Conclusion

The approaches used in this study are very simple but highly effective for shaping a city in the long run keeping in pace with Millennium Development Goal. The study would help Barisal City to achieve MDGs no. 7 ‘‘Ensure Environmental Sustainability’’. This target includes integrating the principles of sustainable development into country policies and programmes; provide sustainable access to safe drinking water and basic sanitation; improving the lives of slum dwellers and reduce biodiversity loss. Only three parameters are used for deriving the best possible options. But the consideration of other criteria like land uses could definitely enhance the result of the study. The outcome of this study can be beneficial for the urban planners, city managers, environmentalist and other concerned professionals.

References

- Allen, P.M. and Sanglier, M. (1979) A dynamic model of urban growth: II, *Journal of Social and Biological Systems*, 2(4): 269–278.
- ESRI, 2010. GIS Solutions for Urban and Regional Planning, 380 New York Street, Redlands, CA 92373-8100 , 909-793-2853, USA.
- Geertman, S. and Stillwell, J. (2004) Planning support systems: an inventory of current practice, *Computers, Environment and Urban Systems*, 28(4): 291–310.
- IWM & BETS, 2009. Water supply, sanitation and drainage project in Sylhet & Barisal City. Unpublished Project, Department of Public Health Engineering, Ministry of Ministry of Local Government, Rural Development and Cooperative, Government of Bangladesh, Dhaka, Bangladesh
- Sudhira, H. S et al 2009. *A Spatial Planning Support System for Managing Bangalore’s Urban Sprawl*, Centre for Ecological Sciences and Centre for Sustainable Technologies, Indian Institute of Science, Bangalore, 560 012, India
- Sharifi, M.A. (2003) Integrated Planning and Decision Support Systems for Sustainable Water, Resources Management: Concepts, Potentials and Limitations, Seminar on *Water Resources Management for Sustainable Agricultural Productivity*, organized by Asian Productivity Organization, Lahore.
- Subudhi, A.P. and Maithani, S. (2001) Modelling urban sprawl and future population prediction, using remote sensing and geographical information system techniques. In: Subudhi, A.P., Sokhi, B.S. and Roy, P.S. (eds.) *Remote Sensing and GIS Application in Urban and Regional Studies*, Human Settlement and Analysis Group, Indian Institute of Remote Sensing, Dehradun.
- UDD, 2010. Preparation of Structure Plan, Master Plan and Detailed Area Plan for Barisal Divisional Town, Unpublished Project, Urban Development Directorate, Ministry of Works, Government of Bangladesh, Dhaka, Bangladesh.

Internet based Applications, Open Source Solutions

OpenLayers: Open Geodata in an Open Source Map Browser

Franz-Josef Behr

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
franz-josef.behr@hft-stuttgart.de

KEYWORDS

Internet/Web, Open Source, OpenLayers, Interoperability, OpenStreetMap

ABSTRACT

Geoinformation deployed over the Internet is characterized by a high level of diversification: many different formats, systems, and services, offered by public and commercial suppliers. Interoperability of the services and of data is gaining increased importance.

In this presentation OpenLayers, a freely available mapping client is presented, where the emphasis is explicitly placed on the integration of different services and formats. The organisational framework conditions of this project may be considered to be exemplary with regard to other successful projects of this category. It is demonstrated how data can be loaded, integrated, and visualized. A general tendency is increasing convergence of open data and freely available software tools; this will be demonstrated with the example of OpenStreetMap, a project of high importance for developed and developing countries, and with the integration of Web Services and commercial data layers

Open Source Desktop GIS - Ready to Use for Teaching at Universities?

Dietrich Schröder

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
dietrich.schroeder@hft-stuttgart.de

ABSTRACT:

The focus of this report is on the evaluation of Free and Open Source Desktop GIS for education at Universities. After defining a list of requirements, the systems are introduced and an evaluation matrix is defined.

1 Introduction

Whereas Free and Open Source GIS (FOSGIS) are successfully applied for Internet-GIS by many municipalities and companies since years, the field of desktop GIS is still occupied by proprietary systems. This situation is mirrored in the software products used in education at Stuttgart University of Applied Sciences (SUAS). Open Source projects are widely used for teaching about WebMapping, whereas data management, analysis of spatial data, and map design is still taught on proprietary systems. As in recent years many interesting desktop FOSGIS projects have been initiated, it is worthwhile evaluating these projects and checking if they have matured to the point where they can be used in education as well. For a lot of reasons, e.g. costs of licences, availability of additional student licences, as well as tackling different versions with often surprising changes, FOSGIS may be an interesting alternative to proprietary software. For this study, gvSIG 1.1/1.9 alpha, uDig 1.1, QGIS 0.11.0, openJump 1.2, GRASS 6, and ILWIS open 3.5 have been evaluated. Most of the products have been tested and evaluated on Windows XP, some information about software or functionality which has not been tested was retrieved from manuals, tutorials, or other literature.

2 Defining Requirements

The education in Geoinformation at SUAS is task and practice oriented. That means the students have to work on projects with a defined goal but where the exact workflow to achieve this goal is not specified. So it is up to the students to find the “best” workflow. This task-oriented approach makes it quite easy to exchange the software used and also to shift from classical talk and chalk teaching to a more student-centered approach of learning.

According to this approach, there is a wide range of requirements which a desktop GIS has to fulfil. Many of these requirements are similar to the usage of GIS in a company or consulting agency, in particular those involved in geoprocessing, i.e. analysis of spatial data and the visualization of the results. On the other hand, in education the usability plays a key role, as many students on the bachelor as well as on the master level are newcomers to GIS, so a steep learning curve is desired and the students should not be bothered with the imperfection of the software.

The most important requirements which have been taken into account for evaluation are the following ones:

- Hybrid system. Visualizing and analysis of vector as well as of raster data should be supported.
- Formats. The most important vector and raster formats should be supported for reading and writing, including de-facto standards like ESRI-shape or Autodesk-DXF, as well as OGC standards like GML.
- GDI client. Web Services based on OGC standards like WMS (Web Mapping service), WFS (Web Feature Service), WFS-T (Web Feature Service with Transaction support), WCS (Web Coverage Service), WMC (Web Map Content), CSW (Web Catalogue Service), and SLD (Styled Layer Descriptor) should be integrated, but also de facto standards like ESRI-ArcIMS should not be neglected.
- Database. Support should be available not only for attributive data, but for the storage of geometries using geo-DBMS like PostGIS, MySQL or Oracle Locator.
- Editing. Updating of vector data as well as attributive data should be possible. This includes interactive georeferencing of images for the use of backdrops, basic functionality for digitizing, as well as tools for editing attributive data. The latter includes a field calculator. It would be nice to have some functionality for designing new features, like in a CAD environment.

- Analysis. Basic functionality for vector (intersection, difference, union, buffer, etc.) as well as for raster (proximity, map algebra, zonal score, etc.) should be provided. In the evaluation of this item, the possible options offered have to be considered as well.
- Map design. The results of analysis have to be presented in appealing maps. So a style editor should be available for defining complex symbology, thematic maps based on classification of attributes should be possible and the design of a final map layout including legend, scale etc. has to be supported.
- Usability. The system should be easy to learn, icons and menus should be self-explanatory for a medium trained GIS user and the workflows to achieve the goals should not include too many workarounds.
- Localization: the training at SUAS is in German as well as in English. Thus at least these two languages should be supported. In the Evaluation matrix Russian was added as a third language.
- Documentation. For achieving the tasks by a student-driven and self-conducted learning approach, good documentation is very important. Thus manuals as well as tutorials should be available, preferably offline.

3 Overview and Evaluation of Desktop Open Source GIS

3.1 gvSIG

The gvSIG project [2] was initiated by the regional government of Valenciana in Spain in 2003. The software is developed by the Spanish company IVER Tecnologías together with the Jaume I University of Castellón based on Java. Though the original language is Spanish, localizations for many languages are available. The system is targeted at professional users of geographic information from the public sector, private industry, sciences and education. gvSIG was developed from the base as a GDI client, thus several OGC services are supported natively. It supports a wide range of data formats and databases, including proprietary ones. It offers some vector analysis functionality, which is handy to use.

Very well integrated is the plug-in SEXTANTE [3], which offers more than 210 tools for raster as well as vector analysis. These modules can be assembled for more sophisticated workflows using an integrated model builder. Additional plug-ins for network analysis as well as for 3D visualization are available.

3.2 uDig

The software uDig [4] has been developed with a similar intention as gvSIG, i.e. as a multipurpose desktop GIS. The development of uDig (User-friendly Desktop Internet GIS) is promoted by the Canadian company Refrations Research. The project started in 2004 and is programmed in Java as an Eclipse-based Rich Client Platform application. Up to now it supports only a small number of vector formats but quite a few raster formats as well as databases are supported. Most of the OGC Web specifications are supported as well. In the uDig core, no geoprocessing is implemented. This is shifted to other OS projects developed as RCPs, which can be integrated easily to the uDig framework. Worth mentioning is in particular JGrass, a Java-based reimplementation of the famous GRASS modules as well as a first alpha version of SEXTANTE for uDig. But some of these plug-ins are not very stable and are tricky to use.

3.3 Quantum GIS (QGIS)

This Open Source project started in 2002; a first Window version is available since 2005 [5]. The development is mainly driven by a team at ETH Zurich, Switzerland. C++ is used as programming language. QGIS is based on the open source library GDAL/OGR, which is used for the support of import and export of different raster and vector formats. It is not intended to add any geoprocessing functionality to QGIS, as there is a plug-in for accessing the very powerful GRASS modules. Thus QGIS often is used as the front-end for GRASS. Unfortunately, this means that for analysis also the not very handy concepts of GRASS data management have to be used. Besides the GRASS plug-in, there are others offering additional functionality like WFS client, GPS support, georeferencing of images, etc.

3.4 openJUMP

Originally initiated by the Canadian Province administration of British Columbia and developed by the Canadian company Vivid, the original project was shifted to the Open Source with a first release in 2006 [6]. As vector formats, mainly shape and GML is supported, other formats like DXF via plug-ins. It supports a number of raster formats. The philosophy of the openJUMP project is not to provide a multi-purpose GIS, but like QGIS to offer a platform for integrating plug-ins with additional functionality. E.g. database support can be integrated using plug-ins, as well as OGC Web Services and re-projection or raster analysis.

3.5 SAGA

SAGA [7]- short hand for "System for Automated Geoscientific Analyses" - is a hybrid GIS software programmed in C++. The development is led by a team at the University of Hamburg, Germany, and started in 2004. Its aim is not to develop a multi-purpose GIS but a framework for scientific spatial analysis. Thus it offers a lot of different geoprocessing tools. As the focus here is on the big variety of tools, the graphical user interface is not very handy. Also no database support or OGC Webservice integration is offered. The documentation as well as the GUI is only available in English.

3.6 GRASS

Started as early as in 1983, the Geographic Resources Analysis Support System (GRASS) [8] is the most complex Open Source-based desktop GIS. The strength of this GIS lies in its modular structure and its utility for raster processing and modelling. GRASS supports an extensive range of raster and vector formats through the binding to GDAL/OGR libraries. Attributes are managed in dbf files or SQL-based DBMS. The system is capable of visualizing 3D vector graphics data and voxel volumes. The recent GRASS release introduces a new topological 2D/3D vector engine and support for vector network analysis. There are over 200 core GRASS modules included in the GRASS distribution, and over 100 add-on modules created by users and offered on the GRASS web site. The GRASS Development Team is a multi-national group consisting of developers at numerous locations.

3.7 ILWIS open

The Integrated Land and Water Information System (ILWIS) [9] is a GIS and Remote Sensing software, developed by ITC (International Institute for Geo-Information Science and Earth Observation, The Netherlands) started in 1988. ILWIS comprises a complete package of image processing, spatial analysis and digital mapping. It is easy to learn and use; it has full on-line help, extensive tutorials for direct use in courses and 25 case studies of various disciplines. As of 2007, ILWIS software has become an OS project under the 52°North initiative, called ILWIS Open. It has been originally programmed in C; the open source version has switched to a MS Visual Studio (.NET) project. It is the only system evaluated for this report which is still not component based.

1 Evaluation matrix

FOSGIS/ rating	formats vector	formats raster	database	localization	GDI client	georef.	edit geometry	design tools	edit attribute	projectio n	map layo ut	usability/ stability	analysis vector	analysis raster	developer/ community
gvSIG (with SEXTANTE)	shp (r/w), GML (r/w), KML (r/w), DXF (r/w), DWG, DGN	ECW, MrSID, JPg (r/w), jp2 (r/w), TIFF, geoTIFF (r/w), PNG (r/w), GIF, bmp (r/w), ...	PostGIS (r/w), MySQL (r/w), Oracle (r/w) ArcSDE	Spanish English ++ German + Russian ++	WFS, WMS, WCS, WMC, CSW, ArcIMS					EPSG proj4			core and sextante	via Sextante*	regional government Valenciana, Spain
<i>rating</i>	++	+	++	o	++	++	+	--	++	+	+	+	+	+	
uDig (with JGRASS)	shp (r/w), GML, WKT	TIF, GeoTIFF, JPEG, GIF, PNG	PostGIS, Oracle, MySQL, DB2, ArcSDE	English German + Russian --	WFS, WFS-T, WMS, WMC, SLD					WKT EPSG			Axios	GRASS* JGrass*, Horton* on mapsets	Refraction Research, Canada
<i>rating</i>	o	o	+	+	+	--	+	--	-	+	o	o	-	o	
QGIS (GRASS)	shp, mif, gml, kml	GDAL	PostGIS, SQLite**	English German ++ Russian ++	WMS	*				proj4			GRASS	GRASS*	ETH Zürich, Switzerland
<i>rating</i>	o	++	o	++	-	++	+	--	-	++	o	+	-	o	
openJump	shp (r/w), GML (r/w), WKT (r/w), DXF*, MIF*	TIF, PNG, GIF, JPEG, ECW	PostGIS, Oracle, ArcSDE	English German + Russian o	WMS, WFS*, WFS- T*, WPS*					CTS*			core	Pirol*	several teams
<i>rating</i>	o	o	o	o	o	--	+	+	++	o	-	o	+	o	
SAGA	shp (r/w),	GDAL		English						geo- TRANS			core	core	Uni Hamburg,

	E00														Germany
<i>rating</i>	-	+	--	-	--	++	o	--	o	++	o	-	++	++	
GRASS	OGR (r/w)	GDAL	PostGIS (r/w), mySQL**, SQLite**	English German ++ Russian ++	WMS, WFS					proj4			modules	modules	Centre for Alpine Ecology Trento, Italy
<i>rating</i>	+	+	o	++	o	o	+	-	+	+	+	o	++	++	
ILWIS open	shp (r/w), E00 (r/w), DXF (r/w)	TIF, GIF, BMP, ... GDAL		English	WMS								core	core	52° North, Germany
<i>rating</i>	o	++	--	-	-	++	o	--	+	o	+	+	o	++	

*Plugins/Extensions not included in the core product

** only attributes

4 Conclusions

As the evaluation shows, a large variety of Open Source desktop GIS are available. The most promising projects as multi-purpose GIS are gvSIG and uDig. With the plug-ins available, most of the task-oriented projects used in GIS education at SUAS could be solved with both of the systems. With some restrictions this holds for openJump as well. SAGA and ILWIS are more focused on scientific applications, whereas QGIS/GRASS offer on the one hand a lot of functionality but on the other hand the concept of GRASS is not very easy to use. Unfortunately, for all the systems, there is still some missing functionality in comparison to the proprietary systems being used now. Even if the functionality is available, there are still less options offered. Some systems have still some problems with stability. A major drawback of most of the systems is the missing or incomplete documentation.

As the development for most of the systems is on-going at a high pace, it is worthwhile to observe the further progress. As there is only a small step missing, may be in the near future these systems can replace proprietary systems in GIS education.

References

- [1] The Free Software Definition. Free Software Foundation. <http://www.fsf.org/licensing/essays/free-sw.html> (8.1.2009)
- [2] Conselleria d'Infraestructures i Transport: gvSIG. <http://www.gvsig.gva.es/> (9.1.2009)
- [3] More about SEXTANTE. (2008). www.sextantegis.com/ (9.1.2009)
- [4] udig. A GIS Framework for Eclipse. <http://udig.refrains.net/> (9.1.2009)
- [5] Quantum GIS. <http://qgis.org/> (9.1.2009)
- [6] What is OpenJump? (2008). <http://openjump.org> (3.1.2009)
- [7] J. Böhner (2008): Saga 2.0. <http://www.saga-gis.org> (9.1.2009)
- [8] GRASS Development Team (2000): GRASS 6.3. <http://netmirror.org/mirror/grass.itc.it/index.php> (13.1.2009)
- [9] Rob Lemmens: ILWIS Community. http://52north.org/index.php?option=com_content&view=category&layout=blog&id=16&Itemid=61. (12.1.2009)

Geocoding: Fundamentals, Techniques, Commercial and Open Services

Franz-Josef Behr

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany

KEYWORDS:

Spatial Data Infrastructure, Interoperability, Web services, Address data, Geocoding, Open Geodata

ABSTRACT:

Referencing information and features to geographic locations is an essential component of daily live. In newspapers and television news we are accustomed to see maps illustrating the location of events in politics, sports, or natural hazards. The process of assigning location to such features is called georeferencing.

Postal addresses, a specific way of locational description, are the essential means by which people express their location in the real world. Also many features in modern societies in administration and business are related to addresses which are used as unique identifiers to geographical locations, often expressed by a pair of longitude and latitude values. Although addresses are data containing location, they do not contain coordinates. Therefore an important feature in many applications especially in Geographic Information Systems (GIS) is the capability to locate addresses, i. e. to geocode to address level.

During the last few years there has been also a significant progress to geocode addresses by commercial Internet mapping application programming interfaces (APIs) and Internet services. Recent advances, mainly in the field of Internet technology, have encouraged people to develop and to integrate addressing data models from variety of addressing system. It has made geocoding to be more popular, and using such online tools, the geocoded address can be retrieved and displayed. The process of geocoding is not perceived as an individual service by the user but as a standard function of WebMapping.

In the beginning of this paper, fundamentals of address schemas and of geocoding approaches are introduced. The process of parsing addresses and matching with reference datasets is explained. Phonetic algorithms including Soundex and Kölner Phonetics, Levenshtein distance are explained in this context.

Generally, geocoding can be performed effectively on the basis of point locations of buildings. If such data is not available, the position can be derived by interpolation along street axes as available in navigation data. These methods are discussed including their pros and cons.

Some open solutions available on the web are reviewed and compared concerning their coverage, the APIs offered, and the response formats:

The paper concludes by reviewing two open, participatory, community oriented projects: one is an open geocoding service (opengeocoding.org), the other an open geodata project: (openaddresses.org). Both projects rely on data derived from volunteer work.

1 Introduction

1.1 Motivation

One particular data set for a spatial information infrastructure is address data, the base of many administrative, statistical and economic processes in every country (Lind 2010). There are a variety of applications relying on address information like marketing, planning, emergency services, route planning, crime mapping, environmental health studies etc. During the last few years there has been also a significant progress to geocode addresses by commercial Internet mapping application programming interfaces (APIs) and Internet services, mainly based on street network data collected for navigation systems. The development of technology, mainly in the field of Internet technology, has encouraged people to develop and to integrate addressing data models from variety of addressing system. It has made geocoding to be more popular and online mapping tools have popularized the concept of using an address as an initial map navigation tool. Using such online tools, the geocoded address can be retrieved and displayed. The process of geocoding is not perceived as an individual service by the user but as initial step for web mapping (Gruber 2006).

While in few countries such data is already provided by state agencies or offered by commercial vendors, in most of the developing countries such data is not available at all (Vivas 2010).

Based on such street data and address ranges, online mapping tools offer since a few years geocoding services, but limited to developed countries. Today, when the Internet has revolutionized the world of Information Technology, users from developing countries should not be left behind: Even in less developed countries addresses should be available as well as a geocoding application to provide locational information (i.e. coordinates) for postal addresses. Such a geocoding application will not only integrate open data with open source code, but will have also its own database to support the demand of users in developed and developing countries.

1.2 Background

1.2.1 Geospatial features and footprints

Georeferencing deals with geospatial features. Such features can be natural physical features, e.g., lake, rivers, or mountains, but geospatial features can be also man-made, physically existing like cities, roads, buildings, or virtual features, like postal code areas, regions, provinces, and countries, or statistical data.

The (geo-) referencing of geospatial features by placenames is, as Hill (2006) denotes, the most commonly used form of referencing a geographic location in human communication. Such placenames (in a strict way we should say: feature names) are used in many circumstances, in daily conversation, titles of books and articles, but also in digital information items, like news feeds, catalogue entries, HTML pages, or messages.

A placename (toponym) can be a single term (the proper name), but in many cases it is extended by additional modifiers (Hill 2006):

- Administrative modifier: Administrative modifiers (like county, state, country) setup a hierarchy of administrative units for different administrative levels. Examples for such modifiers are “Washington D.C.” or “Monte Carlo, Principauté de Monaco”. In some case of ambiguities a city can be identified by administrative modifiers like county or state.
- Locational modifier: A locational modifier helps to identify a certain location by adding an additional placename. This type of modifier is often used in conjunction with city names, like “Stratford-upon-Avon” or “Rothenburg ob der Tauber”.
- Place type or functional modifier. Toponyms are often additionally labeled by a generic part indicating the type of location, e.g., “Lake Constance” or “Mount Titlis”, “New York City”. Quite often however, these terms are implicitly assumed (e.g., “Bavaria” instead of “State of Bavaria”, or “New York, New York”).

For many purposes geospatial features are categorized according to specific fields of application. Some generic categories can be mentioned.

- Topography, representing the terrain, lakes, rivers, etc.;
- Planimetry, like parcels, buildings, and other man-made features;
- Roads, comprising different road types;
- Thematic data, like environmental measurements, or social data.

In some application fields, like in water or electricity supply, we can even find hundreds of different categories.

Geospatial features have a geospatial footprint, e.g., coordinates, which can be categorized (following Open Geospatial Consortium 2006) as Point, LineString, LinearRing, Polygon, and collections of these primitives (MultiPoint, MultiLineString, MultiPolygon, MultiGeometry). The geospatial footprint becomes especially important if maps or computerized representations of geospatial come into play.

Because the footprint of geospatial features can be very detailed and complex, it is often simplified to a generalized form, e. g., a point such as a centroid with latitude and longitude value, or a minimum bounded rectangle (MBR) as it is provided in the GML standard instead of the real geometry. Such representatives can be efficiently stored and evaluated for information retrieval even if the DBMS used is not capable of spatial operations (Hill 2006:88). In addition, they are more easily to obtain, nowadays even for layman using Global Positional System (GPS) sensors. The coordinates are related to a spatial reference system. In most cases we use two-dimensional coordinates, in a generic way called easting and northing. In many cases, especially in context with Internet based applications and the GPS geographical coordinates, e.g., longitude and latitude values are used, in form of decimal degrees. Besides this notation, longitude and latitude values are sometimes traditionally noted in the form of degrees, minutes, seconds.

1.2.2 Gazetteers

Collections of placenames are called *gazetteers*. Typically each entry consists of three parts: the placename, its type / category, and its geospatial footprint. Gazetteers are used to provide the translation (i.e. geocoding) of placenames to spatial footprints and for the reverse process, finding placenames based on coordinates (i.e. reverse geocoding). Gazetteers are set up for many purposes, on local, regional, and world-wide level. Geocoding of placenames can use one or more gazetteers if corresponding interfaces to access their data are offered.

1.2.3 Representing address data

Postal addresses are one often used way of specifying locations. The INSPIRE Directive (INSPIRE 2007) defines this kind of the spatial data theme Addresses as the “Location of properties based on address identifiers, usually by road name, house number, postal code.”

Several approaches exist for the representation of postal address data:

- Point locations: Geocoding relying on individual point locations yields the highest accuracy. In several countries datasets are already offered by public agencies or private companies, in most cases available for high fees only. Addresses can be directly matched to the addresses of the reference data set.
- Street segment interpolation: If no individual point locations are available, the street network of a municipality can modeled based on the center lines for named streets. Each line segment has a start and an end node providing a sense of direction for a named street. Address ranges are assigned for each side (left and right) of the street segment. Typically, it is assumed the even and odd house numbers are on different sides of the street. For many developed countries, such data is already available.
- Addresses firstly have to be matched with the street names of the reference data set. If the match is successful, the house number is interpolated spatially according to address ranges applying the even/odd assumption (cf. Fig. 1). Additionally a mostly fixed offset from the street center line is added. It is evident that the interpolation and the fixed offset can have significant impact on the accuracy of the resulting coordinates. In reality, houses can be spaced unequally, and the offset can vary heavily (cf. Fig. 1). Additional errors can be introduced by missing street segments or overlapping streets (Ahlers & Boll 2009).
- Areal features: Areal features, like countries, states, provinces, or postal code areas are represented by a point. Address geocoding is based solely on this information.

1.3 Related Applications

Based on these approaches, a large number of geocoding and reverse-geocoding applications are available (<http://groups.google.com/group/Google-Maps-API/web/resources-non-google-geocoders>). However, by analyzing such an enumeration it can be noted that most of these services are offered mainly for industrialized countries like USA, Canada, Japan, and the EU states.

Many Web applications, so-called mash-ups, have been implemented based on commercial geocoders. Majewski (2006) explains the geocoding features of a widely used geocoder, Google’s Geocoding Application Programming Interface (API). Addresses can be submitted as a single string and are parsed on server side, expanding abbreviations. The geocoder response uses JSON (see Wagner et al. 2009), KML or xAL format (explained below). Generally, developers are encouraged to build their own client-side caches. Because geocoding via HTTP is supported, it is possible to use AJAX technology (Garrett 2005) on client-side or to call the service from server side applications.

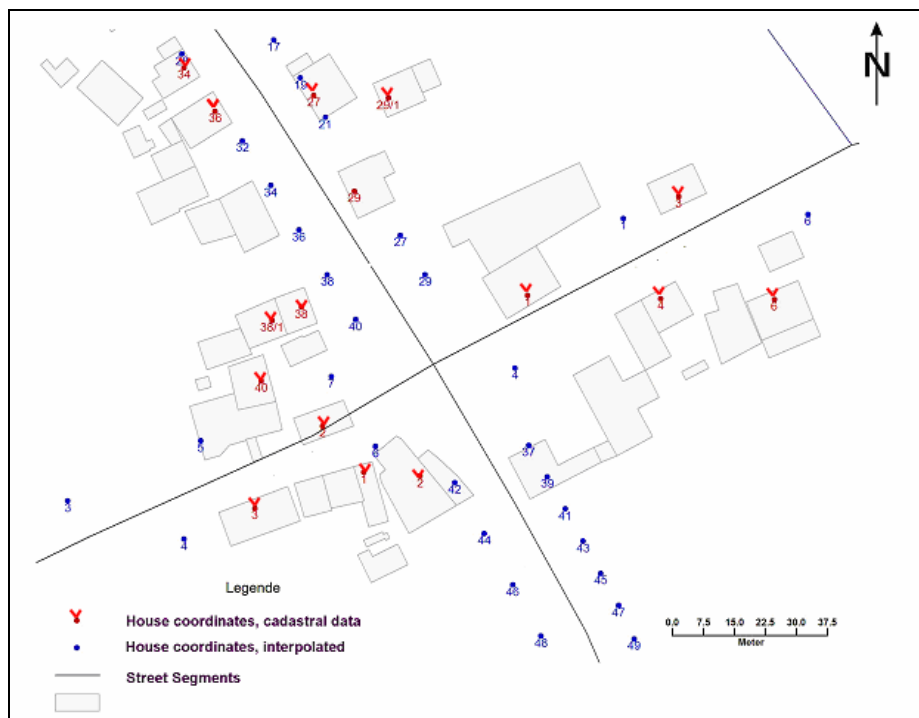


Figure 1: True address location, based on cadastral data, and interpolated address locations (Source: Schäfer & Kunz 2008, reproduced by permission of the authors).

The Geocoding API of Yahoo! Maps supports the Representational State Transfer approach (REST, Richardson & Ruby 2007). The result is delivered in XML format. Similar to Google this service also provides a precision response about the result of the geocoding process. Theurer (2005), for example, describes a mash-up using Yahoo!Maps which offers geocoding addresses, and visualization with additional information layers. The mash-up uses REST to call the Yahoo! API.

Both services have severe limitations concerning the frequent usage. Yahoo's geocoding service is limited to 5,000 queries per IP address per day; Google's service is limited to 50,000 requests per day per API key which is related to a specific domain name.

TeleAtlas, a leading data provider for many commercial mapping services offers its own geocoding service named EZ-Locate (<http://www.geocode.com/>). This service provides fee-based access to the Tele Atlas address database and geocoding technology using web browser, API based requests, or a stand-alone PC-based client.

All these approaches are characterized by their binding to commercial vendors and the data sets offered. In contrary, there are also freely available services. Geonames.org offers a GUI based geocoding and mapping service for country and place names as well as postal codes based on freely available data. Geocoding is also supported by REST web services. Since April 2008, the originally free service has a 50'000 credits limit per day and IP address. For each geocoding service, 1 – 4 credits are needed.

Geocoder.us (<http://geocoder.us/>) is basically a Perl frontend to the TIGER/Line files from the US Census Bureau to geocode US addresses or street intersections. Here it is evident that the data are based on governmental street data provided for the whole country.

Universal Address System, developed by NAC Geographic Products Inc. (<http://www.nacgeo.com/>), is designed to provide a highly efficient unified representation of an address, postal code, area code, geographic coordinates and map grid coordinates. A geocoding service is offered as well.

There are also several free projects of geocoding applications. Geokit (<http://geokit.rubyforge.org/>) is a Ruby on Rails plugin for mapping applications geocoding, location finders, and distance calculation. For geocoding Geokit provides a uniform interface for other geocoders (Google, Yahoo, Geocoder.us.), and IP-based geocoding. Another geocoding library for Ruby as a wrapper for Yahoo!Maps and Geocoder.us is available from <http://geocoder.rubyforge.org/>. Two further, community oriented application are described in more detail in the sub-sections below.

1.4 Geocoding Accuracy

Geographic locations are generally available with a limited accuracy and, in some cases, with some vagueness related to the place name (e.g., Southern France, Bavarian Alps). Such uncertainty can be dealt by the fuzzy-set approach, which is out of scope of this paper. Nevertheless it has to be noticed that the limited accuracy has impacts for the results of a geocoding process and all further analysis steps (Grubestic & Murray 2004). The final accuracy is influenced by the accuracy and errors in the reference data and by the errors introduced by the geocoding processes due to approximation and interpolation.

Several researchers present results of accuracy assessment. Ratcliffe (2001) compared TIGER-type geocoded address data in relation to cadastral and census areal units. As a result from a study of over 20000 addresses in Sydney, Australia, he detects that 5–7.5% are misallocated to census tracts, and that more than 50% may be geocoded within a different land parcel.

Davis et al. (2003) discussed how geocoding can be implemented over incomplete and possibly inaccurate addressing data (which is the prevailing situation in many municipalities in developing countries) and how users can benefit from the results of such a process.

In the context of public health research Cayo & Talbot (2003) determined positional errors for 3,000 residential addresses using the distance between each geocoded point and its true location as determined with aerial imagery. Error was found to increase as population density decreased. The mean error for rural areas was 614 m, 143 m for suburban areas and 58 m in the urban area.

Nicoara (2005) investigated the lack of standardization in addressing, differing ways of defining location, ownership and property definitions, the complexities of multi-source location and data address maintenance, differing address table schemas and models, and the variety of algorithms used. Borges et al. (2007) examined the automatic process of recognition, extracting, and geocoding of addresses in Web pages. They report successful geocoding (for city granularity) up to 77%.

Stark (2008) reports an accuracy assessment of the geocoding engines of Google Maps and Microsoft Virtual Earth. 90,000 addresses of the Canton of Solothurn were used to test the geocoding accuracy; official cadastral data – assumed as true location – were used as reference. The Google Maps geocoder returned 93% of the addresses with an accuracy level of 6 (street level accuracy). Only 53% of these addresses have a distance less or equal to 20m from the true location. With Microsoft Virtual Earth, 86% of the addresses were classified as a good match; in this case, only 41% show a distance lower or equal to 20m compared to the reference data.

Ahlers & Boll (2009) report their approach of combination of two of the most widely used geocoders for obtaining optimal results. Strength of single services can be used, while weaknesses can be avoided.

1.5 Address Standards

If placenames consist of postal addresses, the diversity of address formats come into play. Interoperability problems have encouraged many organizations and developers to create specifications for name and address data (Cover 2008). The Organization for the Advancement of Structured Information Standards (OASIS, <http://www.oasis.org>) has quite successfully developed the extensible Customer Information Quality Specifications Version 3.0 (CIG) which includes formal definitions for Name (xNL), Address (xAL), Name and Address (xNAL) and Party (xPIL). One of them, xAL, is used as part of Google Gecoder's response. Based on xNL and xAL specifications, xAL uses hierarchical nested XML elements to encode address information and is also available as one of OpenGeocoding's output options (section 3.6). Besides these general address formats, local developments are found. The Government of Victoria (2006), for example, published its own address standard.

Also international boards are involved in case of standardization. Schulzrinne (2006) proposed a Request for Comments to IETF. Information about the country, administrative units such as states, provinces, and cities, as well as street addresses, postal community names, and building information is suggested as part of the Dynamic Host Configuration Protocol. The option allows multiple renditions of the same address in different scripts and languages. Issues related to an international address standard, its relevance for postal and Spatial Data Infrastructures and other topics were part of ISO/TC 211's agenda during a workshop in Copenhagen end of May 2008 (http://www.isotc211.org/Address/Copenhagen_Address_Workshop/index.htm).

Addresses are also treated in INSPIRE Directive Annex 1. INSPIRE (2010) describes the INSPIRE Data Specification on Addresses

Addresses and geocoding are also part of the standardization work of the Open Geospatial Consortium (OGC). Its OpenGIS Location Services (OpenLS) document⁶ describes an abstract, standardized set of interfaces, protocols and data types for implementations of Location Based Services (LBS) including geocoding and reverse geocoding.

2 The Geocoding Process

The process of geocoding generally includes several steps:

1. Parsing and structuring the address / locational description;
2. Matching process, including quality assessment;
3. Proliferation of the results.

A typical input for this process is a semi-structured addresses that contains single items for address components such as street, postcode, and city name.

A formal concept of postal addresses is given by Borges et al. (2007). The basic address supplies a street type, its name, and a house number. The second, optional part provides a complement to the basic address, including neighborhood name. The third part, called location, consists of three identifiers applicable in urban context: postal code, phone number, and/or city/state. An address can be complete (having all necessary parts), incomplete (having the basic part plus one of the identifiers), or partial (consisting only of one of the three identifiers). It is worth noting that there is some redundancy in postal addresses. Hence the postal code can give additional support if a city name is ambiguous.

A similar concept is given by INSPIRE (2010:7). The address must be associated with a number of “address components” that define its location within a certain geographic area. Each of the address components represents a spatial identifier as for example the name of a road, district, postcode, municipality, region or country. Remarkable, an address has besides the geographic position which enables an application to locate the address spatially a “locator”, e.g. a unique identifier (e.g., address number and / or name) distinguishing it from the neighbour addresses and temporal attributes.

Language and cultural specific ordering of the parts has to be taken into account while parsing an address (“Geo-parsing”) to convert this semi-structured address to a structured address according to the structure required by the matching process.

According to Mikheev et al. (1999, cited by Clough 2005) simple list lookup for place names already can perform well, but during the matching process some obstacles may occur (Ding et al. 2000):

- Alteration of placenames over the time: For example “Saint Petersburg” formerly was called “Leningrad” and “Petrograd”. People might use different names for the same location.
- Aliasing: Different names are commonly used for the same location on a regional level and in different languages as well. For example “München” is called “Munich” in English, and “Monaco” in Italian.
- Ambiguity: A placename refers to locations in different countries. For example “Saint Petersburg” exists in Russia, and in the USA in Florida, and Pennsylvania. Such a placename without further qualifier is inherently ambiguous; a ranking can be done according to the probability that a certain location is the right one.
- Vernacular naming: “Hyères” in France is called “Iero” or “Ieras”, depending on the Provençal dialect, spoken in the south of France.
- Multilingual naming: Address matching becomes even more difficult in regions that are bi- or multilingual. I.e. in Switzerland there are cities where street names exist both in German and French. Additionally the order of addressing parts are altered according to the language used, e.g., while the German version of an address is street name, house number. the French way is house number, followed by street name.
- Spelling errors: Spelling may not be equal to the one in the reference data-set. Missing accents have to be considered as well as diacritical marks,
- Omissions: Quite often in daily practice, placename modifiers are omitted. Such anticipations can cause additional efforts if a geocoding application has to support different languages.

⁶ <http://www.opengeospatial.org/standards/ols>

Such particularities and ambiguities need to be taken into consideration when geocoding addresses by applying specific string matching algorithms. Sometimes spelling errors or different variations in spelling of an address may occur and the geocoding mechanism must deal with the problem and decide whether the textual location description can be matched to the reference data-set at all or whether some minor or major changes in the spelling need to take place. For the matching entries in the gazetteer, the matching quality can be expressed, for example as an indicator from 0 to 100.

Geocoding addresses is generally not perfectly successful due to spelling differences or errors either in the reference data-set or in the address itself. Unmatched addresses or matches on a lower level of granularity furthermore decrease the accuracy of the matching process, e.g., if resolution does not go down to building level but only to street, postal code or even place-name or country level,

Further challenges in address-parsing and validating are discussed in Ahlers & Boll (2008).

2.1 Granularity, Accuracy, and Precision

While assessing the results of a geocoding process, it is necessary to distinguish several, related terms which are used varyingly by different authors and services.

Generally, the geocoding can be achieved on different levels of spatial granularity. It can be performed on base of a street address, or on a lower level of granularity, like city or country. The number of levels of granularity can be different, depending on the geocoding engine used. Ahlers & Boll (2008) introduce five classes in terms of spatial granularity (country, region, city, street, building). Google's geocoder report one of 10 different levels of granularity⁷ (0 = unknown location, 1 = country level, 2 = region (state, province, prefecture, etc.), 3 = sub-region (county, municipality, etc.), 4 = town (city, village), 5 = postal code (zip code), 6 = street, 7 = intersection, 8 = address, 9 = premise (building name, property name, shopping center, etc.)). Yahoo instead uses 7 level of granularity⁸.

Accuracy indicates the level to which the resulting coordinates adhere to the real location. In most instances, the geocoding engine can not report this value because the accuracy can not predicted in many cases. However, as illustrated in section 1.4, several authors examined the accuracy of geocoding services and obtained results which depend on service, area, and other parameters.

Precision, finally, refers to the exactness of the digital representation of the resulting coordinates, often recognizable by the number of digits. As Hill (2006) emphasizes "that precision does not imply accuracy in any way".

2.2 Placename Matching

String matching algorithms have been discussed by many authors. Exemplarily White (2005) gives a good overview of different matching approaches. Basically there are two main categories of matching algorithm for string similarity: the first class comprises equivalence methods and the second similarity ranking methods.

Equivalence methods compare two strings and return a value: true or false. The decision is made according to whether the method judges two strings to be equivalent. Word stemming is one approach which tries to find a common stem of two terms. Also synonyms are used to find perfect matches. With addresses especially abbreviations are used to say the same thing: blvd for boulevard, st for street, ave for avenue, rd for road and others.

Similarity ranking methods on the other hand compare a given string to a set of strings ranking these reference strings in terms of similarity. A numeric measure indicates the measure of similarity for each comparison.

Levenshtein distance⁹, for example, is defined as the minimal number of edit operations necessary to convert one string into the other. Edit operations can either be insertions, deletions or substitutions of characters. Cologne phonetics is a phonetic string matching algorithm, originally published by Postel (1969), and finally

⁷ called by this service "accuracy", see <http://code.google.com/intl/en/apis/maps/documentation/reference.html#GGeoAddressAccuracy>. The term accuracy here indicates however only the level of granularity, not a metrical accuracy.

⁸ called by this service context "precision" <http://developer.yahoo.com/maps/rest/V1/geocode.html>

⁹ <http://www.levenshtein.de/>

typewriter distance takes the position of keys on the typewriter-keyboard and their relative alignment to each other into consideration.

Another approach for matching algorithms is phonetic string matching algorithms. Zobel and Dart (1996) introduce the use and mode of operation of phonetic string matching algorithms. With these algorithms spelling is used to identify other strings that presumably have the same pronunciation. Implicitly this is very much language-dependant. Soundex (Knuth 1998) is one of the most popular representatives of this class (Zobel & Dart 1996, Zandbergen 2007).

Some of the algorithms mentioned above, especially the Soundex algorithm, are implemented in Database Management Systems like MySQL®, PostgreSQL and Oracle® and programming languages.

3 Open Address and Geocoding Services

3.1 Overview

The idea of Open Address and Geocoding services is simple and follows the approach of volunteered geographic information, well described by Goodchild (2008) as Volunteered Geographic Information (VGI), one of the characteristics of the Web 2.0. Every person has a local spatial expertise and can act as a "living sensor". If this local knowledge is globally collected in a structured way a valuable and comprehensive information repository is established for a multitude of use cases. Therefore the challenge is to connect these "sensors", collect and organize their expertise and finally provide this accumulated information for public usage, as described by Amelunxen (2010).

3.2 OpenStreetMap.org

The OpenStreetMap (OSM) project currently is one of the most successful VGI projects. Up to now approximately 1.8 billion GPS points have been uploaded by more than 250.000 registered users, converted to 700 Million nodes (points) and 56 Million ways (linear features)¹⁰.

Users can tag the data entered by key value pairs in a very flexible way, e.g. key=amenity, value=Parking or key=Highway, Value=Residential. In addition they are uniquely identified by an identification number. Following these rules, an address schema was established owning its own addr namespace comprising the keys housenumber, housename, street, postcode, city, country and additional keys for interpolation along ways (kind of interpolation, e.g. all, even, odd, alphabetically, number increment used for interpolation, and a tag to indicate the accuracy level of survey used to create the address interpolation, e.g. actual/estimate/potential)¹¹.



Figure 2: Example of tagging a postal address in OpenStreetMap.

Based on these proposed tags users collected more than 4 Million nodes tagged with values for city, country etc. and more than 5 Million with values for street name¹². Address input is done using one of the editors like Potlatch, JOSM, or Merkaartor without special interface for address input.

For Geocoding several approaches are supported by OpenStreetMap¹³.

¹⁰ http://www.openstreetmap.org/stats/data_stats.html [2010-07-21]

¹¹ <http://wiki.openstreetmap.org/wiki/Addresses> [2010-07-21]

¹² <http://tagwatch.stoecker.eu/Planet-100707/En/grouplist.html> [2010-07-21]

¹³ In addition The geocoding of Geonames.org is integrated in the OpenStreetMap portal.

Nominatim is the current search system for OpenStreetMap. In addition to a GUI-based research (from the OpenStreetMap website) REST-based requests are supported offering interoperability with other web services. As result formats, HTML, XML, and JSON can be requested.

Name Finder (http://wiki.openstreetmap.org/wiki/Name_finder) offers search for place names, postal addresses, and supports a “near” search. Name Finder was historically used for geocoding based on the OSM database. Currently the Name Finder database is dated “9 Jan 2009”¹⁴ and seems out of synchronization with the central OSM database.

3.3 OpenAddresses.org

Unlike other volunteered approaches for geodata capturing, OpenAddresses.org does not necessarily require data-capture with a GPS sensor in the field. Data collection is primarily organized through a mapping mash-up, currently based on OpenLayers¹⁵. The user interface is currently available in several languages (e.g. German, English, French, Italian, Polish, etc.). The collected data is provided under GNU General Public License version 3.

Figure 3 shows how the user is required to interactively and manually digitize the locations of the corresponding addresses using the map data layers presented in the mash-up (OpenStreetMap and Yahoo imagery). Stark (2009) presents the predecessor project whole project in more detail.

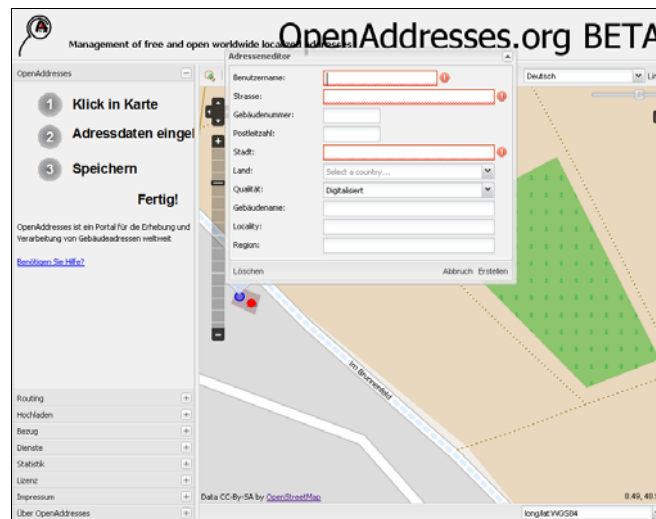


Figure 3: The OpenAddresses.org website offers an interactive geocoding function.

REST-based requests are supported offering interoperability with other web services. As result formats, HTML, XML, and JSON can be requested.

3.4 OpenGeocoding.org

While in developed countries usually governmental or private nation-wide efforts to generate and maintain address data exist, in less developed countries such geodata are still not available (Vivas 2010). The main objective of this project is the generation of a worldwide free address database for developed *and* developing countries – the application aims to collect data as well to provide a geocoding service.

- Data collection: Registered users submit address information and corresponding coordinates (figure 4). Special attention is given to the validation to prevent spamming, by comparison of entered data to existing reference data on country, province and city level. Supported by AJAX based techniques (Garret 2005) for map display and user friendly auto-completion, users can enter addresses and the corresponding coordinates which could be derived from maps or GPS measurements. Alternatively, users can use a Web map (currently with several data layers, e. g., OpenStreetMap and different Google® Maps layers) to digitize the

¹⁴ <http://gazetteer.openstreetmap.org/namefinder/> [2010-07-21]

¹⁵ <http://openlayers.org/>

location. After successful data validation, address and coordinates are saved in the database and can be used for geocoding, interactively or through the API.

- Geocoding: The geocoding service for postal addresses is offered based on this data free of charge through a REST based web service (Richardson & Ruby 2007). The geocoding application does not rely only to its internal database filled by the community. If no relevant results can be retrieved, the request will be forwarded to other publicly available geocoding services. The result will be checked, parsed, and forwarded to the user in a specified format.

Besides offering REST services, the application additionally offers a Web based user-friendly front-end for interactive address submission.

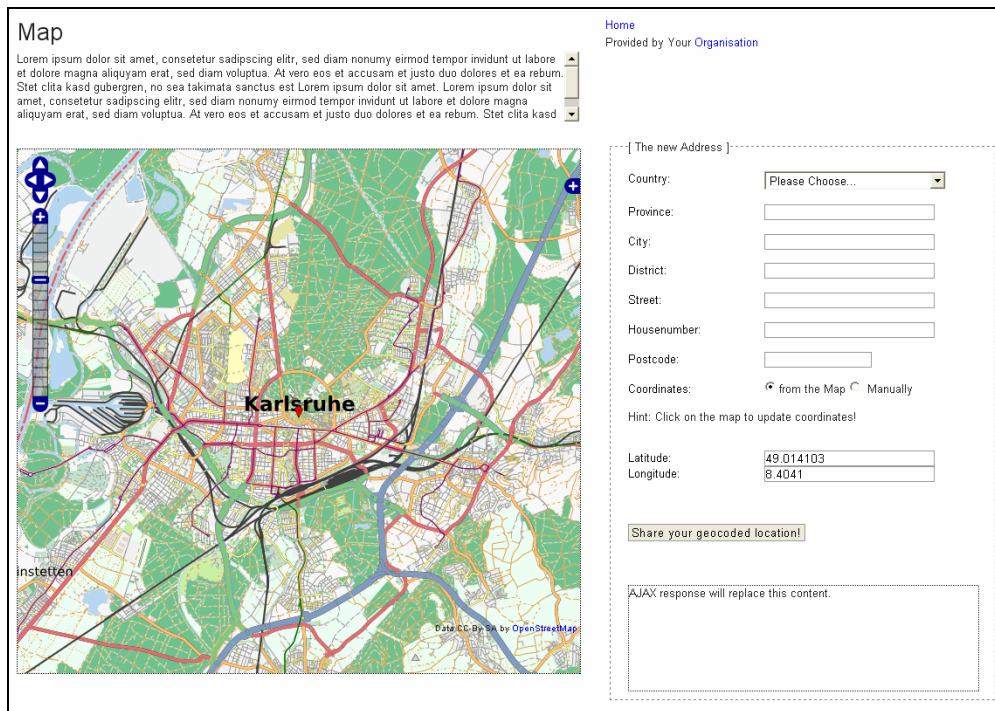


Figure 4: Geocoding using OpenGeocoding.org.

3.4.1 Technologies used

The portal site is based on the Openlayers API, JavaScript and AJAX. PHP is the corresponding server-side scripting language for database access, geocoding and creation of web pages. MySQL, since February 2008 owned by SUN Microsystems, Inc. (<http://www.mysql.com/>) is used as a database management system.

3.4.2 Output formats

The service provides results in XML, KML, GML, JSON, and CSV format (see Wagner et al. 2009). The XML format corresponds to the aforementioned xAL standard, embedded in a KML envelop generated according to the OpenGIS® KML Encoding Standard (<http://www.opengeospatial.org/standards/kml/>). GML output contains coordinates according to the OpenGIS Geography Markup Language (GML) Encoding Standard (http://portal.opengeospatial.org/files/?artifact_id=20509).

4 Conclusion

The “worldwide” geocoding problem is far from being solved for a variety of reasons. From the technical point of view it involves the development of standards, availability of reliable data in a digital format and associated issue of accuracy. In many countries, the main problem is the availability of data (especially in a digital format) since many municipalities often do not have street names! The applications presented here are an approach to overcome these restrictions by participatory, community based data capture. These applications can be an alternative to the existing geocoding services, independent from the commercial solutions provided that a critical mass of participants can be achieved.

References

- Ahlers D., Boll S. 2008: Retrieving Address-based Locations from the Web. Proc. Second International Workshop on Geographic Information Retrieval, GIR'08, October 29–30, 2008, Napa Valley, California, USA.
- Ahlers D., Boll S. 2009: On the Accuracy of Online Geocoders. Proc. GEOINFORMATIK 2009, Osnabrück, Amelunxen C. 2010: An Approach to Geocoding based on Volunteered Spatial Data. <http://www.geoinformatik2010.de/public/abstracts/amelunxen.pdf> [2010-07-19]
- Borges, K. A. V., Laender, A. H. F., Medeiros, C. B., Clodoveu A. D. Jr. 2007: Discovering Geographic Locations in Web Pages Using Urban Addresses. GIR'07, November 9, 2007, Lisbon, Portugal
- Cayo M. R., Talbot T. O. (2003): Positional error in automated geocoding of residential addresses. <http://www.ij-healthgeographics.com/content/2/1/10> [2009-05-27]
- Davis Jr., C. A., Fonseca, F. T., De Vasconcelos Borges, K. A. (2003). A Flexible Addressing System for Approximate Geocoding. Proceedings Geoinfo 2003. Available online at <http://www.geoinfo.info/geoinfo2003/papers/geoinfo2003-25.pdf>. [2009-05-27]
- Davis Jr., C. A., Emerson de Salles E. (2007): Approximate String Matching for Geographic Names and Personal Names. IX Brazilian Symposium on GeoInformatics, Campos do Jordão, Brazil, November 25-28, 2007, INPE, p. 49-60.
- Cover, R, (2008): Markup Languages for Names and Addresses. <http://xml.coverpages.org/namesAndAddresses.html> [2009-05-27]
- Clodoveu A. Davis Jr., Emerson de Salles (2007): Approximate String Matching for Geographic Names and Personal Names. IX Brazilian Symposium on GeoInformatics, Campos do Jordão, Brazil, November 25-28, 2007, INPE, p. 49-60. <http://www.geoinfo.info/geoinfo2007/papers/S3P2.pdf> [2009-06-08]
- Clough, P. 2005. Extracting metadata for spatially-aware information retrieval on the internet. In Proceedings of the 2005 Workshop on Geographic information Retrieval (Bremen, Germany, November 04 - 04, 2005). GIR '05. ACM, New York, NY, 25-30. <http://ir.shef.ac.uk/cloughie/papers/cikm2005-geo.pdf>
- Ding, J., Gravano, L., Shivakumar, N. 2000. Computing Geographical Scopes of Web Resources. In *Proceedings of the 26th international Conference on Very Large Data Bases* (September 10 - 14, 2000). A. E. Abbadi, M. L. Brodie, S. Chakravarthy, U. Dayal, N. Kamel, G. Schlageter, and K. Whang, Eds. Very Large Data Bases. Morgan Kaufmann Publishers, San Francisco, CA, 545-556.
- .Garrett, J. J. (2005). AJAX: A New Approach to Web Applications. Technical report, Adaptive Path. <http://www.adaptivepath.com/publications/essays/archives/000385.php>. [2009-05-27]
- Goodchild M. F. (2008): Volunteered geographic Information, GEOconnexion International Magazin, October, 2008
- Government of Victoria, 2006. Whole of Victorian Government Standard - Street Address Data Standard. A standard for the storage, interchange and validation of street address data by Victorian Government agencies. [http://www.dtf.vic.gov.au/CA25713E0002EF43/WebObj/-DStreetAddressData/\\$File/DI%20Street%20Address%20Data.pdf](http://www.dtf.vic.gov.au/CA25713E0002EF43/WebObj/-DStreetAddressData/$File/DI%20Street%20Address%20Data.pdf) [2009-05-27]
- Grubestic T. H., Murray A. T. 2004: Assessing positional uncertainty in Geocoded Data. Proc. Urban Data Management Symposium, 2004
- Gruber, F. 2006. Comparing the Mapping Services. <http://www.techcrunch.com/2006/04/17/comparing-the-mapping-services> [2008-05-09]
- Hill, L. L., 2006. Georeferencing: The Geographic Associations of Information. MIT Press, Cambridge, Massachusetts
- INSPIRE 2007: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) 14.03.2007. <http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2007:108:SOM:EN:HTML> [2010-07-19]
- INSPIRE 2009: Draft COMMISSION REGULATION implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services 14.12.2009. <http://ec.europa.eu/transparency/regcomitology/>

- index.cfm?do=Search.getPDF&1A6b4z6edALEzOuvQ2DQwEuYwr24bl+u6M8oCwqlYrvB7EJR+poTzWZ/2wT/z/JFTr7x0HnynbCJdi/BzR4ZvdPpAur0FOHhej8jYcN49FA=^ [2010-07-19]
- INSPIRE 2010: INSPIRE Data Specifications on Addresses - Guidelines v 3.0.1 03.05.2010.
http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.0.1.pdf [2010-07-19]
- Knuth, D. E. 1998: *The Art of Computer Programming, Vol. 3: Sorting and Searching*. 2. Aufl., Reading/Mass.: Addison-Wesley.
- Levenshtein, V. I. (1965). "Binary Codes Capable of Correcting Spurious Insertions and Deletions of Ones." *Probl. Inf. Transmission* 1: 8-17.
- Lind, M. 2010: Benefits for society of a strong address infrastructure - Danish case study. European Address Forum Brussels, 15 June 2010, http://www.europeanaddressforum.eu/-EAF/index.php?option=com_attachments&task=download&id=39 [2010-07-19]
- Mikheev A., Moens M. and Grover C. (1999) Named Entity recognition without gazetteers. In *Proceedings of the Annual Meeting of the European Association for Computational Linguistics EACL'99*, Bergen, Norway, 1-8.
- Nicoara, G., 2005. Exploring the Geocoding Process: A Municipal Case Study using Crime Data.
http://charlotte.utdallas.edu/mgis/prj_mstrs/2005/Summer/greta/Nicoara_Masters/Website/main.html. [2008-05-09]
- Open Geospatial Consortium 2006: OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture http://portal.opengeospatial.org/files/?artifact_id=18241 [accessed 2009-05-22]
- Postel, H. J. (1969). *Die Kölner Phonetik. Ein Verfahren zur Identifizierung von Personennamen auf der Grundlage der Gestaltanalyse*. IBM-Nachrichten, 19, 925-931.
- Ratcliffe, J. H., 2001. On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units. *International Journal of Geographical Information Science*, Vol. 15, no. 5, 473 – 485
- Richardson L., Ruby S. (2007): *RESTful Web Services*. O'Reilly
- Stark H.-J. (2008): Open Geodata - am Beispiel von OpenAddresses.ch, in: Strobl, J. et al. (Hrsg.), *Angewandte Geoinformatik 2008. Beiträge zum 20. AGIT-Symposium Salzburg*. Heidelberg, 2008
- Stark H.-J. (2009) OpenAddresses - Free geocoded street addresses. *Proc. Applied Geoinformatics for Society and Environment (AGSE) 2009*, <http://applied-geoinformatics.org/> [2009-06-14]
- Schäfer B., Kunz C. 2008: Geocoding. Unpublished project Stuttgart University of Applied Sciences
- Theurer, D. 2007: How to build a Maps Mash-up. <http://www.theurer.cc/blog/2005/11/03/how-to-build-a-maps-mash-up/> [2010-07-20]
- Vivas, P. 2010: "Addressing the world – An address for everyone". European Address Forum Brussels, 15 June 2010, http://www.europeanaddressforum.eu/-EAF/index.php?option=com_attachments&task=download&id=35
- Wagner, D., Zlotnikova, R., Behr, F.-J. 2009: XML-Based and Other Georelated Encodings: Overview of Main Existing Geocoding Formats. In: Behr, F.-J., Schröder, D., Pradeepkumar, A. P. (Editors) 2009: *Proceedings Applied Geoinformatics for Society and Environment (AGSE 2009)*. Publications of the Stuttgart University of Applied Sciences, Hochschule für Technik Stuttgart, Volume 103 (2009), 317 pp., ISBN 978-3-940670-13-7, http://www.gis-news.de/papers/AGSE_Proceedings_2009_07_01_1025.pdf [2010-07-20]
- White S. (2005): Tame the Beast by Matching Similar Strings. http://www.catalysoft.com/-articles/MatchingSimilarStrings.html?article=Tame_the_Beast_by_Matching_Similar_Strings_14 [2009-05-27]
- Zandbergen P. A. 2007: A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems*, Volume 32, Issue 3, May 2008, Pages 214-232
- Zobel J., Dart P. (1996): *Phonetic String Matching: Lessons from Information Retrieval*. www.cs.rmit.edu.au/~jz/fulltext/sigir96.pdf [2009-05-27]

OpenStreetMap: Data Model, Crowdsourcing, and Technology

Franz-Josef Behr

Department of Geomatics, Computer Science and Mathematics,
University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
franz-josef.behr@hft-stuttgart.de

KEYWORDS:

OpenStreetMap, CrowdSourcing, Mashups, Google Maps, Open Source, OpenLayers, JavaScript

EXTENDED ABSTRACT

Volunteered Geographic Information (VGI) is one of the characteristics of the Web 2.0. Every person having local spatial expertise can act as a 'living sensor'. If this local knowledge is collected on a regional or global level in a structured way a valuable and comprehensive information repository is established for a multitude of use cases.

The OpenStreetMap (OSM) project started recently is one of the most successful VGI projects and an excellent example of crowdsourcing. Up to now approximately 1.8 billion GPS points have been uploaded by more than 250000 registered users, converted to 700 million nodes (points) and 56 million ways (linear features).

In contrast to the approach used in Geographic Information Systems (GIS) where geo-objects (features) are categorized into feature classes or layers, OSM is based on a surprisingly simple data model consisting of just four data types: Nodes, representing points of interest (POIs) and acting as vertices, ways, and relations which aggregate nodes, ways and other relations to larger units. Relations are used for geo-features comprising several sub-features, similar to a feature collection in the Geographic Markup Language (GML), and are among others used to model turn restrictions for further routing. These structural elements are used to build polygons as well as topological relationships.

In order to create map features users can tag nodes, ways, and relations in a very flexible way by key value pairs, e.g. key = amenity, value = Parking or key = Highway, Value = Residential. In addition they are uniquely identified by an identification number.

The tags are based on a consensus principle; i. e. there is no fixed catalogue of features classes which is defined by a custodian instance. In contrary, it is based on user agreement and is subject to some user debates, modifications, and developments encompassing the evolvement of the whole project. For orientation a list of commonly used features is listed in the wiki (<http://wiki.openstreetmap.org/wiki/Tag>). This flexibility, of course, has some disadvantages but enables easy capturing of new kind of features.

Typically data collection is based on using GPS devices. In addition it is possible to use licence-free maps or records for data collections. In addition Yahoo!'s geo-imagery or freely available web map services can be used for digitization. According to different level of user experience and requirements different editors for data capture and updating are available. For many purposes, the Flash-enabled browser-based editor Potlach can be used. The Merkaartor editor offers some more functionality which is still exceeded by the Potlach editor, at the expense of complexity.

All data are stored in a common geo-database. Through an application program interface it is possible to obtain the whole dataset (the so-called planet file), incremental updates, features for certain regions on the earth, and features with specific tags or other characteristics. Also a web-based graphical user interface enables data export in different formats.

OpenStreetMap has its own data format, mirroring the structure of nodes, ways, relations, and tags in its own XML-based markup language. Conversion programs support the transformation into other formats and the bulk loading into geo-databases. Some Geographical Information Systems already support import of OSM data.

Closely related to OpenStreetMap are some software packages (e. g., Osmarender, Mapnik) supporting the rendering of the data. The visualization is based on style rules which can be adapted for many cases of cartographic needs. The generation and provision of tiles allows the integration into mapping applications, e. g. the browser-based OpenLayers mapping solution.

Its data usage is not restricted by any restrictive licence. This flexibility and freedom (see Creative-Commons-Attribution-ShareAlike-Licence (CC-BY-SA, <http://creativecommons.org/licenses/by-sa/2.0/>)) encourages the

usage of OpenStreetMap data and tools in many fields of application. One core theme is the application for humanitarian needs, e. g. natural hazards or man-made crisis situations.

References

Topf, J., Behr, F.-J. (2008): Vermessene Freiheit. GIS BUSINESS, 5/2008, S. 38 - 41

Goodchild M. F. (2008): Volunteered Geographic Information, GEOconnexion International Magazine, October, 2008

Ramm, F., Topf, J. 2008: OpenStreetMap. Die freie Weltkarte nutzen und mitgestalten. Jahr: 2008. Lehmanns Media, 288 Seiten, Preis: 29,95 Euro, ISBN: 3-86541-262-9

Data Visualization Using Google Maps: the Hard Way and the Easy Way

Shailesh Shrestha ^{a*} and Franz-Josef Behr ^b

^aPhotogrammetry and Geoinformatics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart (Germany), shreshai@yahoo.com

^bDepartment of Geomatics, Computer Science and Mathematics,
University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart (Germany), - franz-josef.behr@hft-stuttgart.de

KEYWORDS:

Data Visualization, Google Maps API, Google Fusion Table

ABSTRACT:

Commercial mapping tools and Geographic Information Systems (GIS) are often quite costly and out of reach of many Non Governmental Organisations (NGOs) which operate on low budget. However, emergence of freely available spatial tools over past few years has propagated the use of spatial information to a broad range of users. One of these tools, Google Map API has revolutionized online mapping service to display and integration of spatial data in a lot of different ways. But it is still a daunting task to display spatial data in Google Map for people who do not have knowledge about programming languages.

This paper aims to introduce Google's Fusion Table, cloud-based service for data management and integration, as an easy tool to visualize data quickly and easily without almost any prior programming knowledge required. At the beginning, brief description of displaying point data with the use of Google Maps API based on Flash will be given which is followed by much simple and straight forward method using Google Fusion Table visualization approach.

There are some limitations associated with Fusion Table but this is not a hindrance for people who are happy to visualize the data with few clicks in a graphical way with no database setup costs. Therefore Google Fusion Table can be especially advantageous for people with no programming experience and for organizations operating on low budget.

1 Introduction

Commercial mapping tools and Geographic Information Systems (GIS) are often quite costly and therefore out of reach of many Non Governmental Organisations (NGOs) which operate on low budget. However, the emergence of freely available spatial tools and web mapping applications over the past few years has propagated the use of spatial information to a broad range of users. Nowadays, there are several free tools and so-called mash-ups (Purvis et al. 2006, Udell 2008) exposing their application development interface (API), which can be utilized for disseminating and processing of spatial information. One of these tools, Google Maps API has revolutionized online mapping service to display and integration of spatial data in a lot of different ways. It allows users to embed Google Map in own websites and with some programming knowledge opens a wide array of possibilities to create robust and interactive maps.

Users and developers have widely embraced this API according to their requirements and hence Google Map has established itself as a pioneer among the online mapping service. Since its inception, Google Maps API makes use of JavaScript as a scripting language for programming. In Nov 2008, Google has released new version of Google Maps API based on Flash (<http://code.google.com/intl/en-EN/apis/maps/-documentation/flash/>) which uses XML based vocabulary language named MXML and ActionScript. But it is still a daunting task to display spatial data in Google Map for people who do not have knowledge about programming languages.

This paper aims to introduce Google's Fusion Table (<http://tables.googlelabs.com/>), a cloud-based service for data management and integration, as an easy tool to visualize data quickly and easily without almost any prior

* Corresponding author

programming knowledge required. Google Fusion Table is a cloud based data management and integration service supporting collaboration among multiple users and has since received considerable use (Gonzalez et al., 2010). The term cloud computing denotes the concept to use different kinds of computing resources (e.g., services, storage ...) through the web¹⁶.

For the examples provided here, a text file in CSV format is used containing information about 76 points of interest (POIs) around the Stuttgart University of Applied Sciences containing attributes such as latitude, longitude, address, website and type of place.

2 Data visualization using Google Maps API for Flash

The Google Map API for Flash (<http://code.google.com/intl/en-EN/apis/maps/documentation/flash/>) allows users to add more dynamic and interactive mapping application to their website. To use Google Maps API for Flash, users must obtain a mandatory API Key from Google. The process is rather simple and straight forward. Users need to provide the URL of the site that would use the application to be developed and need to have a Google account. After this process, SDK was downloaded from Google's website which contains all necessary class libraries (ActionScript library made of public class of Maps API as SWC file) necessary to develop application utilizing the Flash environment. For development of simple interface and embedding of Google Maps API for Flash, Flex Builder 3 software was used. It provides WYSIWYG type of design interface where lots of pre-built components can be simply drag and drop to create an interface. It uses the XML based mark-up vocabulary MXML for creating interface and the programming language called ActionScript for capturing user interaction and response. A good introduction to develop application with Flex can be found in Cole (2008). The overall steps for creating a simple application using Flex Builder 3 utilizing Google Maps API for Flash is shown in figure 1. Each communication is denoted by sequential numbering in which it takes place.

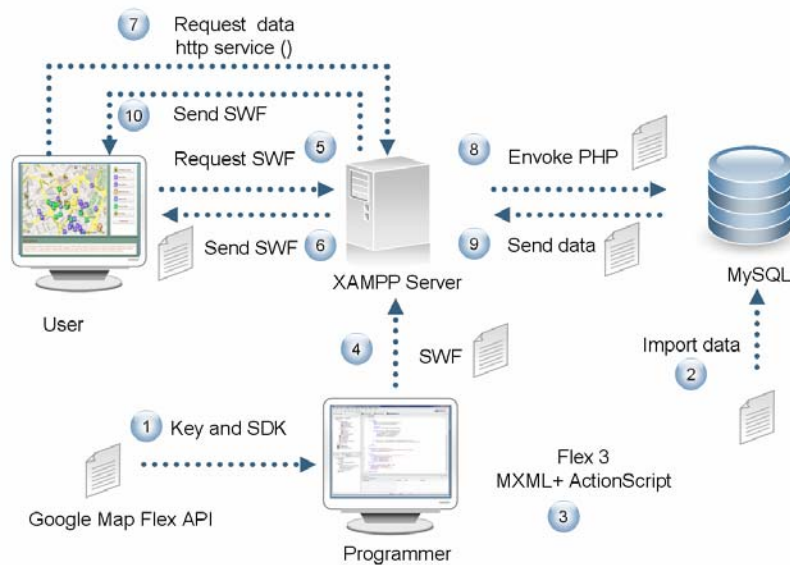


Figure 1: Steps in preparing simple webpage with Google Maps API for Flash with MySQL backend

2.1 Establishment of XAMPP and MySQL Database

The web server tier is established by downloading XAMPP installer (<http://www.apachefriends.org/en/xampp-windows.html>) which is a free and open source cross-platform webserver package. The XAMPP package comes with MySQL database management system and PHPMyAdmin. Before importing data in the MySQL database, it is very important to set up a structure of a table specifying the data types for each column. Afterwards the data were imported into MySQL database.

¹⁶ Another representative of cloud computing is Amazon Elastic Compute Cloud, <http://aws.amazon.com/de/ec2/>

2.2 Outputting Data as XML with PHP

To retrieve the data from MySQL database, a PHP script was written. For the Flex environment, it is only possible to pass external data as XML format. Creating a XML format data from the database is possible either using server's PHP *dom_xml* functionality which will result in saved XML data or using PHP's *echo* function. The second method was chosen over the first one because of familiarity of the author with the second method. A PHP file was created utilizing simple *mysql_connect*, *mysql_select_db* and *my_sql_query* commands of PHP.

```
<?PHP
mysql_connect("localhost","root","");
mysql_select_db("googlemaps");
$result = mysql_query('select * from data');
echo "<?xml version='1.0'?'><map>";
while ($row = mysql_fetch_assoc($result))
{
    echo "<loc><lat>".$row["lat"]."</lat>";
    echo "<lon>".$row["lon"]."</lon>";
    echo "<name>".$row["name"]."</name>";
    echo "<type>".$row["type"]."</type>";
    echo "<address>".$row["address"]."</address>";
    echo "<website>".$row["website"]."</website></loc>";
}
echo "</map>";
?>
```

Figure 2: PHP file to extract data from MySQL database

2.3 Initialization of Map

A single panel was created in Flex Builder, and Google Map was added as child of application tag (see Fig. 3) . A namespace was assigned so that to link with ActionScript. The height and width attributes of the map child were set. The displayed map is still conveying standard data layers.

```
<!-- Panel to contain Google map -->
<mx:Panel title="Google Maps application using flex for students portal HFT" height="490" width="930" x="9.0" y="9.0">
    <maps:Map id="map" xmlns:maps="com.google.maps.*" width="910" height="440"
        mapevent_mapready="onMapReady(event)"
        key="ABQIAAAA6MkmLRCxSAoD1SHyS05djRT2yXp_ZAY8_uFC3CFXhHIEiNvwKxSy3U4nb-sqbdmmNPYiBcwh8_rf-A"/>
</mx:Panel>

//Initialize googlemaps with specified center containing zoom control and type control
public function onMapReady(event:MapEvent):void {
    map.setCenter(new LatLng(48.780309,9.172602), 15, MapType.NORMAL_MAP_TYPE);
    map.addControl(new ZoomControl());
    map.addControl(new MapTypeControl());
}
```

Figure 3: Embedding map (MXML) and initializing map with added controls (ActionScript)

To extract the data from the server and display those data as custom markers, a *httpservice ()* function of Flex which allows to communicate with the server was defined. With this service, the application calls the PHP file described above to access the data residing in the MySQL database which returns an array of data as output. To loop through that array and extract the data, a function which reads the array line by line and stores the required values in a variable was written. For the purpose of displaying markers as different recognizable icon based on type, some images were downloaded from a website maintained by Mollet and Stasiak (2010) which contains huge collection of beautiful custom markers and were embedded in Flex as an object array. A custom function named *createMarkes ()* which takes five parameters of different types as input and actually is used to create the markers. In additions, it creates the information window that will pop-up if a mouse click event occurs on the marker (see Fig. 4 and Fig. 5).

```
// Function that creates markers and infowindow page for information
public function createMarker(latlng:LatLng, name:String, address:String, type:String, website:String): Marker {
    var marker:Marker = new Marker(latlng,new MarkerOptions({icon: new customIcons[type]}));
    var html:String = "<b>" + name + "</b> <br/>" + address + "<br/>" + website ;
    //var html:String = "<b>" + name + "</b> <br/>" + address + "<br/>" + "<alt>" + website;
    marker.addEventListener(MouseEvent.CLICK, function(e:MouseEvent):void {
        marker.openInfoWindow(new InfoWindowOptions({contentHTML:html}));
    });
    return marker;
}
```

Figure 4: Custom function to creates markers with different icons



Figure 5: Google Map displaying different places of interest in custom icons

As shown above each marker is displayed with a custom icon indicating the corresponding category of the POI. For example, bars are represented by a blue icon containing a glass full of beer, and libraries are represented by an orange icon containing stacks of books etc.

As can be seen, displaying point data with Google Maps API for Flash encompasses several steps and is not an easy and straight forward process. One must be familiar with different programming language like PHP, MXML and ActionScript together with server side technology, e.g., Apache HTTP server and MySQL database management system. Many experts from field of Geoinformatics may be familiar with these technologies, whereas for a person working in a NGO with limited educational background, such concepts might be too complex.

In the following sections, data visualization of same point data is shown but now with a simple and straightforward process with Google Fusion Table.

3 Data visualization using Google Fusion Table

Google Fusion Table is the newest addition in the wide range of already available Google's Geo Tools. With the introduction of this approach in June 2009 by Google Labs, Google has opened its computing capabilities for web based data management system. User can upload their data and can instantly create business graphics like bar charts, pie charts, timeline etc. and other map visualization to display spatial datasets on Google Maps. Primarily developed as a data management system it enables collaboration among different organizations and facilitates integration of data from multiple resources by means of joining tables to gain new insight into already existing data. In current version of Google Fusion Table, user can upload data up to 100 MB per data set limited to 250 MB per user (Halevy 2009).

3.1 Uploading the Data

The first step is to upload the data in the Google Fusion Table. For this, users need to have a Google account. After logging in users have an opportunity to upload data in various formats like xls, xlsx, csv, ods and kml. The limit for file size is different for each format, ranging from 1 MB for ods, xls and xlsx to 100 MB for kml and CSV data type. Our data which is in CSV format can be easily loaded in Google Fusion Table. The snap shot of loaded data in Google Fusion Table is shown in figure 6. After loading the data, users can instantly visualize data in different charts form using different columns of the data. It is also possible to merge the data with already existing data that are shared by other people. These parts are not covered in this paper.



Figure 6: Importing data into Google Fusion Table

3.2 Column Configuration

Next step is to do some configuration that Google Fusion Table knows which column is used to geocode the POIs. In most cases, Google Fusion Table automatically detect when a column contains information that specifies location information such as latitude-longitude pair or an address. But it is always good to manually define columns which have location information or at least to cross check. In our case, we have location

information in two different columns named Longitude and Latitude. To define these columns as location information, *Modify column* from *Edit menu* is clicked and Longitude and Latitude columns are assigned as type location (c.f. figure 7Figure). Since location information in the form of addresses is also available, we can also use that information to geocode instead. However for that purpose the address information has to be in one column such that street, city and zip information are separated by space. This can be easily achieved in Microsoft Excel by using *concatenate* function. For example, formula *concatenate (E2, " ", F2)* will concatenate the text in column E and F with a space in between.

3.3 Visualize the Map

Now everything is set for visualizing the data. Directly The graphical user interface allows the immediate visualization of the POIs including controls such as zooming, panning and layer selection. Points are displayed as default red markers. When clicked, all the information associated with that markers pop up. With few clicks, without use of any programming knowledge and short time, users now have a tool at their disposal which can be used for different purposes. However it has to be realized that it has some limitations such as it cannot be customized as in a way which is possible in Google Maps API and Google Map API for Flash. For example user cannot choose a specific point as centre of map display and specify different types of markers for different objects as shown in Figure . Nevertheless most often people are more than happy to have an ability to display spatial data in the form as shown in figure 8. With the map, user can control the information to be displayed in *Info Window* by using six inbuilt templates by using *Configure Info Window* or user can easily get embedding code by clicking *Get Embeddable Link* or user can export it as a KML (Keyhole Markup Language) to view the data in Google Earth and perform some customization such as changing the icon type and adding organization's logo as static layer.

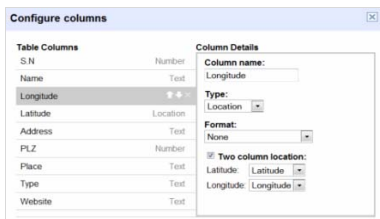


Figure 7: Configure columns in Fusion Table

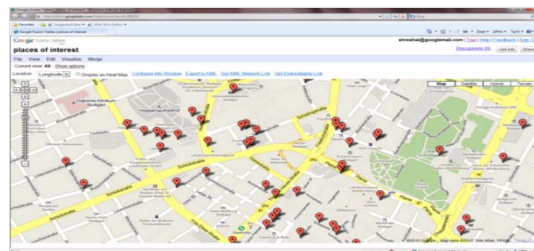


Figure 8: Visualization in Google Map using Fusion Table

3.3.1 Configure Info Window

If a user is familiar with the HTML, many customizations can be done on *Info Window* of markers to display information once a marker is clicked. Wide array of possibilities are there such as adding descriptive text, link and images with little bit of HTML coding. All the columns in a table in Google Fusion Table are available as fields which can be initiated in the HTML code. A small customization example is given in figure 9. With few HTML tags, only Name, Type and Website information as hyperlink are displayed in the *Info Window*. If we have a column containing images location associated with markers that are available through internet, then it could be easily incorporated in the *Info Window* with single addition of HTML tag ``.



Figure 9: Customization of Info Window with HTML tags

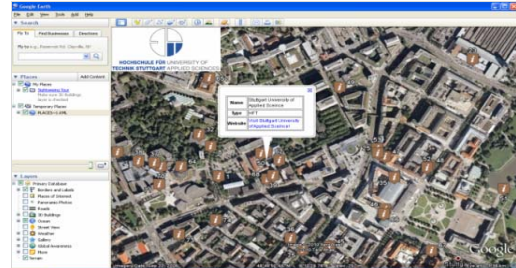
3.3.2 Export to KML

There is an option for exporting the map created in the Google Fusion Table as KML which can be viewed on any computer with internet access with installed Google Earth application. Thus exported KML file can be easily viewed by just double clicking in the file and also can easily shared with other people. KML is an XML

format originally created in 2001 which has since evolved as international standard for visualization of geographic data (Wernecke, 2008). With simple addition of few lines, user has the possibility of changing the marker type and adding institutional logo as overlay on Google Earth which is shown in figure 10. In the code *HftLogo.gif* is a custom logo of HFT Stuttgart and *info.png* is a custom type marker downloaded from website maintained by Mollet and Stasiak (2010). These files should reside within the same folder containing the KML file and can be zipped as KMZ file. Zipping as KMZ file ensures that every linked file with KML file is also transferred with the KMZ file. Wernecke (2008) gives excellent detailed information about KML with amazing things that can be achieved with KML.

```
<ScreenOverlay>
  <name>Shrestha KML example</name>
  <Icon>
    <href>HftLogo.gif
    </href>
  </Icon>
  <overlayXY x="0" xunits="fraction" y="1" yunits="fraction"/>
  <screenXY x="0" xunits="fraction" y="1" yunits="fraction"/>
</ScreenOverlay>

<IconStyle>
<color>FFFFFF</color>
<scale>1.1</scale>
<hotSpot x="0.5" y="0" xunits="fraction" yunits="fraction"/><Icon>
<href>info.png</href>
</Icon>
</IconStyle>
```



Adding custom institutional logo and custom marker

Visualization of KML export from Google Fusion

Figure 10: Little customization in the KML

3.3.3 Get embedded link and integrate with user website

After creating map with Google Fusion Table, it is very easy to embed it into own website. It is enough just to copy the URL from the *Get Embeddable Link* and paste it in a desired website. The example of an embedded map, presented in figure 11, can be seen online by visiting <http://shreshai.blogspot.com/2010/07/embedded-link.html>. Such a blogging facility can be convenient for inexperienced users – the time needed for basic setup of the blog and to incorporate the embed link is less than five minutes.

3.4 Visualization with Google Fusion Table Choropleth

Choropleth, which are frequently used in cartography, is a colour coded map comparing region with each other. With Google Fusion Table, it is possible to display a choropleth map and embed it in a website with ease. Choropleth for the whole world or for a particular country can be created. Basically, the name of the country or states is used for the geocoding purpose. For the illustration purpose, data related to area of each states of the Germany which is available at <http://www.citypopulation.de/Deutschland-Agglo.html#Land> was directly accessed from Microsoft Excel with *From Web* tool (xls file format). The file was uploaded to Google Fusion Table and visualized as *intensity map*. Data can be displayed as *Regions* or as *Markers* as shown in Figure. Both choropleths can be viewed on the authors blog mentioned above.



Data displayed as regions



Data displayed as markers

Figure 11: Choropleth map made with Google Fusion Table

In such maps, user can hover each region and instantly view the data associated. Some inaccuracies were observed in the both maps. In the region based choroplethmap, one state was not geocoded at all whereas when the same data is displayed as markers, three states (!) were geocoded within the state of Bavaria. However such discrepancies were not observed when choropleth maps in the global scale were made.

If such a choropleth map should be made in Google Maps API for Flash using scripting language, the developer should be highly proficient in MXML and ActionScript as well. In addition, it would also take considerable long

time. Process of making choropleth using Google Fusion Tables is very easy. Nevertheless, it is sad that Google has not provided option of selecting different colour scheme. It always use default green colour scheme. Moreover, legends which are visible when visualizing in Google Fusion table vanishes when the embedded code is copied and pasted in user's own website.

4 Issue of data security and state of Google Fusion Table

Initial step for visualizing own data is to upload the data in the Google Fusion Table. It can be shared to public or can be shared to limited people desired by the up loader. User can later delete the data however data might never be deleted from the Google database. Hence, there is a prime concern about the data security issue and many people are often reluctant to use such services. Therefore it is wise not to use Google Fusion Table for visualization of very confidential data. In addition when requesting for embeddable link, it is mandatory to choose visibility option of the table as *Public* or *Unlisted*. In the *Unlisted* mode only person who has link of the table (Fusion Table provides link for each table) can access the data but there is still a risk that a person with link can forward that to an unauthorized person.

As mentioned in the introduction, Google Fusion Table was released by Google Labs, which is a platform to test different ideas by Google Engineers. Hence it is officially in "Beta Version". Therefore it is suggested not to plan big projects using Google Fusion Table. There is a possibility that any of the experiments may be unavailable in near future or even removed completely.

5 Conclusion

Developing Google Map applications using either Flex or Javascript involves different programming languages. However it offers much more customization possibilities than with Google Fusion Table. There is also no concern about the data security when using Google Map API as the data resides on the local server. On the other hand Google Fusion Table offers a simple and straight forward interface to create interactive maps for non-programmers with no need of database setup. There are some limitations associated with Fusion Table, but this is not a hindrance for people who are happy to visualize the data with few clicks in a graphical way at no cost. Therefore Google Fusion Table can be especially advantageous for people with no programming experience and for organizations operating on low budget. Really cool applications can be made by just using the graphical interface of Fusion Table. However it is possible to integrate Fusion Table with Google Maps API and Google Maps API for flash through Google Fusion Table API. But doing so, Fusion Table will also be yet another "rocket science for non programmers". Since, Google Fusion Table is still in the development stage, there are some deficiencies as illustrated by the chloropleth map of states of Germany.

References:

- Cole, A., 2008. Learning Flex 3: Getting Up to Speed with Rich Internet Applications. O'Reilly Media, First edition
- Gonzalez, Alon Y. Halevy, Christian S. Jensen, Anno Langen, Jayant Madhavan, Rebecca Shapley, Warren Shen, Jonathan Goldberg-Kidon 2010: Google Fusion Tables: Web-Centered Data Management and Collaboration. <http://www.google.com/buzz/goog.research.buzz/3Nf2QjBos2s/Google-Fusion-Tables-Web-Centered-Data-Management> [2010-07-22]
- Halevy, A., 2009. Google Fusion Tables, User Experience <http://googleresearch.blogspot.com/2009/06/google-fusion-tables.html> [20 July 2010]
- Mollet, N. and Stasiak, M., 2010. Google - Maps-icons. <http://code.google.com/p/google-maps-icons/> [accessed 20 July 2010]
- Purvis M., Sambells J., Turner C., 2006: Beginning Google Maps Applications with PHP and Ajax: From Novice to Professional. Apress
- Udell S., 2008: Beginning Google Maps Mashups with Mapplets, KML, and GeorSS: From Novice to Professional. Apress
- Wernecke, J., 2008. The KML Handbook: Geographic Visualization for the Web. Addison-Wesley Professional, First edition

Technology Innovations

Mobile Mapping – A Carfinder Application for a Modern Carsharing System

Paul Rawiel

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
paul.rawiel@hft-stuttgart.de

KEYWORDS:

Mobile mapping, Routing, GPS

ABSTRACT:

Big cities nowadays have many problems to solve. One of them is the big amount of traffic that causes air pollution, noise and transport problems. The politics tries to handle these problems with different kinds of measures. One possibility to decrease the number of cars running in the cities is to provide a pool of cars shared by many people. Daimler realized a new concept of a modern carsharing model, that aims to guarantee mobility to people without the need to have an own car. To reach that goal, a certain number of cars have to be distributed in the area of the city and it should be easy and flexible for people to rent a car. The problem in big city centers may be to find a car, even if it is not far away, but around a corner not visible for the client. A mobile application to find the next available car is the solution that shall be presented here. In a student project such a carfinder was developed in cooperation with Daimler TSS. During the project, an analysis of possible infrastructures had to be done, a database had to be configured and a communication with a Daimler TSS server had to be established to get the coordinates of free cars. Then another server had to be configured for the routing from a current position to the next available car. On top of that, a web application that serves as middle ware between the web client and the routing server and an application for mobile devices had to be developed. This was realized in several teams with different responsibilities. The project and its results are presented here.

Car Navigation Systems: Technical Principles and Trends

Cesar Amador Ocampo Lecaros

Reichenhaller Str. 8, 70372 Stuttgart, Germany,- cepeoca@hotmail.com

KEY WORDS:

Car navigation systems, GPS, TMC, Car sensors, advanced driver assistant systems

ABSTRACT:

A navigation system in the automotive branch is a complete hardware and software system. It is designed to guide the drivers to their desired destination. Two categories of navigation systems could be set: one integrated into the infotainment systems in the automobiles and the other portable systems that are normally called PNA (portable navigation assistant). The focus of this paper will be on the integrated systems. These systems are more complex in hardware and software. In the following sections the most common components of all car navigation systems and the current trends are described.

1 Main SW Components of a Navigation System

1.1 Map Data

The map data is a basic component for a navigation system. In the map database, mainly the motorways with their different attributes (e.g. form of way, functional road class, speed category, etc.) are stored and also some polygonal information could be included, as for example the land use or street blocks (mainly used for HMI proposes) and finally some points of interest (POI) representing special locations that the user could ask for; normally the user could edit it and add new POIs to the system.

1.2 Input Data Module

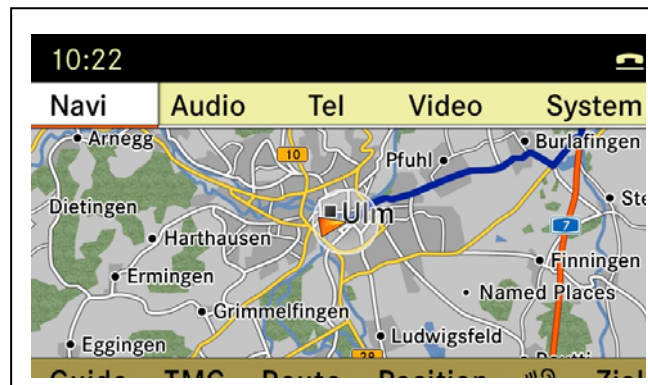


Figure 1: Menu of navigation input.

The user has many possibilities to enter data or to control the system ((Figure 1). The most common are: the DDS used by Audi and many after mark systems, the CCE use by Daimler and BMW, touch screen used for many after mark systems and PDA navigation systems and finally the speech dialogue system (SDS) which allow the user to introduce the destination or navigate on the menu system using voice commands.

1.3 Route Guidance

The system will guide the driver using visual or acoustic route guidance. The visual route guidance could be in the form of an arrow navigation in which the driver will see the route guidance as arrows and other symbols; this is the lower cost method. The more expensive way is the map navigation in which the user will have the

possibility to visualize the complete map and commonly in colors in addition to the arrow information. The map navigation could be in 2D or 3D depending on the processing and graphical capabilities of the system (Figure 2).

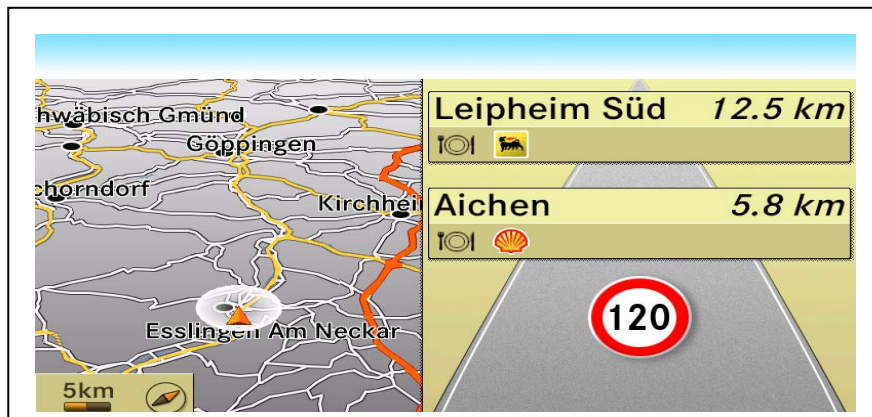


Figure 2: Visual route guidance on a car navigation system.

Besides or in addition to the visual route guidance there is also acoustic route guidance. Typically there are three types of notifications to the driver. The first two are to prevent the driver from an action that should only be executed on the way ahead; these two notifications are given at different distance intervals depending on the systems and the location of the car (inside cities or highways), the third notification is the action notification and this indicates to the driver to conduct an action immediately. Besides the notifications the navigation system normally includes a repeat, that allows the user to again hear the latest voice advice.

1.4 Scout module

This module is present in all automotive navigation systems and is responsible for calculating, evaluating and finding an optimal route from a starting point to a destination point. Optimal route means a route that fulfills the determined conditions, as for example the shorter way or the less time consuming route.

1.5 Sensors Module

This module is responsible for collecting all the information from the sensors around the car. In case of simple systems the module could be a replacement for a GPS module. In the most complex system this module will collect information from the gyrometer, radar, laser sensor, etc.

1.6 TMC Module

This software module is an optional one that is used to interpret the traffic message channel (TMC) that is used in many European countries to send valuable information regarding the traffic conditions on the different routes (figure 3). The traffic announcements (TA) are transmitted using the radio data system (RDS). The TA are transmitted digitally and are not audible to the driver. This makes it possible to transmit the TA without having to interrupt the radio program. The TMC is nowadays completely available in Germany, The Netherlands, France, Austria, Switzerland and Great Britain. The TMC are transmitted using the ALERT-C protocol (Advice and problem location for European Road Traffic version C).

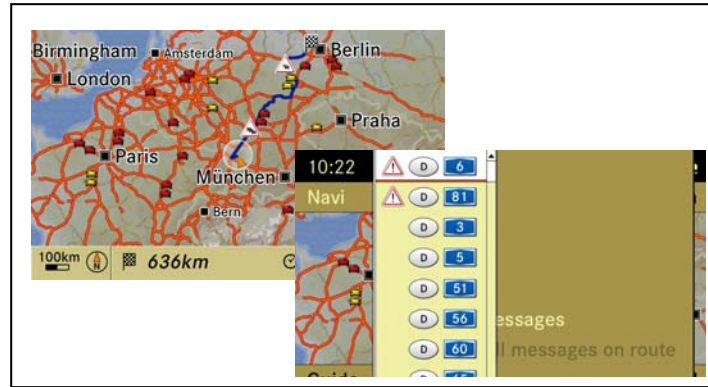


Figure 3: TMC screen shots on a car navigation system

2 Main Hardware Components of an Automotive Navigation System

Besides the GPS the more complex navigation systems include more hardware to allow the system to determine the position of the car (Figure 4). In this section the most common hardware components of a car navigation system are described. Of course this complexity in the systems means also an increase in the cost compared to aftermarket or portable systems. It is important also to mention here that many of these hardware components are not used exclusively by the navigation systems and many are used for other systems in the vehicle and are independent electronic control units (ECU) communicating with each other using mainly a controlled area network (CAN) bus.

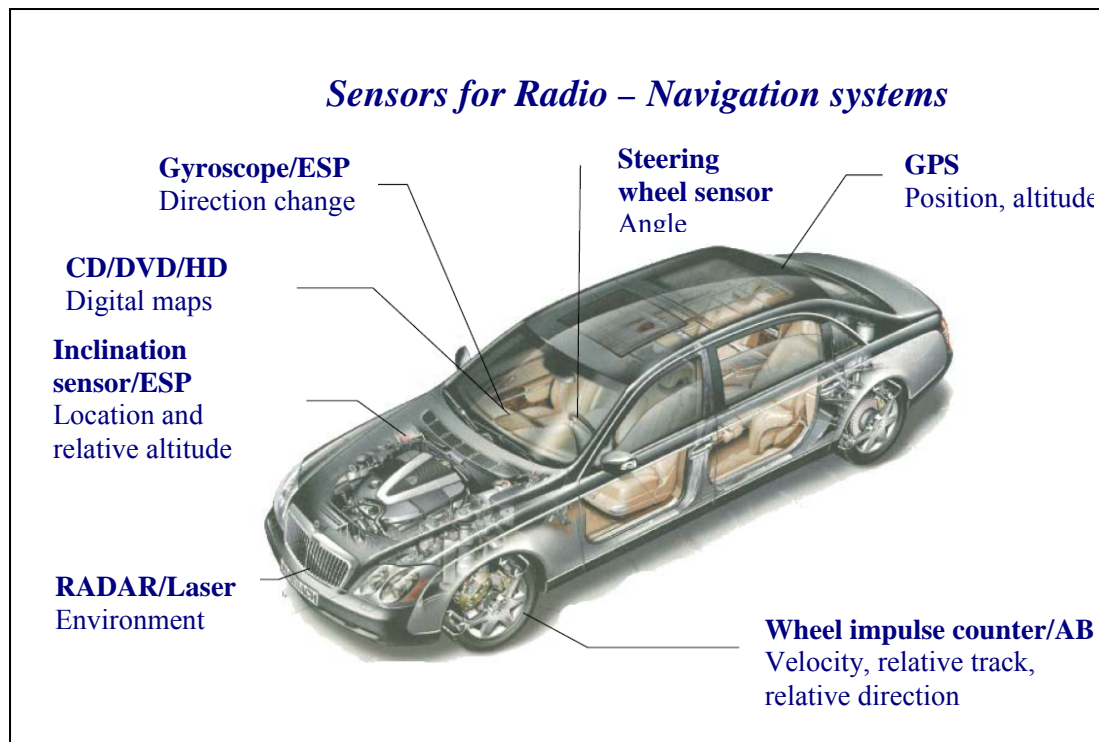


Figure 4: Hardware components of a car with modern navigation systems.

2.1 GPS

The global positioning system (GPS) is used in the navigation systems to determine the current car position. The US defense ministry developed this system in 1973 and it is based on satellite signals from satellites around

the word, a control segment based on a master control station located in Colorado Springs and four other stations around the world, these stations are charged with the complete control of the system; the final component of the GPS is the GPS receiver. All the three components, working through mathematical manipulations make it possible to determine the position of an object on the earth.



Figure 5: The gyroscope.

2.2 Gyroscope

The changes in direction are detected by the gyrometer (Figure 5). This component transforms the changes of position into electrical impulses which the computer analyses to know if the car accelerates, stops or changes direction. The advantage of this component is the small size and the minimum influence on the distortion fields present in the car. This component is very important when GPS signal is lost, for example, in a tunnel. Using the information from the gyrometer the system can compare the route of the car and the map information to continue with the correct route guidance. In case the car reaches a position that is not included in the map data, the gyroscope could be used to determine the car direction and alert the driver as to the best way to return to the route indicated in the map information.

2.3 Wheel Impulse Counter/ ABS

In the first generation navigation systems, for the positioning of the navigation system, it was very usual to use two inductive wheel rotation counters to determine the route and change in the direction of the vehicle, also the earth magnetic field sensor was used to determine the absolute driving direction. The GPS signals were used more commonly to correct strong sensor distortions or after driving in a ‘non-digitized’ map zone to return to the saved route net. But the newer systems use an easy route signal from an electronic tachometer, just as it is used in the speedy automatic volume regulation of auto radios. The tachometer sends drive-speed-proportional impulses that can be interpreted in the navigation systems.

2.4 Steering Wheel Sensor

Some systems include this component that could be used to determine the car’s direction. It is a very simple component that registers the angle in which the steering wheel is rotated and this information can be used by the navigation system to calculate the driving direction.

2.5 Inclination Sensor/ESP

This component could be used in combination with the wheel impulse counter to determine the altitude of the car. This component transmits to the system the inclination angle of the vehicle and the system calculates the relative altitude of the vehicle.

2.6 RADAR/ Laser

Nowadays many vehicles include radar and laser sensors to determine the condition on the way. These systems provide information to the driver regarding the route, as for example a vehicle in front braking hard. All these information could be shown as an alert message in the navigation system or as an acoustic alert message to the driver. Some sensors also provide information regarding the weather condition and this information is processed by the navigation system and provided to the user as alert messages.

3 Trends of the Car Navigation Systems

To show the most important trends of the car navigation systems, it is again important to make a distinction between the PNA and the systems integrated into the cars.

3.1 Trends in the PNAs

Nowadays the market is pushing the classic PNA manufactures to increase the performance and to include more features (Figure 6). Many PNAs are including photorealistic 3D visualizations. Some manufacturers include an information system called HD traffic; these systems allow the user to have the most recent information of the traffic conditions on the way and update it constantly. Other new features on the PNA are the entertainment features like music player, DVB-T tuner, photo viewer, etc. As an alternative to the PNA, new generation of smart cell phones are entering the navigation systems market; many cell phone companies are offering navigation systems within the cell phones and most of these phones have already integrated a GPS receiver. In the case of the cell phones the user needs a connection to internet via the cell phone in order to calculate the route (all map data is in storage on the server) and in this case it could carry an extra cost to the user if he does not have a flat rate connection to the internet. 'Off line' systems exist where the user receives the cell phone with the data installed or pre-installed in a storage device like the micro SD cards, in this case a connection to the server is not needed to calculate a route but it is needed if the user wants to update the map data.



Figure 6: PNA.

The real problem for the classic PNA is that there are many free navigation solutions for the cell phones and the trend is that the map data will be updated as well for free; and in the case of the PNA the user has to pay for a map data update. For example Nokia will provide the Ovi-maps for free for the users of some Nokia smart phones and Google and Apple have developed a software that converts the iPhone into a high-end navigation system.

3.2 Trends of the car-integrated systems

In case of the car navigation systems integrated in the car infotainment systems the trend is going in another direction because many of the new features on the PNA are already included in the infotainment systems. In the infotainment systems it is appreciated that the trend is that more valuable information is included in the map data, for example the POIs will contain information like pictures, videos or text information; this information could be stored in the hard disk of the systems or can be called from internet (web services). The car manufactures are introducing a concept of multiuser navigation system (multiseat); this will allow the passengers to look for information on the route or to calculate an alternative route in parallel to the route calculated by the driver.

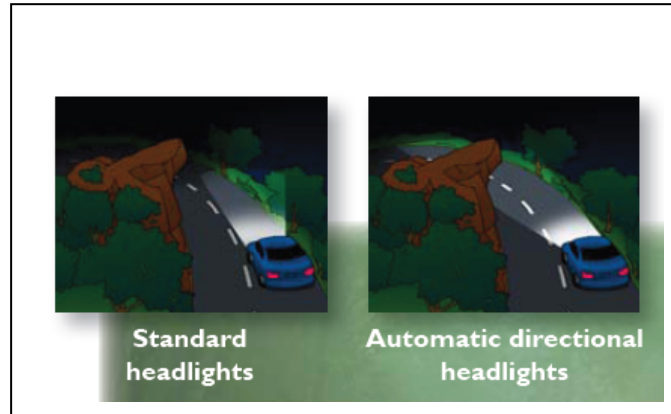


Figure 7: Headlight steering system.

The most high tech systems are going to include features that are going to provide information from the navigation system, most precisely from the map data to the other car systems. The navigation systems will support the advanced driver assistant systems (ADAS), for example the navigation will provide lane information to support lane change assistance, the navigation will provide information of the road classification to support cruise control functions or in combination with multipurpose cameras the navigation speed information could be used to improve the traffic sign recognition for speed limit.

Nowadays the integrated navigation systems are introducing more and more 3D data to the navigation data set and many Original Equipment Manufacturers (OEMs) are introducing highly accurate digital elevation models (DEM); this 3D data allows the interaction with other systems on the car, like for example the headlight steering that anticipates road curves and slopes (Figure 7); this improves dramatically the driver visibility ; this 3D data improves the adaptive cruise control systems by anticipating slopes and curves, thereby reducing false alarms from optical or radar sensors; the 3D data also assist the lane-keeping systems that anticipate threatening situations, warn the driver and in extreme situations, brake and steer the car out of danger. These DEMs have also ecological applications in the automobile industry, like for example the interaction with the power train control systems; provided with elevation information, the truck can intelligently shift its gears and make optimum use of its transmission for uphill and downhill slopes. The data allows for the selection of the most efficient gear ration and engine performance for varying grades, maintaining speed without wasting fuel. The introduction of this 3D data is also used on the hybrid electric vehicles (HEVs) because the vehicle 'knows the road ahead', the precise 3D road vector data could be used to help the HEVs anticipate upcoming grades and enable active energy management, quickly switching between power sources to conserve as much fuel as is possible.

Spatial Data Infrastructures, E-Cadastre

Spatial Data Infrastructures – Status 2010

Hardy Lehmkuhler

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
hardy.lehmkuhler@hft-stuttgart.de

KEYWORDS:

Spatial Data Infrastructures (SDI), E-Government

ABSTRACT

Like Globalization Spatial Data Infrastructures (SDI's) can be regarded as a phenomenon. They would have somehow emerged even if we hadn't given such a name to them. Despite this it makes sense to stick to our definitions and see what has emerged until now. The biannual GSIDI conferences (G for global, last in 2009) lately showed some look back which is reported and extended here. The principal result is: systems converge more and more.

After a brief introduction of SDI's with their technical components, the necessary legal and other settings a newly required orientation about workflows and cooperation using geodata in SDI's is described. It requires SDI's to change with the needs of e.g. government purposes. Such needs come up far out of the scope of SDI specialists. Providing the 3 mentioned views can be seen as passing milestones between 3 generations of SDI's in the last 15 years.

A first example then deals with the difficulties that we currently face about the multinational and hierarchical European ESDI initiative. Another example of a regional and expressively not "national" SDI shows the many benefits and potentials of cooperation

Redefinition of Coordinate Reference for GIS: Expected and Unexpected Effects

Hardy Lehmkuhler

Department of Geomatics, Computer Science and Mathematics,
University of Applied Sciences Stuttgart, Schellingstraße 24, D-70174 Stuttgart, Germany
hardy.lehmkuhler@hft-stuttgart.de

KEYWORDS:

Shift of geometric reference in GIS, Coordinate System, Geodetic Datum in GIS

ABSTRACT:

The geometric reference (gr) is an important characteristic of a GIS data set. It tells us which coordinate system is used for the coordinates of the data elements in it. In GIS the gr is usually declared for a whole data set or feature type and further on it is assumed not to change.

The gr can however definitely change when it is newly defined as e.g. currently in Germany where a new unique gr has been introduced (ETRS 89 with UTM). The gr can also change when data sets of different origins are brought together, do not match and have to be geometrically harmonized. This can obviously happen everywhere.

The presentation explains the consequences of the German case for GIS data sets. It points out the lessons that can be learned for any project with a gr shift. It explains why especially for transformations from older data sets to modern coordinate systems transformations often depend on location. Such local transformations need a special grid-based representation form. They cannot be solved by standard "on the fly"-transformation inside of GIS but require modification of each saved coordinate value. With this background the practical conduction of the gr shift is described.

It is worth knowing in advance the potential effects of such projects. Those start with deformation of shapes, include artificial edges in formerly straight lines and end for example with area values calculated from geometries to change.

Disponibilidad de información geoespacial para la investigación mediante la Implementación de una IDE¹⁷

Eddy J. De la Cruz Aquino

Dirección de Geomatica, Dirección Técnica de Aplicaciones Espaciales, Agencia Espacial de Perú, Av. Luis Felipe Villaran (Perú), - edelacruz@conida.gob.pe

KEYWORDS:

IDE, Datos Geoespaciales, Mapserver, Geonetwork

ABSTRACT:

La demanda de información geoespacial va creciendo día a día y la disponibilidad de información confiable se ve inmersa en una búsqueda bibliográfica tediosa y muy a menudo no satisfactoria; para ello la solución de la publicación de información de carácter espacial se hace necesaria para la ayuda o complemento de los diferentes temas que se pueden laborar en gabinete, para tal objetivo se emplea los servicios de Web Map Service (WMS) y Catalog Web Service (CWS) y en algunos casos los Web Coverage Service (WCS) que está mas orientado al empleo integramente de las imágenes raster.

Existen diversos medios de publicación en línea tanto licenciados como de carácter libre, la cual es de libre elección según la disponibilidad de recursos en la adquisición de una ó de otra. Para este caso hemos optado por emplear dos software el MAPSERVER actualmente soportado por OSGeo y GEONETWORK. Todas las instituciones deberían tener implementado su IDE para la agilizar su organización de información y evitar la duplicidad de trabajos en la generación de productos.

1 Infraestructura de Datos Espaciales

1.1 Objetivos de una IDE

La Infraestructura de Datos Espaciales es un conjunto de políticas propias de cada institución para el desarrollo de protocolos de libre acceso a la información espacial que se tiene en la base de datos de dicha institución, estándares y recursos tecnológicos para facilitar la obtención, uso, producción a diversas fuentes de información espacial, ello para una oportuna toma de decisiones.

El objetivo principal de una IDE institucional es prevalecer el orden de la información para satisfacer las peticiones del usuario final en la obtención de la misma con un mínimo error para las diversas aplicaciones que se le puede dar. La forma más práctica de difundir ésta información es a través del servicio de catálogo en línea y servicios web de mapas.

1.2 Implementando una IDE Institucional

La implementación de una Infraestructura de datos espaciales parte en conceptualizar en el manejo y organización de la base de datos espacial, es decir, modelar la base de datos según las necesidades básicas que tiene la institución productora de información.

El Modelado de las bases de datos relacionales parte por conocer y discriminar el tipo de información que se posee ya que no es de buena practica tener información textual junto con la información espacial dentro de una misma tabla, para ello se emplea diferentes tablas ó diferentes bases de datos según la complejidad y el volumen que se proyecta a manejar, todo modelado se asemeja a la realidad y se debe tener en cuenta que va a crecer en el tiempo y para ello tiene que ser un modelo escalable, es decir, que al insertar mas tablas dentro de una base de datos o afectará el flujo de información entre la misma, en el caso de tener registros dentro de un programa tipo hoja de cálculo se debe aplicar las formas de normalización para evitar redundancias y duplicidad de información así como un espacio innecesario en el almacenamiento de la información. Cuando ya se ha implementado las tablas con un número

¹⁷ Basado en experiencias universitarias del área de ingeniería en desarrollo de trabajos de investigación.

suficiente de registros llenos se procede a realizar las pruebas funcionales para determinar si la distribución de la información es la correcta o se debe realizar algunas modificaciones en la estructura de la misma, para finalmente completar el llenado de los registros de toda la información disponible en la institución proveedora de información.

Los pasos para lograr una IDE institucional escalable y de buen rendimiento se puede confiar en la utilización de software libre y programación manual la cual evita fallos no reconocibles, el procedimiento aplicado es la siguiente:

- Modelamiento de la base de datos espacial y no espacial.
- Implementación de un servidor y sus programas necesarios para la publicación.
- Implementación del modelo de la base de datos en el sistema web.
- Programación de los archivos configurables (WMS, WFS y WCS)
- Instalación del software cliente de los servicios publicados.
- Llenado de Metadatos y harvesting con otras instituciones.

1.3 Software Necesario Para la Implementación

Para obtener toda la información publicada en Internet existen varios programas libres que ayudan de forma práctica el funcionamiento de la IDE, en este caso se empleó los siguientes programas libres:

- DB Designer (Modelamiento de Base de Datos)
- Apache (Servidor Web)
- PostgreSQL (Manejador de Base de Datos)
- PostGIS (Modulo geoespacial para el PostgreSQL)
- PHP (Lenguaje de Programación orientado a objetos)
- MapServer (Software de publicación de mapas y de WMS, WFS, WCS)
- Geonetwork (Catálogo en línea basado en ISO-19139 - Metadatos)
- GvSIG (Software de escritorio y de aplicación SIG, cliente de servicios WMS, WFS, WCS, GeoDB)

2 CASO GEOMATICA – CONIDA²

En la Comisión Nacional de Investigación y Desarrollo Aeroespacial CONIDA se implementa una IDE para solucionar parcialmente la demanda de información satelital por parte de la comunidad universitaria, municipios así como otras instituciones que requieran la necesidad de este tipo de información, para ello se pone en marcha la IDE con la aplicación del MAPSERVER y GEONETWORK que son los paquetes de mayor uso por su fácil publicación en línea y bajos costo de implementación, para ello se cuantifica la información actual y se proyecta el crecimiento anual para modelar la base de datos relacional.

¹ Trabajo desarrollado por Eddy De la Cruz A.

CONIDA tiene un almacenamiento aproximado de 6 Terabytes y un crecimiento de 0.5 TB anual tanto por producción así como adquisición de información tanto espacial como documentaria.

Teniendo esto como base se proyecta a realizar el modelado de datos, distribución de la información separando la documentaria de la espacial así mismo la información vectorial es almacenada en el módulo PostGIS de PostgreSQL que es una buena opción para agilizar la visualización y manejo de grandes volúmenes de información.

La publicación realizada en la institución se ve afectada por una semiredundancia ya que se ha implementado dos formas de almacenar Metadatos que es la parte esencial de toda información espacial y no espacial para una búsqueda eficaz, las ISOS 19125 – 19139 son parte de la vértebra de la IDE, la ISO 19125 es la que almacena alrededor de 300 campos de metadatos para cada imagen, geodatabase, texto ó información importante en almacenamiento, la ISO 19139 es la estructura de

cómo publicar dicho metadato en la cual actualmente existe una decadencia en la parte de información satelital ya que no cubre los campos necesarios para una consulta técnica o de precisión por lo cual se opta en la implementación aparte de registros especiales para las imágenes de satélite en el servidor de Mapas de CONIDA es por ello que se dice casi redundante ya que comparten varios campos de información repetida dentro del Catálogo Geonetwork como en la búsqueda de imágenes dentro de la base de datos.

Este servicio de publicación de información se continúa desarrollando para ampliar las aplicaciones y utilidades a estos servicios mejorando e incrementando el potencial que brinda la información geoespacial. Toda la implementación se ha realizado con software libre bajo la plataforma Linux para un mejor desempeño y mayor control de la información puesta en línea, para acceder a estos servicios bastará con el llenado de una pequeña ficha de registro para la obtención de su usuario y contraseña en el ingreso al Servidor de Mapas.

Para las pruebas y manipulación de la información vectorial se emplea el Software GvSIG desarrollado por Iver Technologies de origen español.

2.1 INCOMPATIBILIDADES QUE SE PUEDEN PRESENTAR

- En la configuración también existen incompatibilidades con versiones anteriores versus actuales, la decisión de que componentes se necesitan instalar depende de la versión con la cual se piensa trabajar.
- Como es de esperarse siempre existen inconvenientes tanto en la implementación como en el mantenimiento de los servicios, un problema principal es la ausencia de estándares para la información existente en la institución, no hay reglas de escritura para nombrar los archivos, carpetas o discos de almacenamiento.
- La información no es manipulada por una única persona, es muy frecuente que una información vectorial sea manejada por una persona y modificada y luego guardada con otro nombre porque posee un objetivo específico pero no se informa que ya existen dos versiones de la misma información y no se puede determinar cual se debe almacenar y cual de debe borrar o en su defecto guardar las dos informaciones como complemento una de otra.
- La información vectorial no es almacenada correctamente, este es otro problema muy frecuente debido al inadecuado almacenamiento sin los requisitos mínimos que se deben tomar en cuenta como que la información tenga sistema de proyección y se encuentra en coordenadas reales.
- Inexistencia de archivos de Metadatos que puedan facilitar la información almacenada.

2.2 SOLUCIONES SATISFACTORIAS

- En la plataforma Linux tomar en cuenta los repositorios disponibles para las versiones actuales, lo recomendable es no utilizar la última versión sino al anterior a la última ya que siempre la última aun no ha sido probada en su totalidad y posee algunos errores que aun se encuentran en desarrollo.
- Proponer estándares para la codificación de cada clase de información que se va disponer en el servidor implementando un sistema de codificación y etiquetado para los discos de almacenamiento.
- Comunicar la forma de trabajo para evitar redundancias de información y duplicidad de trabajos en el procesamiento de la información vectorial, en el mejor de los casos guardar la información con un archivo de metadato para identificar que es un producto o secuencia de otra información.
- Implementar flujo de trabajo para la información vectorial de tal modo de asegurar que toda la información almacenada cuenta con los procesos mínimos y necesarios para trabajos futuros.
- Generar Metadatos obligatoriamente para grupos de archivos e información importante o de mayor demanda de consulta.

3 MODELADO E IMPLEMENTACION DE BASE DE DATOS ESPACIALES

El modelado se genera a partir de la necesidad de ¿cómo obtener la información y de que manera deseo visualizarla?, para ello pensamos en lo siguiente:

- a) Debe tener un identificador único en caso de que existan información de igual características pero con un elemento que varía como es el caso de igual tipo pero de diferente fecha como indica la figura siguiente.

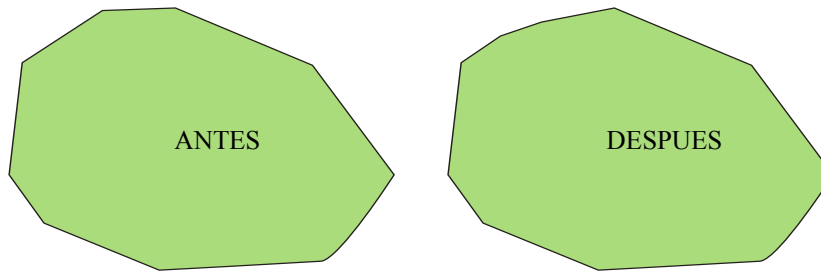


Fig. 1: Un área es de tipo poligonal pero en el tiempo puede variar su forma y permanece del mismo tipo.

- b) Contar con los atributos necesarios para no confundir la información vectorial con la raster y/o textual, entendamos que un atributo es la característica especial del objeto de quien se trata de describir, se debe tomar en cuenta la longitud del texto a insertar y prever posibles nombres largos.
- c) Determinar los tipos de relaciones entre las tablas para su óptimo funcionamiento en las consultas anidadas complejas, satisfaciendo todas las consultas que se realizan cotidianamente.
- d) Instalar el PostgreSQL para la implementación del modelo realizado e insertar por lo menos 30 registros para realizar las pruebas necesarias para su funcionamiento.

³ Modelo desarrollado e implementado por Eddy De la Cruz A.

- e) Al instalar esta base de datos se debe decidir cuantas personas se piensa que administrarán la base de datos, dado que unos cuantos ingresan datos, otro tanto actualiza; para ello se debe decidir el tipo de rol para un grupo de usuarios a fin de evitar errores que pueden dar inestabilidad a nuestra base de datos como es el caso de que por accidente se borre un registro que va anidado a otro en otra tabla y al realizar consultas no encuentre dicho registro y nos producirá un error y tomará tiempo solucionar adecuadamente la base de datos que se trabaja.
- f) Las consultas correctas garantizan el buen funcionamiento de la base de datos relacional así como la búsqueda eficaz de la información almacenada.
- g) Al finalizar todo esto como parte de las consultas espaciales se procede a la instalación del módulo PostGIS para el ingreso de datos geoespaciales.

4 INSTALACION DEL SERVIDOR WEB – PHP Y MAPSERVER⁴

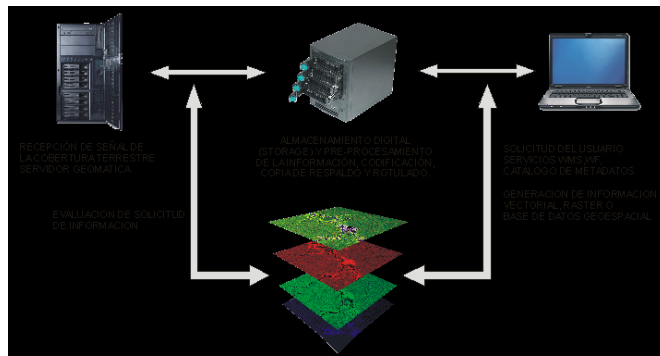
Al culminar la implementación de la base de datos relacional es necesario contar con un servidor web para poner en funcionamiento las aplicaciones de consulta de toda la información ingresada en intranet ó Internet; el servidor que se emplea en este caso es APACHE, el cual su instalación es practica y de fácil manejo pero este servidor en plataforma LINUX se debe configurar con soporte de módulos “.so”, este tipo de instalación se puede ubicar en Internet.

Seguidamente es necesario instalar un lenguaje de parte del servidor, es decir, que se genere en el servidor y se muestre el resultado en el lado del cliente (usuario) para ello hemos elegido el lenguaje PHP HyperText Processor para emplear el PHP MAPSCRIPT como módulo ejecutor para la

publicación de la información geoespacial almacenada en la base de datos. Al instalar el MAPSERVER en Linux se deberá compilarlo con opciones de soporte postgres, postgis, php y apache para que funcione correctamente todo, para mayor información deojo un link a un manual de instalación recomendada y probada con la penúltima versión del MAPSERVER, al finalizar la instalación podremos ejecutar un ejemplo simple de PHP MAPSCRIPT¹⁸ el cual nos dirá si está instalado el soporte para php_mapscript.

```
<?php
if (PHP_OS == "WINNT" || PHP_OS == "WIN32")
{
    dl("php_mapscript.dll");
}
else
{
    dl("php_mapscript.so");
}
phpinfo();
?>
```

Al finalizar la instalación de estos paquetes se procede a configurar los archivos map los cuales carga toda la configuración para mostrar con los servicios wms, wfs, wcs y conexión geodatabase, el siguiente cuadro muestra un archivo map básico para la conexión necesaria, una conexión a base de datos y una para archivos shp.



Map File¹⁹

Map → Sin esta entrada el Mapserver no reconocerá que es un archivo map
extent -203119 7800000 1190961 10120000 → limite X mínimo, Y mínimo, X máximo, Y máximo
imagecolor 255 200 240 → color que mostrará la imagen de fondo
imagetype png24 → tipo de imagen de salida
name "Servidor de Mapas" → nombre del archivo
shapepath "/home/usr/local/apache2/htdocs/mapserver/shapes/" → ruta completa donde se almacena los archivos shp
size 800 500
status on
units meters
projection → es necesaria para la proyección al vuelo y combinar varias capas en diferentes zonas
"init=epsg:32718"
end

¹⁸ Manual referencial de http://www.cartografia.cl/download/compilacion_mapserver_postgis.pdf.

¹⁹ Configuración empleada en el MapServer de CONIDA.

WEB → no es necesaria salvo en casos de utilizacion en modo web

END

layer

```
connectiontype postgis
name "Curvas de Nivel"
connection "user=postgres password=***** dbname=postgis host=localhost"
data "the_geom from mi_tabla"
status on
labelcache on
type line
class
color 196 110 48
end
projection
"init=epsg:32718"
end
```

end

layer

```
data "/home/usr/local/apache2/htdocs/mapserver/shapes/depart.shp"
labelcache on
labelitem "DEPNOM"
name "Departamentos"
status on
type polygon
class
name "Departamentos"
outlinecolor 0 128 128
label
color 0 128 255
position cc
size small
```

5 INSTALACION DEL GEONETWORK - GVSIG

Al finalizar con la parte de publicación de mapas, también es necesario tener un catálogo en línea para los metadatos generados de toda la información el cual será de uso extensivo para la búsqueda de información, así mismo su instalación es simple no requiere ningún conocimiento especial, para mayor información es anexada en la bibliografía.

El GvSIG juega un papel importante debido a que es software libre el cual se puede utilizar sin ningún problema y de bajo costo, esta herramienta es empleada para el manejo de los servicios WMS, WFS, WCS y GeoDatabase que es la información almacenada en el módulo PostGIS, con ello podremos adquirir otros servicios nacionales como internacionales a fin de continuar con la investigación y tener la mayor información disponible y con los estándares que garanticen la calidad de la información.

Bibliografía

DOCUMENTACIÓN CONIDA. <http://200.37.254.55:8080>, Software y manuals_Último acceso 05/07/2010

Comité Coordinador de la Infraestructura de Datos Espaciales del Perú - CCIDEP

<http://www.ccidep.gob.pe/>. Nodo de portales institucionales con visores de mapas. Último acceso 05/07/2010

INSTALACIÓN MAPSERVER HISPANO.

http://www.cartografia.cl/download/compilacion_mapserver_postgis.pdf. René F. Viancos S. rviancos@gmail.com. Último acceso 05/07/2010

BEGINNING MAPSERVER: OPEN SOURCE GIS DEVELOPMENT. Bil Kropla. 448 pp. Publicado en Agosto 2005

WEB MAPPING ILLUSTRATED. Tyler Mitchell Publicador: O'Reilly Media, Inc. Fecha de Publicación: Junio 17 del 2005 368 pp.

Modelo de integración y preprocesamiento de información para el descubrimiento de conocimiento en bases de datos federadas utilizando ontologías

Geovanna Evelyn Espinoza Taype

Universidad Católica de Santa María, Programa Profesional de ingeniería de sistemas,
evelynzyz8@gmail.com

KEYWORDS:

Descubrimiento de conocimiento en bases de Datos (KDD), Minería de Datos, Bases de Datos Federadas, Administración de Bases de datos, Knowledge discovery in data bases (KDD), Data Mining (DM), Federated Databases, Systems Management Database (SGBD).

ABSTRACT

El aumento de la información en las empresas ha provocado un incremento de datos en sus sistemas de información, sin embargo no ha implicado un aumento proporcional del conocimiento disponible. Para obtener este conocimiento es necesario avanzar en el proceso de descubrimiento de conocimiento en bases de datos. El presente modelo propone el uso de ontologías en las fases previas a la minería de datos dentro de un proceso de descubrimiento de conocimiento. Una de las ventajas más importantes de este enfoque es poder acceder de forma homogénea a información de varias fuentes heterogéneas unificadas y semánticamente coherentes.

The increase of information in companies has caused an increment of data in her information system, however these hasn't involved an increase proportional of knowledge available, for get these knowledge is necessary improve in the process of knowledge discovery in databases. The present model propose the use of ontology's in the first phases in the data mining into to process of knowledge discovery. The advantage more important is the Access to databases heterogeneous through a Homogeneous interface with databases unified and coherent semantically.

6 INTRODUCCIÓN

Durante los últimos años se ha producido un aumento exponencial de la información que producen las organizaciones, lo que ha impulsado el desarrollo de soluciones cada vez más eficientes para su almacenamiento. Sin embargo, el conocimiento utilizado en los procesos de toma de decisiones no ha experimentado un incremento proporcional al de la información disponible. El descubrimiento de nuevo conocimiento impulsa, por tanto, la investigación de este modelo, con la aplicación en métodos avanzados de búsqueda, integración y procesamiento de información. Esta investigación es patrocinada por Microsoft Student Partners y Bitcomm Research Center, el cual propone un modelo para la integración y preprocesamiento de información en bases de datos federadas proponiendo la utilización de Ontologías en las fases previas a la minería de datos dentro de un proceso de descubrimiento de conocimiento.

7 MATERIAL Y MÉTODOS

7.1 DEFINICIÓN DEL PROBLEMA

El acceso a grandes cantidades de datos, con diferentes esquemas, formatos y contextos ha impulsado el desarrollo de soluciones para su almacenamiento, Sin embargo el conocimiento utilizado en los procesos de toma de decisiones no ha experimentado un incremento proporcional al de la información disponible, Es por tal motivo el descubrimiento de nuevo conocimiento, impulsa la investigación en métodos avanzados de búsqueda, integración y preprocesamiento de información.

7.2 OBJETIVO GENERAL

Plantear un modelo de integración y preprocesamiento de información para el descubrimiento de conocimiento en base de datos federadas utilizando ontologías.

7.3 OBJETIVOS ESPECÍFICOS

- Desarrollar un método para la homogeneización de esquemas heterogéneos de las diferentes bases de datos utilizando Ontologías.
- Desarrollar artefactos virtuales para la integración de información.
- Desarrollar un método para la integración de esquemas a través del uso de ontologías.
- Desarrollar y mejorar métodos para solucionar problemas de inconsistencias de datos.
- Precisar un modelo adecuado que permita integrar fuentes de datos y ofrecer una interfaz homogénea a los usuarios.

8 HIPÓTESIS

Es probable que el uso de ontologías para la integración, clasificación y el procesamiento de inconsistencias en las fases previas a la minería de datos mejore la calidad de los patrones obtenidos mediante un proceso de descubrimiento de conocimiento o KDD.

9 MODELO DE INTEGRACIÓN Y PROCESAMIENTO DE LA INFORMACIÓN

El modelo propuesto para la integración y preprocesamiento de la información se basa en la utilización de ontologías para alcanzar la mayor expresividad posible. Las ontologías se utilizan como mecanismos para resolver diversos tipos de heterogeneidades presentes, incluyendo las semánticas en fuentes estructuradas, para manejar las relaciones entre los datos. El modelo propuesto aborda la integración en dos niveles:

- (i) Integración de esquemas, en la que se centran la mayoría de los proyectos de investigación en el área, y donde se suelen construir nuevas vistas virtuales de las bases de datos físicas.
- (ii) Integración de instancias y preprocesamiento, donde se realizan las transformaciones necesarias de los valores contenidos en las bases de datos para homogeneizar los resultados. En este último caso se incluyen también cuestiones del preprocesamiento de datos, ya que aunque son fases que no tienen por qué abordarse al mismo tiempo, comparten objetivos desde un punto de vista funcional.

El modelo propuesto se guiará en base al siguiente esquema, y tomando como punto de partida, véase la figura 1 donde se muestra el modelo general propuesto para la integración y preprocesamiento de bases de datos federadas.

Este modelo incluye la integración y preprocesamiento de la información y tiene como entradas:

- (i) Fuentes de datos,
- (ii) Consultas de los usuarios y
- (iii) Ontologías que dirigen las transformaciones.

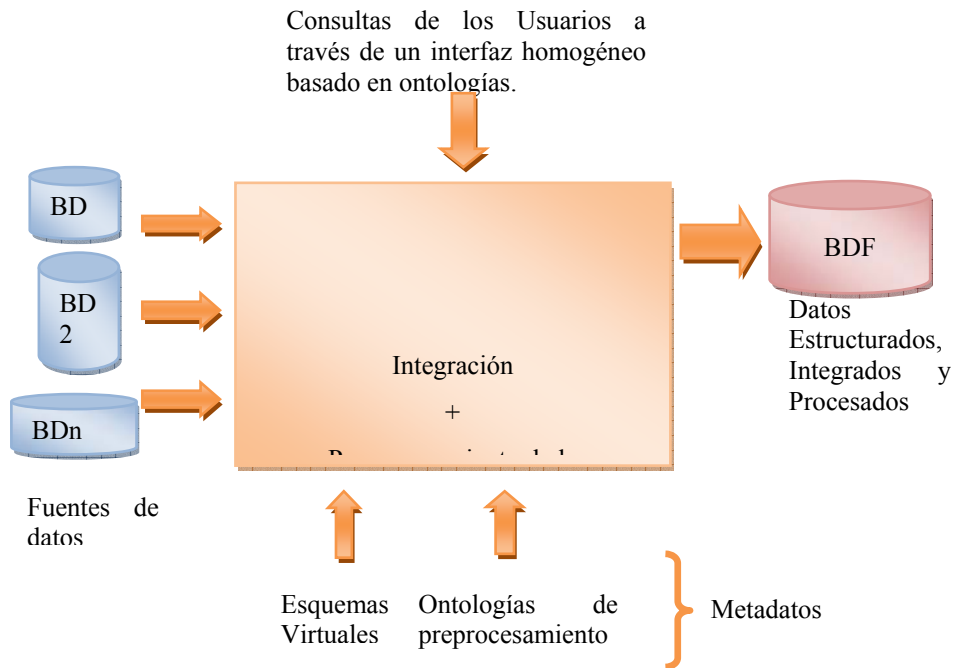


Figura 1. Modelo propuesto de KDD basado en Ontologías

Este modelo incluye la integración y preprocesamiento de la información y tiene como entradas:

- (i) Fuentes de datos,
- (ii) Consultas de los usuarios y
- (iii) Ontologías que dirigen las transformaciones.

Como salida devuelve una base de datos general con datos integrados y preprocesados de forma transparente al usuario.

Para conseguir este objetivo, tanto en la integración de esquemas como en la de instancias, se propone una metodología que consta de tres pasos secuenciales:

- (i) La integración de esquemas,
- (ii) la detección de inconsistencias de una o varias fuentes de datos, en un proceso manual y
- (iii) la resolución de dichas inconsistencias automáticamente.

En la Figura 2 que se muestra a continuación, se puede observar esta metodología de tratamiento de inconsistencias.

La detección de inconsistencias en la integración de esquemas se lleva a cabo de forma manual mediante la construcción de esquemas virtuales que representan virtualmente los datos estructurados. La integración de instancias se lleva a cabo mediante la generación de ontologías de preprocesamiento asociadas a cada esquema virtual, que posteriormente se incluyen en el proceso de unificación.

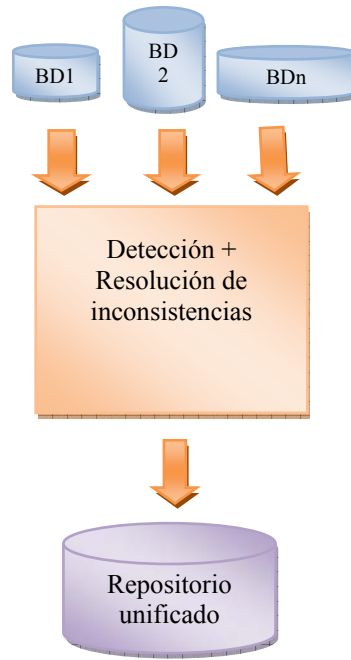


Figura 2. Tratamiento de inconsistencias.

9.1 UNIFICACIÓN DE REPOSITORIOS VIRTUALES

Una vez detectadas las inconsistencias de las distintas fuentes de datos, y almacenadas en sus correspondientes ontologías, el proceso de su unificación es automático. Consiste en la fusión de distintos esquemas virtuales correspondientes a varias fuentes de datos en un esquema virtual unificado. Estos esquemas virtuales unificados son ontologías que reflejan la estructura conceptual de la información almacenada en varias bases de datos de forma transparente al usuario. La Figura 3 ilustra el funcionamiento general de los procesos de “mapping” y unificación, así como los elementos que intervienen.

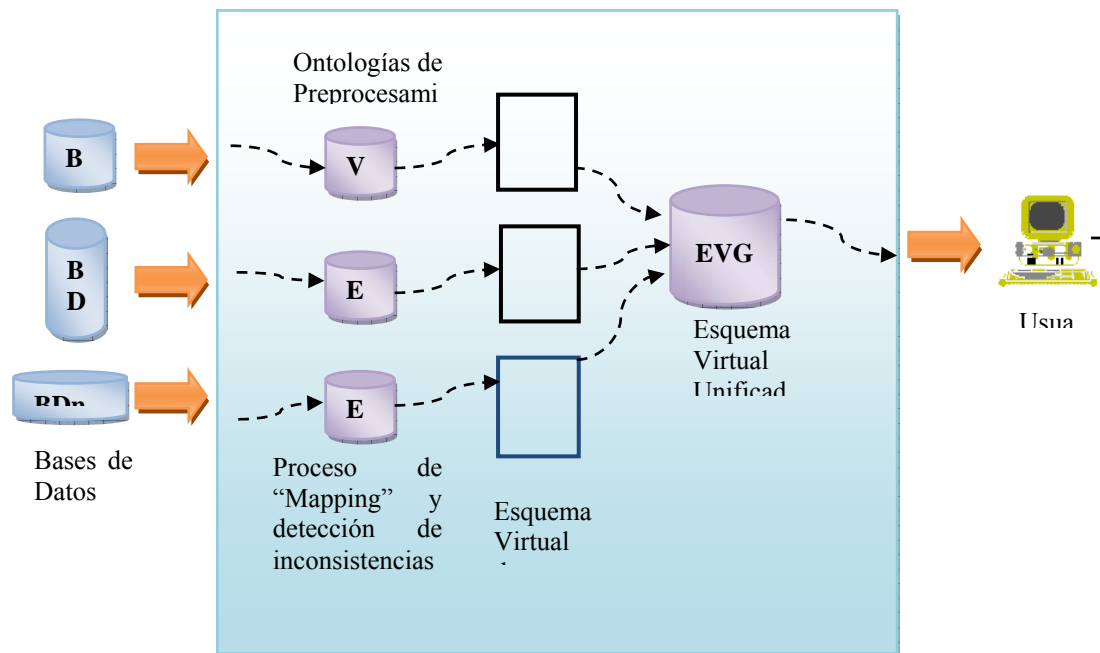


Figura 3. Procesos de integración y preprocesamiento.

La primera fila desde abajo representa las fuentes de datos físicas incluidas en el modelo. La siguiente representa el proceso de “mapping” y la detección de inconsistencias dando lugar a las ontologías de preprocesamiento y los esquemas virtuales de primer nivel. En la parte superior de la figura se puede observar el proceso de unificación que integra varios esquemas virtuales en nuevos esquemas virtuales unificados, permitiendo una jerarquización en varios niveles. Con este enfoque se pueden crear distintos esquemas virtuales, que cubran necesidades específicas de datos.

Para integrar fuentes de datos en la solución propuesta, es necesario que los esquemas virtuales se hayan creado utilizando una ontología de dominio común. Para identificar qué conceptos deben ser unificados, el algoritmo empleado comprueba la ontología de dominio. Si dos o más términos están enlazados con el mismo concepto de la ontología de dominio, se unifican, incluyendo los atributos asociados. Con este método de unificación, el administrador tan sólo ha de seleccionar un conjunto de esquemas virtuales y la ontología de dominio común que se ha utilizado para su desarrollo. El proceso posterior de unificación es completamente automático. Las ontologías de preprocesamiento encargadas de tratar las inconsistencias de instancia, son transparentes a este proceso de unificación, ya que se asocian a cada esquema virtual de primer nivel.

9.2 INCONSISTENCIAS DE ESQUEMA: ESQUEMAS VIRTUALES

El objetivo de la integración de esquemas basada en ontologías es resolver las heterogeneidades estructurales y semánticas entre los esquemas de las fuentes de datos. Para ello, las similitudes entre distintos conceptos deben ser reconocidas y expuestas a los usuarios. Los esquemas virtuales son ontologías que representan la estructura de la información contenida en una base de datos dada a nivel conceptual, es decir, cada concepto en una base de datos física se enlaza con un concepto de una ontología específica de una ontología de dominio. La finalidad es facilitar su posterior unificación con los demás repositorios virtuales en un entorno homogéneo, produciendo nuevos esquemas virtuales unificados, a los que puedan acceder los usuarios finales a través de una interfaz uniforme, intuitiva y basada en ontologías. El proceso de “mapping” de repositorios virtuales se lleva a cabo de manera manual. Aunque existen proyectos que investigan la automatización de este proceso. Para llevar a cabo el desarrollo ideal de la asociación de ontologías se requiere la necesidad de supervisión de un experto en el dominio de los datos, que será el encargado de identificar las relaciones entre los elementos de la base de datos y la ontología de dominio.

Los esquemas de una base de datos no poseen un entorno conceptual, por lo que los usuarios no pueden interactuar con el significado subyacente a los datos. Hacer el “mapping” de una nueva base de datos es un proceso que enlaza un esquema físico con un marco conceptual. Cuando se desarrolla un nuevo esquema virtual, los nombres de conceptos no se pueden elegir arbitrariamente. Deben seleccionarse de la respectiva ontología de dominio, específicas para cada campo y que suministran el mencionado marco conceptual. Además, proporcionan los enlaces semánticos entre distintos esquemas, guiando al experto en los datos en el proceso de construir nuevos esquemas virtuales. Las ontologías de dominio actúan como vocabularios estandarizados e incluyen jerarquías, conceptos, atributos y relaciones, y dependencias entre ellos. El objetivo de una ontología de dominio es garantizar que sólo conceptos generales y nombres ampliamente aceptados se utilicen para describir los elementos de un esquema virtual. También dirigen el proceso de unificación de esquemas virtuales, ya que las ontologías aseguran que todos los elementos de distintos esquemas, pero semánticamente equivalentes, utilicen el mismo nombre de concepto. Seguidamente se muestra el proceso de construcción de un esquema virtual con los elementos involucrados.

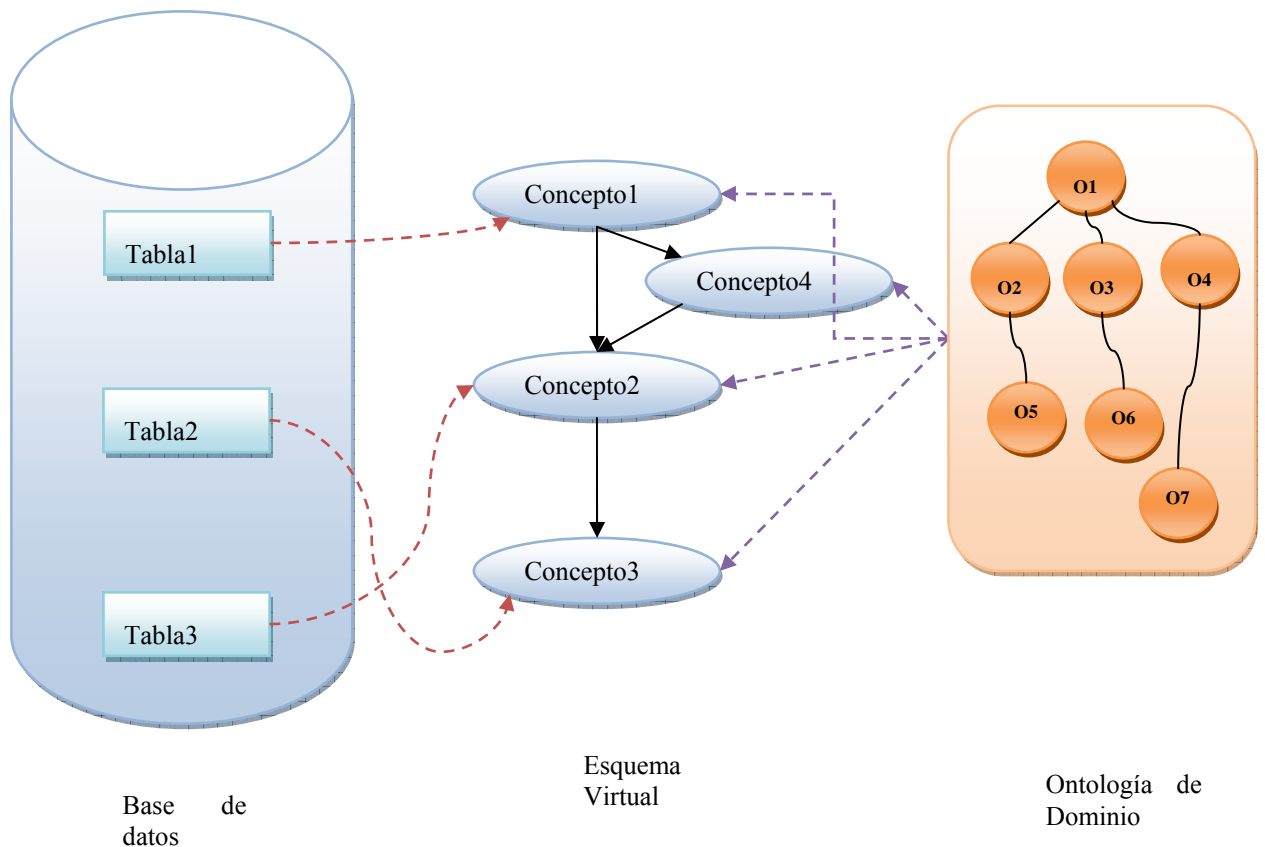


Figura 4. Construcción de ontologías de esquemas virtuales

Como se muestra en la Figura 4, los elementos de los esquemas virtuales se obtienen de una ontología de dominio adecuada. Las tablas y atributos del esquema de la base de datos física se enlazan a conceptos, atributos y relaciones virtuales. Esto forma un esquema virtual definido por los usuarios que suministra una nueva vista de una base de datos física, lo que permite a su vez modificar el esquema de la fuente de datos si es necesario. En la siguiente sección se describen las posibilidades de este rediseño de bases de datos, pasando a continuación a describir las distintas metodologías de construcción de esquemas virtuales dependiendo de la ontología de dominio utilizada.

9.3 DESARROLLO DEL MODELO INTERNO

Para realizar este modelo el cual crea un esquema virtual se hizo uso de recursos de software: como gestores de base de datos Sqlserver, Mysqserver y Postgres, entorno de desarrollo integrado Visual estudio, y el Sistema Operativo Windows Vista. Además se hizo uso de Datos almacenados en bases de datos heterogéneas. En la figura 5 muestra el esquema interno del modelo, donde tenemos como entradas bases de datos con sus respectivos gestores en un proceso de extracción, procesamiento e integración que se realiza en el servidor.

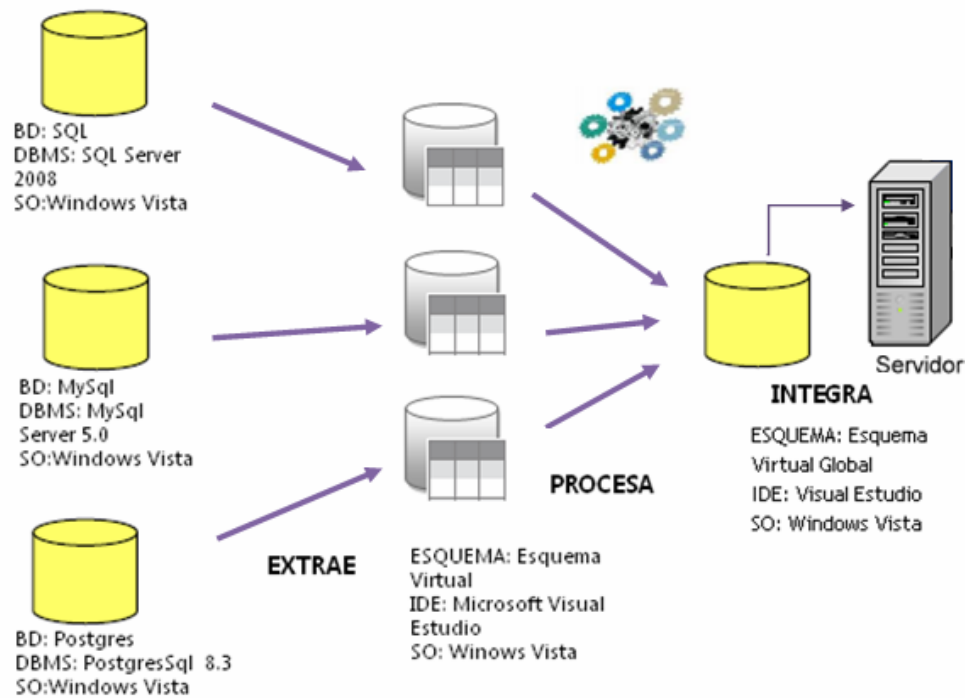


Fig. 5 Desarrollo del modelo interno.

10 RESULTADOS

Se ha observado que las distintas baterías de consultas ejecutadas sobre la unificación virtual de fuentes heterogéneas devuelven datos correctamente integrados y preprocesados. Lo que demuestra la viabilidad del modelo propuesto basado en ontologías. Aparte del hecho de que las instancias de datos recuperadas sean correctas con respecto a las fuentes. Este proceso prueba la fiabilidad del modelo propuesto de “mapping” y unificación. Sin embargo, de las pruebas también se ha concluido que no es viable integrar fuentes de datos de dominio muy distintos, ya que las unificaciones obtenidas serían demasiado genéricas, y se podría perder la correspondencia entre los esquemas virtuales que describen los datos y los datos físicos. La mayor ventaja en todo caso, es poder acceder de forma homogénea a la información de varias fuentes heterogéneas unificadas.

11 CONCLUSIONES

- El planteamiento de un modelo de integración y preprocesamiento de Información ha permitido integrar esquemas e instancias de datos que ha permitido obtener datos de calidad.
- El método para homogenizar los esquemas heterogéneos ha permitido la obtención de un esquema general global el cual funge como meta modelo.
- El desarrollo de artefactos virtuales para la integración de la información ha permitido desarrollar un enfoque híbrido, teniendo por cada base de datos un esquema virtual particular, el cual en su integración ha dado como resultado el esquema virtual global.
- Las Ontologías juegan un papel muy importante en la integración de bases de datos federadas, a esta afirmación se ha llegado tras la creación, implementación y evaluación de un nuevo modelo de KDD basado en ontologías, centrado en el preprocesamiento e integración de esquemas e instancias de fuentes de datos heterogéneas.
- Para obtener mejores resultados en el proceso de KDD se requiere tener datos de calidad, es por ello que la integración de instancias utilizando ontologías nos permite obtener datos consistentes y de calidad.

- Debemos darle mayor importancia a las primeras etapas de KDD debido a que en esta fase se obtiene la esencia de los datos y de esto dependerá las futuras decisiones que se lleguen a tomar en un proceso de Data Mining.

AGRADECIMIENTOS

Microsoft Student Partners y Bitcomm Research Center, por su apoyo y por las licencias para el desarrollo del modelo. Gracias a Supervisor: Ing. Karim Guevara Puente de la Vega, Ing. Percy Huertas Niquen.

12 REFERENCIAS Y BIBLIOGRAFÍA

- [BAL02] Baluarte, Araya César. Tecnología de Información (texto universitario). 2002.
- [GHJ97] Gary W. Hansen, James V. Hansen. Diseño y Administración de bases de datos. Segunda Edición. ISBN: 0-13-308800-6. Madrid - 1997.
- [OGX05] Oscar Nigro Héctor, González Císaro Sandra, Xodo Daniel. Ontologías en el Proceso de Descubrimiento de Conocimiento en Bases de Datos. INTIA- UNICEN. [Publicación] Argentina - 2005.
- [PDD07] Pérez del Rey David. Sobre Un modelo de integración y preprocesamiento de información distribuida basado en ontología. [Doctorado], Madrid - 2007.
- [RAM02] Ramez A. Elmasri Shamkant B. Navathe Fundamentos de sistemas de bases de datos. Tercera Edición - Pearson Education S.A, Madrid - 2002. ISBN: 84-7829-051-6.
- [RSG05] Reyes Saldaña José Fernando, García Flores Rodolfo, El proceso de descubrimiento de conocimiento en bases de datos. [Postgrado], 2005.
- [SPA03] Saavedra Places M. Ángeles. Arquitectura para Federación de Bases de Datos Documentales basada en Ontologías. [Doctorado] A Coruña - 2003.
- [VSJ06] Vallejos Sofia J. Diseño y Administración de Datos. Universidad Nacional del Nordeste [Publicación] Argentina – 2006
- [CJ008] Carbonell Javier . Tecnologías Semánticas: situación y tendencias actuales. [Publicación] 2008. <http://www.lacofa.es/index.php/tag/ontologias>
- [AMA07] Abián Miguel Ángel. 1a eb Semántica: Además De Metadatos (Ontologías), Lógica Y Confianza. [Publicación] 2007. <http://www.wshoy.sidar.org/index.php? - la-web-semántica-metadatos-ontologías-lógica-y-confianza>.
- [WIK09] Wikipedia. Ontología Informática. [Publicación] 2009. [http://es.wikipedia.org/wiki/Ontolog%C3%ADA_\(Inform%C3%A1tica\)](http://es.wikipedia.org/wiki/Ontolog%C3%ADA_(Inform%C3%A1tica)).
- [AMA05] Abián Miguel Ángel. Ontologías: Qué Son Y Para Qué Sirven. [Publicación] 2005. <http://www.wshoy.sidar.org/index.php?>.
- [MFL06] Molina Félix Luis Carlos. Extracción de Conocimiento en Bases de Datos. ICMSC – [Publicación]. 2006. <http://www.lsi.upc.es/~lcmolina/SC/html/paper/unam98.pdf>
- [MMM06] Marín Rafael, Begoña Martínez, Merenciano Josep M., Miramón David, Pérez Gema, Valentín Lluís. El desarrollo de una ontología a base de conocimiento enciclopédico parcialmente estructurado. [Publicación] España – 2006. <http://www.fiv.upv.es/jotri/Ponencias/Desarrollo.pdf>
- [ECO06] Espinoza Neira Lucero, Carpio Apaza Eliana, Ortiz Garate Marcos. Modelo De Integración De Bases De Datos Federadas. [Publicación] 2006.
- [SGJ07] Sudarshan Chawathe, Garcia Molina Hector, Joachin Hammer, Ineland Kelly. Papakonstantinou, Jeffrey Ullman, Jennifer Widom. Integrationof heterogenous

- information sources. [Publicación]. España - 2007.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.9238>
- [CGM05] G. de la Calle, M. García, V. Maojo. Structuring European Biomedical Informatics to Support Individualised Healthcare. UPM. IST-507585. [Publicación] 2005.
<http://www.infobiomed.org>
- [CGM05] G. de la Calle, M. García, V. Maojo. Structuring European Biomedical Informatics to Support Individualised Healthcare. UPM. IST-507585. [Publicación] 2005.
<http://www.infobiomed.org>
- [EKS04] Vadim Ermolayev, Natalya Keberle, Vladimir Shapar, Vladimir Vladimirov. Semantically Reinforced Web Services for Wrapping Autonomous Information Resources. [Publicación]. 2004. http://ermolayev.com/eva_personal/PS/VKhNU-EKSV-04-draft.pdf
- [WIG09] Wikipedia. Sistema de gestión de base de dato. [Publicación] 2009.
<http://es.wikipedia.org/wiki/DBMS>
- [GAC04] Geográfico Agustín Codazzi. Sistema gestión de bases de datos. IGAC. [Publicación] 2004.
http://www.igac.gov.co:8080/igac_web/UserFiles/File/ciaf/TutorialSIG_2005_26_02/paginas/ctr_sistema_de_gestion_de_basededatos.htm
- [RCZ06] Ramírez Céspedes Zulia. Las Ontologías como herramienta en la Gestión del Conocimiento. [Publicación]. Cuba - 2006.
<http://www.intempres.pco.cu/Intempres2006/Intempres2006/Ponencias/208.pdf>
- [ASI02] Asisod. Integración de esquemas en bases de dato. [Publicación] 2002.
<http://www.monografias.com/trabajos37/arquitectura-de-sistemas/arquitectura-de-sistemas.shtml>
- [RSG04] Reyes Saldaña Jose Fernando, Garcia Dlores Rodolfo. El proceso de descubrimiento de conocimiento en bases de datos federadas. [Postgrado] 2004.
http://ingenierias.uanl.mx/26/pdfs/26_el_proceso.pdf
- [RMM99] Romero Martinez M. Lenguaje de consultas para una Multibase de Datos. [Publicación] Américas – Puebla. 1999.
http://catarina.udlap.mx/u_dl_a/tales/documentos/msp/romero_m_m/capitulo5.pdf
- [RMI99] Regina Motz InCo. Arquitectura para BDMS distribuidos y heterogéneos. [Publicación] 1999.
http://www.fing.edu.uy/inco/grupos/csi/esp/Cursos/cursos_posg/Inter2002/transparencias/arquitecturas.pdf
- [EAH07] Elimaira Añez .H. Bases de datos Heterogéneas. [Publicación] C.I: 18.411.943. 2007.
<http://alfa.facyt.uc.edu.ve/computacion/pensum/cs0347/download/exposiciones/2/BDH.pdf>
- [CC000] Castilla Carmen. Arquitectura de los Sistemas Distribuidos de Bases de Datos. [Publicación] 2006.
http://antares.itmorelia.edu.mx/~antolino/Curso_Tit2006/Curso_Tit_BDD_ArqIII.ppt
- [NSC05] Navarro Sandra, Castellano Carlos. Bases de Datos heterogéneas. [Publicación]. 2005. <http://www.todoprogramas.com/manuales/ficheros/2008/9.6275.8924.pdf>
- [EIE06] Erica Isabel Enriquez. Evaluación de Rendimiento de Motores de Bases de Datos Multiplataform. [Licenciatura] L.U.: 29063. Argentina – 2006.
<http://exa.unne.edu.ar/depar/areas/informatica/SistemasOperativos/monerica.pdf>
- [ACG99] Alvarez Carrion G. Propuesta de integración de esquemas. [Maestría] 1999.
http://catarina.udlap.mx/u_dl_a/tales//documentos/msp/alvarez_c_g/capitulo4.pdf
- [ACG99] Alvarez Carrion G. Introducción a las bases de datos distribuidas y sistemas multi base de datos. [Maestría] 1999.
http://catarina.udlap.mx/u_dl_a/tales/documentos/msp/alvarez_c_g/capitulo1.pdf

- [VB002] Vineet Buch. Database Architecture: Federated vs. Clustered. [Publicación] 2002.
<http://www.oracle.com/technology/tech/windows/rdbms/ClusterComp.pdf>
- [TSG03] Tomas Boguele, Sebastian Hubner, Gerhard Schister. Buster- An information Broker for the semantic web. [Publicación] ISSN 0933. 2003.
http://www.kuenstliche-intelligenz.de/fileadmin/template/main/archiv/2003_3/buster.pdf
- [FSH07] Federico Sequeda Sanclemente Juan y Hurtado Bustamante Diana Paola. Propuesta del Uso de Ontologías para la Búsqueda Semántica en Laboratorios de Investigación. [Publicación] Colombia - 2007.
<http://libertad.univalle.edu.co/~jsequeda/ontology/Poster-Propuesta%20del%20Uso%20de%20Ontologias%20para%20la%20Busqueda%20Semantica%20en%20Laboratorios%20de%20Investigacion.pdf>
- [SDT06] Sánchez de la Torre. Sistemas Avanzados de Recuperación de la Información (SARI). [Licenciatura]. Madrid - 2006
http://es.geocities.com/lenguajes_ontologicos_en_la_web/Licenciatura.
- [GGG03] García García Grecia. RDF y RDF schema. [Publicación] 2003.
http://www.matem.unam.mx/~grecia/semantic_web/rdf.html.
- [SPA03] Saavedra Places M. Ángeles. Arquitectura para Federación de Bases de Datos Documentales basada en Ontologías. [Doctorado]. A curuña - 2003.
http://rosalia.dc.fi.udc.es/lbd/.%5CPlainConsulta%5CArchivosAdministracion%5CContrabajos_investigadores%5Ctrabajos%5C576965170V.pdf
- [GS08] Galiano Sebastian. Datos en la Web. LCC UNR. [Publicación] 2008.
<http://www.slideshare.net/sebastiangaliano/Datos-en-la-Web-Clase-3>
- [BVG09] Borox Vergara Rosell Alberto García Matías Marcos. Modelos avanzados de bases de datos bases de datos federadas y bases de datos móviles. universidad de castilla la mancha. España – 2009. <http://alarcos.inf-cr.uclm.es/doc/bbddavanzadas/08-09/Distribuidas2.pdf>
- [NSC08] Navarro Sandra, Castellano Carlos. Bases de Datos Heterogéneas. [Publicación] 2008. <http://www.todoprogramas.com/manuales/ficheros/2008/9.6275.8924.pdf>
- [FRA00] Francisca Campins Verger, Pedro Fiol Noguera. Creación de un Sistema de bases de datos federadas a través de Tecnología Java. [Publicación] España - 2000 ,
<http://www.turismo.uma.es/turitec/turitec99/pdf/bd3.pdf>
- [AC999] Alvares carrion. integración de esquemas en bases de datos heterogéneas fuertemente acopladas. [Publicación] 1999.
http://catarina.udlap.mx/u_dl_a/tales//documentos/msp/alvarez_c_g/capitulo4.pdf
- [MVP01] Morocho Zurita Villie, Saltor Fèlix, Pérez Vidal Lluís. Hacia la Integración de Bases de Datos Espaciales. [Publicación] España – 2001,
http://www.lsi.upc.edu/~vmorocho/reportes/hetero_GISpbl.pdf

GIS for Municipalities: A Review of the State-of-the-Art in Germany

Dietrich Schröder

Department of Geomatics, Computer Science and Mathematics
University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
dietrich.schroeder@hft-stuttgart.de

KEYWORDS:

Municipality GIS, GIS architectures, base data

ABSTRACT

The last few years have brought recognition of the meaning of geo information for almost all areas of society. It is estimated that approximately 80% of all decisions in public and private life have a spatial reference. One of the main users of geo data of the national mapping agencies are municipalities. In the presentation different architectures are discussed how geo data in a municipality is distributed to the different departments and to the public. In addition, the base data used in municipalities and a number of applications based on these data are discussed.

A Prototype Digital Cadastral Information System for the Survey of Kenya

David Kuria²⁰,*, Nancy Mwangi⁺ and Moses Murimi*

*Jomo Kenyatta University of Agriculture and Technology

⁺Kenya Institute of Surveying and Mapping

KEYWORDS:

Cadastral Information System, cadastre, survey documents

ABSTRACT:

Survey of Kenya (SoK) is the mapping agency for the country. The director of surveys under provisions of Survey Act (Chapter 299) and Registered Lands Act (Chapter 300) has the responsibility of providing safe custody of, maintaining, updating, retrieving, dissemination and managing cadastral information in the country. The Cadastral Branch is responsible for processing survey records for purposes of land registration and ensuring quality control. These survey records form fundamental component of cadastre. Cadastral survey started way back in 1903. Over the years, volumes of records have increased. Demand for land due to population growth has greatly increased and need for land information for other purposes has grown.

The current system in analogue form has resulted in the system's inability to face the existing challenges and cope with these needs. This research analyses the existing cadastral system in SoK, analyzing the problems from the perspectives of the staff and users. A sample conceptual model for a new digital system is proposed in which logical relations are defined and its implementation as a prototype GIS is discussed.

The proposed system is capable of storing parcel information, e.g. owner, area, locality, date of authentication etc., allows for fast retrieval of these information, including retrieval of survey plans as scanned images, retrieval of boundary definition, and tracking of survey jobs.

The system addresses the problems of storage and tracking the progress of survey jobs within the Cadastral Branch as well.

1 Introduction

Presently cadastral system in Kenya is still in analogue form apart for very small percentage in various departments of Ministry of Lands which have been digitized. The analogue records started being generated way back in 1903 and continue to be generated to date. No system design has been done so far and specifications of data formats, software and hardware are yet to be done (Muchoki, 2008). Parcel information is in analogue form with paper maps and conventional land registers giving information on location, tenure, use, encumbrances, ownership and distribution of land.

The classical cadastral system has become insufficient owing to the fact that land records have greatly increased in volumes, they are in paper form thus the process of storage, access and retrieval is a great challenge (Mulaku, 2000). In the course of economic development the demand for more varied information has become increasingly pressing, and data stored in this conventional cadastral system fail to meet requirements related with supervision, management, decision-making, forecasting and development planning.

Due to economic growth there is wider use of spatial data. Policy makers, planners, environmental scientists, administrators, developers and individual citizens are in need of accurate land information. Accessibility of land information is currently a difficult task due to the form in which it is in even with the Internet services which allow easy access and sharing of data.

Increased population and rural urban migration has led to higher demand for land both in rural and urban areas. The pressure on efficiency in land management is more critical in the cities and urban areas due to urban sprawl which has had great impacts on land use, health, environment, and cost of living among others.

The need for digital cadastre cannot be underestimated. Land being the greatest resource, and with much of the economy based on land as a resource, efficiency in land administration, planning and management is of

²⁰

Corresponding author: dn.kuria@gmail.com

paramount importance (Kaufmann and Steudler, 1998). The issue of land ownership, use and rights is a very sensitive issue in Kenya. With the growing interest in land as a valuable resource and population growth, there has been increasing demand for this limited resource. There is therefore a dire need for efficient land management to provide security of tenure, enhance social economic growth, ensure political stability, and for sustainable development.

The objectives of this study were therefore, (i) identification of the shortcomings/problems with the existing cadastral information system, (ii) identification of the obstacles/constraints to development of a Digital Cadastral Information System (DCIS) for the Survey of Kenya and (iii) development of a customized Geographic Information System (GIS) based system that offers solution to some of the identified problems.

2 Methodology

2.1 The Scenario

The usual procedure in processing a surveying project at the Survey office proceeds as follows:

- a surveyor submits the completed job (survey notes, computations, plans) to the survey office. They are noted and given a tracking number.
- The job then undergoes various checks for confirming that it was professionally done.
- If the computations and job execution are satisfactory, then the submitted plan will be used to amend existing maps to incorporate the changes in dimensions and acreage. At this point the job is said to be authenticated
- it is subsequently submitted to the registrar of lands for issuance of titles depending on the registration legislation in force for the particular type of job.
- Once the title is ready, then the job is said to be fully completed and duly registered.

While these processes may seem straightforward and obvious, the time taken to complete this process can take up to five years and can be frustrating for the surveyor (and their clients) as they seek to track the progress of their work, and for members of the survey office as they track the many jobs that are submitted.

2.2 Analysis of the Problem

To aid in understanding the problem, three types of questionnaire were prepared in order to collect data which could be analyzed.

The first questionnaire collected information from officers interacting with the existing non computerized system on a daily basis and provided information on existing problems, causes of these problems and what they thought would be appropriate solutions.

The second questionnaire targeted the top management in survey of Kenya most of whom have worked in the cadastral division. These provided more information on the efforts and the obstacles towards the establishment of digital cadastral system.

The third questionnaire targeted the users. In this case, licensed and other private surveyors were the only users given the questionnaires. This was because they form the greatest percentage of users. This provided information on what problems are experienced by users, their user needs, what they perceive as the causes and what solutions they would suggest.

Random sampling was done in the distribution of the questionnaires, statistical analysis using the Statistical Package for the Social Sciences (SPSS) software.

2.3 System Design

From the questionnaires, the greatest problem for members of staff was the lack of a computerized system for storing and retrieving land records. These questionnaires also informed the decision on the various processes that take place at the Survey of Kenya, in the various sections, and the various stages that a submitted survey job passes through until the final approval. The existing digitized data comprises of a collection of shapefiles without a proper automated solution for their utilization. To enhance information flow (for both spatial and non-

spatial information) a centralized database was used. This leverages the strengths of database systems within the Survey office. Since this is a large project, it was important to develop a good conceptual model of the entire system before implementing it.

The development of the spatial database consisted of preparation of the conceptual model of the cadastral system, creation of entity relationships, creation of tables, and implementation of a prototype. The conceptual model of the proposed system comprised of identifying the key entities in the system, identification of attributes,, relationships and other information requirements. The logical design consisted of the mapping of the conceptual model to the relational schema to come up with normalized relations (tables). The relations were created in ArcCatalog (see figure 2)

Figure 1 shows the key entities and the relationships amongst the entities. The parcel is the basic entity. The relations of the entities with the parcel were identified and they include 1:1 as well as 1: many relationship. These entities were then mapped to correspond to database tables. The tables identified that needed to be created were nine. These included:-

- Parcel: - The parcel table comprise of information of the parcel such as parcel number, area, value, encumbrances, owner id etc. It also comprises of an array of point ids of all points that define the parcel boundary.
- Approval: - This table provides information on the approvals given from government ministries and departments for any survey job to be done.
- Locality: - The table provides administrative information of a parcel e.g. registration district, province etc.
- Deed plan: - Table provides information on the deed plan number, the parcel id, date of issue etc.
- Registry Index Maps (RIM): - A registry map is created by section 18 (1) of the Registered Land Act of the Laws of Kenya, and its production is vested in the Director of Surveys. The Act states that “The Director of Surveys shall prepare and thereafter maintain a map or series of maps to be called the registry map for every registration district.” This registry map is what is referred to as the RIM.

The RIM table indicates the details of an RIM that constitute given parcels.

- Surveyor: - This table provides information of the surveyor who carried out the survey that created given survey records of a parcel e.g. surveyors id, First, owner’s names, physical address etc.
- Owner:- This table provides information of the parcel owner
- Comps and Folio Register (FR): - This table provides information of computation file number, field book number and folio registration number of survey plan. Further gives information on date survey job was received and date authenticated. Tracking of the progress of survey job is done by indicating the status. (i.e., at Survey Records Office (SRO) or at preliminary checking, etc.) and the date and time it was received at the section.

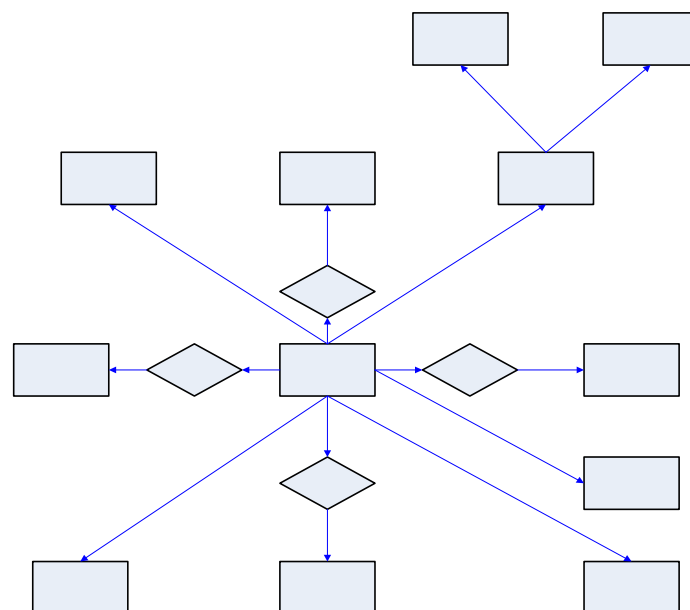


Figure 1: Entities comprising the system and their entity – relationships

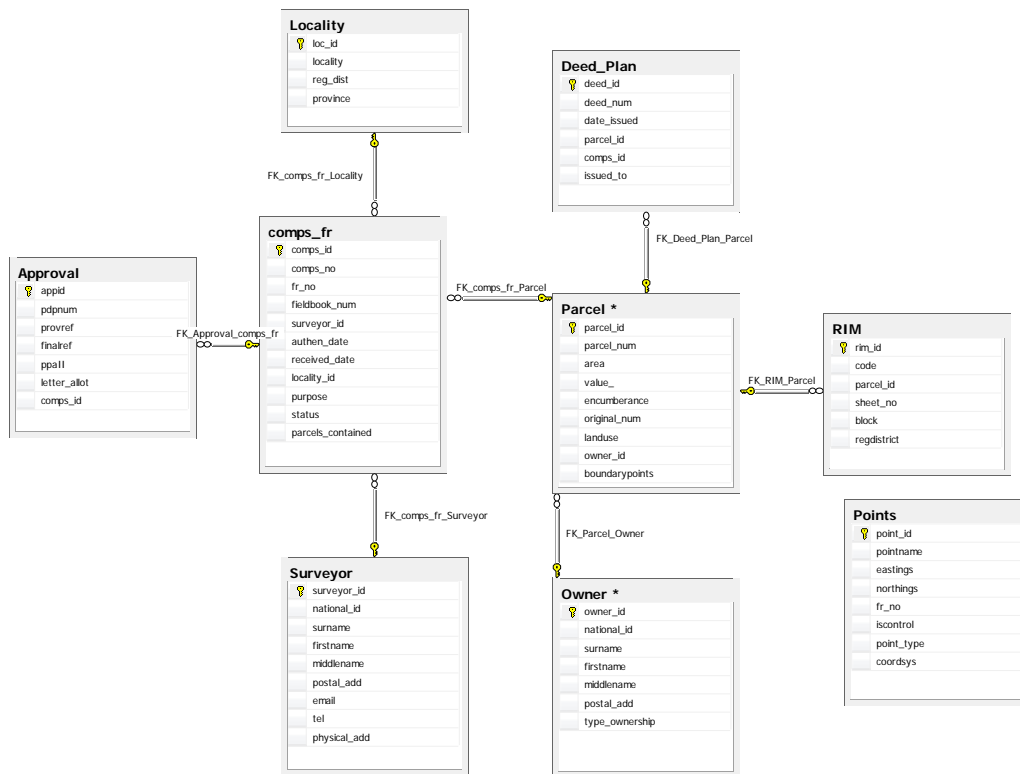


Figure 2: Tables and relationships developed using ArcCatalog[™]

- Points (Corner and Control Points):- This table provides information of points ids, point names, Eastings, Northings, point type etc.

For each table, the fields were determined and primary keys indicated. The choice of the fields was dependent on the information that the database should be able to provide both to SoK staff and users. Redundancy in the original tables was addressed through normalization thereby enhancing integrity of database (Codd, 1970). Figure 2 shows these tables with the various relationships as done using ArcCatalog[™]

An interface for searching, and retrieval of information from the database was developed as an ArcGIS extension. This extension allows interrogation of the database on parcel information and tracking the progress in processing cadastral survey jobs. The search field and search value are entered to enable the system do a search. When a record is found, all information regarding the parcel and its relation to the other entities can be displayed.

3 Results

3.1 Analysis of Problems and Solutions

During the analysis of the existing system, various problems, causes and solutions were identified by both users and officers as shown in figure 3. For the staff most opined that inadequate storage facilities were the greatest hindrance while most users were unhappy with lengthy processes in processing of cadastral jobs. Other identified problems show the need to put in place an efficient solution (missing and worn out records, difficulties in retrievals etc)

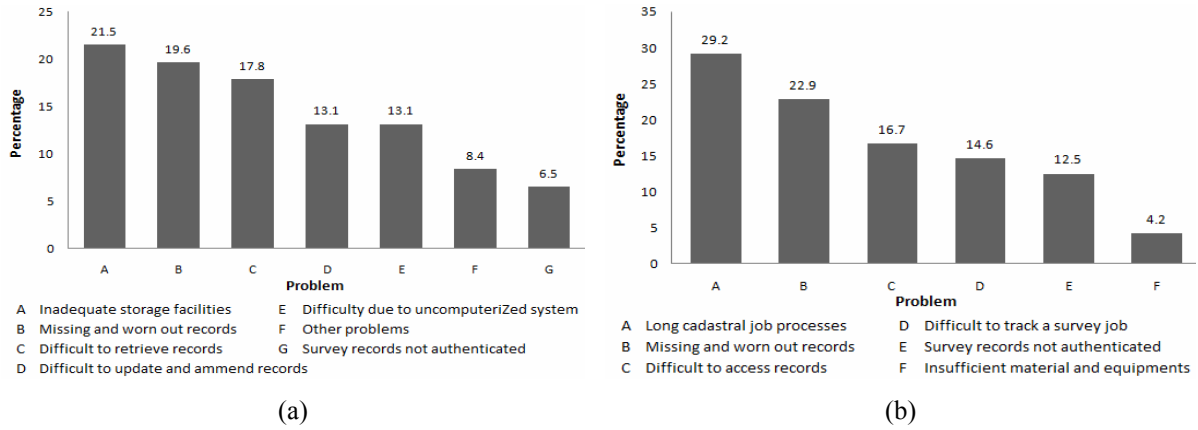


Figure 3: Identification of problems from questionnaire survey a) from SoK staff and b) from users

From the survey of possible causes of the problems identified in figure 3, it was clear that lack of computerization was the most severe cause for both members of staff and users. This information is presented in figure 4 from which more than a third of the respondents cited lack of computerization as the cause for these problems.

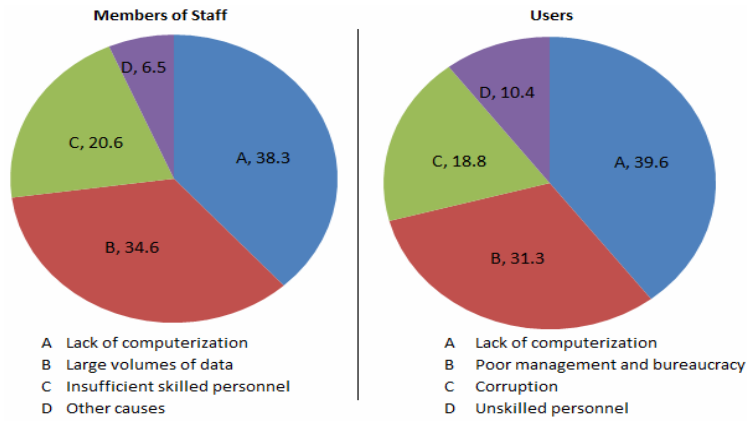
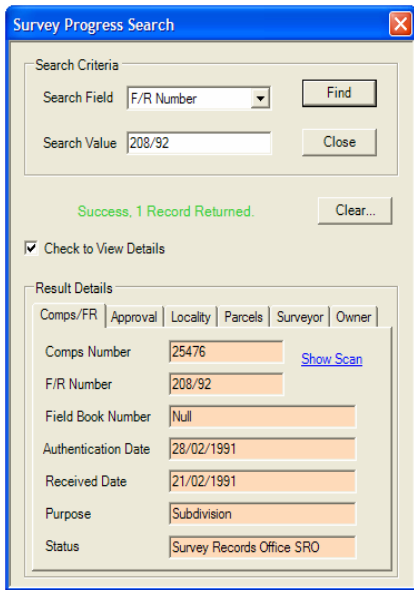


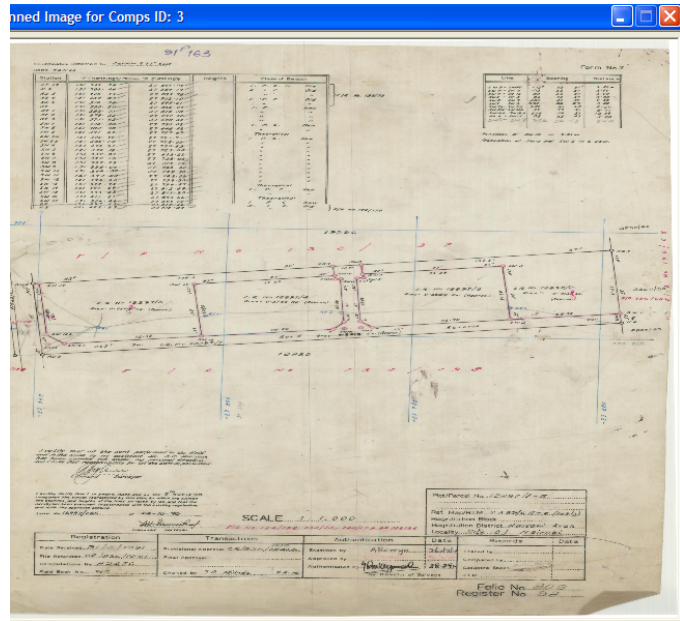
Figure 4: Cited causes to the identified problems

3.2 Spatial Database Query Results

Figure 5 (a) shows the query interface which features a dialog type application with five tabs representing the tables in the database which have information that is connected to a particular parcel. Figure 5 (b) shows the retrieval of a scanned plan



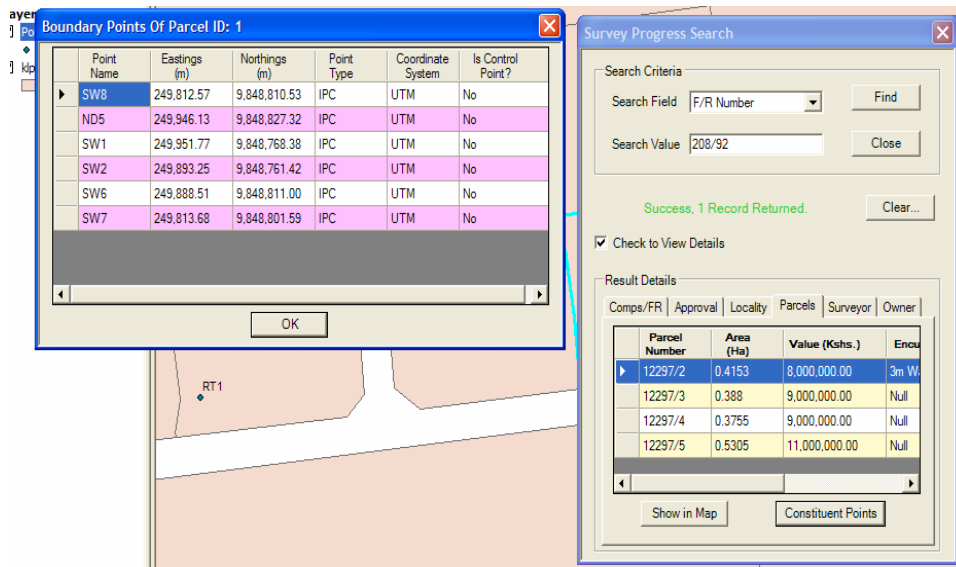
(a)



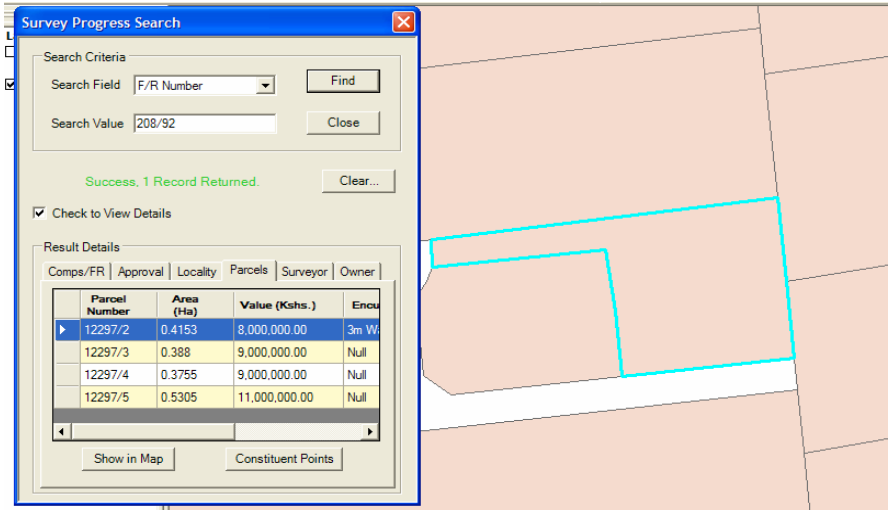
(b)

Figure 5 Query interface (a) and results – scanned plan (b)

Figure 6 shows the retrieval of points comprising the parcel boundaries and a zoomed in selection of the affected parcel in the map window.



(a)



(b)

Figure 6: Retrieval of boundary points information (a) and selection of the parcel in the map window (b)

4 Conclusion

The existing problems have been identified using the questionnaires administered to both officers and users. It is clear that computerization is the most appropriate solution to the identified existing problems. To aid in this computerization, entities of a data model of the system have been successfully identified and relationships defined in a way that ensures database integrity. The prototype system designed enhances storage, retrieval and access of survey records/data/information, and also provides quicker means of tracking a survey job within the office. This system allows fast retrieval of both raster (scanned plans) and vector data (digitized parcel boundaries) and associated attributes and be used to perform cadastral jobs tracking, owner inventories amongst other capabilities.

This prototype can be further developed to fully automate the work-flows within the cadastral division. With full automation, the system can be extended to allow access through Internet thus enhancing efficiency in service delivery.

REFERENCES

- Codd, E. F. (1970). "A relational Model of Data for Large Shared Data Banks." *Communications of the ACM*, 13(6), pp 377-387
- Kaufmann, J. and Steudler, D. (1998). "Cadastre 2014, A Vision for A Future cadastral System." Working Group 7.1 FIG Commission 7.
- Mulaku G. C. (2000). "Setting up a land information system: a systematic approach." *Suid-Afrikaanse Tydskrif vir Opemeting en Geo-Inligting*, 1(5 & 6), pp 260-264
- Muchoki C.M, 2008. "Towards an automated business at SOK- the challenges." Institution of surveyors of Kenya.
- The Kenya Law Review, Laws of Kenya, online http://www.kenyalaw.org/kenyalaw/klr_home/ accessed on 22nd July 2010

Workshops

Geostatistics – an Introduction

Hardy Lehmkuhler

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
hardy.lehmkuhler@hft-stuttgart.de

KEYWORDS:

Geostatistics, Geodata, Kriging

ABSTRACT:

The tutorial "Geostatistics -- An Introduction" is meant to motivate further dealing with the issue. After explaining the specialty of spatial statistics a first completely regular data set is considered. Here ESRI's ArcGIS GeospatialAnalyst is used and demonstrated as one of several tools of analysis. While descriptive statistics foundations are simple and clear an important and somehow strange issue will be the spatial self-similarity of measured quantities or expressively correlation. A second data set consisting of several irregularly and differently distributed quantities points out some important questions of estimation. Such questions deal with representative mean values or distributions. The estimation of point values (interpolation) or larger block values follows. The kriging method for interpolation is addressed. A final view is laid upon the theory of cokriging where unmeasured quantities are derived from measured ones. The tutorial is also preparatory for the workshop "R for geospatial analysis".

Workshop on Open Source Desktop GIS

Dietrich Schröder

Department of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
dietrich.schroeder@hft-stuttgart.de

1 Desktop GIS and Free and Open Source Software

A **geographic information system (GIS)** is an information system for capturing, storing, analyzing, managing and presenting data which is spatially referenced, i.e. linked to location.

GIS applications are tools that allow users to create interactive queries (user created searches), analyze spatial information, edit data, maps, and present the results of all these operations. Geographic information system technology can be used for scientific investigations, resource management, asset management, archaeology, environmental impact assessment, urban planning, cartography, criminology, marketing, and logistics to name a few.

A **desktop GIS** is a mapping software system that is installed onto and runs on a personal computer and allows users to display, query, update, and analyze data about geographic locations and the information linked to those locations.²¹

That is, the GIS software is not executed on a server and remotely accessed or controlled from or by a different computer. Another definition focus on the functionality:

A desktop GIS is a GIS program with an interactive GUI and reduced GIS functionality, mainly used for visualizing of GIS data or developed only for special applications.²²

Of course the borders of the different types of GIS is blurred, there are data viewers which offer additional functionality, where as nearly all desktop GIS nowadays offer the possibility of consuming OGC services like WMS or WFS, so they are using resources from the Internet. On the other hand, workhorses like GRASS are still running on one single personal computer but offering full GIS functionality.

Over the last few years the world of free and open source geospatial software has experienced some major changes. For instance, the website www.freegis.org of the German association FOSSGIS e.V. currently lists more than 350 GIS related projects. With the broad use of non-proprietary and open data formats such as the Shape File format for vector data and the Geotiff format for raster data, as well as the adoption of OGC standards for networked servers, development of open source software continues to evolve, not only for web and web service oriented applications, but also desktop GIS have some kind of renaissance. Well-known open source desktop GIS software includes GRASS GIS, Quantum GIS, uDig, OpenJUMP, gvSIG and many others. Despite the diversity of projects an exchange among projects can be clearly noticed. This holds true especially for the use of software libraries that enable data format conversions (e.g. GDAL/OGR, GeoTools), provide coordinate projections (Proj4), or offer basic geometric algorithms (e.g. GEOS, JTS). These libraries are used by open source and commercial software alike to provide basic functionality.

There are several dozen Open Source based Desktop GIS applications available to choose from, most of which are optimized for specific application areas. The relevant developments are mostly based on Java or C++ and provide a high degree of flexibility due to standardized interfaces and the capability to be deployed within service oriented architectures.

Nevertheless, desktop GIS are powerful tools and but still require a solid background and good training to unleash its full potential. After clarifying the meaning of free and open source software development, the most important desktop GIS with an active developer and user community will be introduced.

Open source²³ is a development methodology, which offers practical accessibility to a product's source (goods and knowledge). Some consider open source as one of various possible design approaches, while others consider it a critical strategic element of their operations. Before open source became widely adopted, developers and

²¹ ESRI *GIS Dictionary* [online]. [http://support.esri.com/index.cfm?fa=knowledgebase.gisDictionary.search&searchTerm= desktop%20GIS](http://support.esri.com/index.cfm?fa=knowledgebase.gisDictionary.search&searchTerm=desktop%20GIS)

²² Buhmann/Wiesel [2007]. GIS Report. Bernhard Harzer Verlag, Karlsruhe, Germany (in German)

²³ see http://en.wikipedia.org/wiki/Open_source and http://en.wikipedia.org/wiki/GNU_General_Public_License

producers used a variety of phrases to describe the concept; the term open source gained popularity with the rise of the Internet, which provided access to diverse production models, communication paths, and interactive communities. The following Open Source Definition is used by the Open Source Initiative (OSI) to determine whether or not a software license can be considered open source²⁴:

1. Free Redistribution

The license shall not restrict any party from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources. The license shall not require a royalty or other fee for such sale.

2. Source Code

The program must include source code, and must allow distribution in source code as well as compiled form. Where some form of a product is not distributed with source code, there must be a well-publicized means of obtaining the source code for no more than a reasonable reproduction cost preferably, downloading via the Internet without charge. The source code must be the preferred form in which a programmer would modify the program. Deliberately obfuscated source code is not allowed. Intermediate forms such as the output of a preprocessor or translator are not allowed.

3. Derived Works

The license must allow modifications and derived works, and must allow them to be distributed under the same terms as the license of the original software.

4. Integrity of The Author's Source Code

The license may restrict source-code from being distributed in modified form only if the license allows the distribution of "patch files" with the source code for the purpose of modifying the program at build time. The license must explicitly permit distribution of software built from modified source code. The license may require derived works to carry a different name or version number from the original software.

5. No Discrimination Against Persons or Groups

The license must not discriminate against any person or group of persons.

6. No Discrimination Against Fields of Endeavor

The license must not restrict anyone from making use of the program in a specific field of endeavor. For example, it may not restrict the program from being used in a business, or from being used for genetic research.

7. Distribution of License

The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties.

8. License Must Not Be Specific to a Product

The rights attached to the program must not depend on the program's being part of a particular software distribution. If the program is extracted from that distribution and used or distributed within the terms of the program's license, all parties to whom the program is redistributed should have the same rights as those that are granted in conjunction with the original software distribution.

9. License Must Not Restrict Other Software

The license must not place restrictions on other software that is distributed along with the licensed software. For example, the license must not insist that all other programs distributed on the same medium must be open-source software.

10. License Must Be Technology-Neutral

No provision of the license may be predicated on any individual technology or style of interface.

A slightly different understanding has the Free Software Movement. The fundamental difference between the two movements is in their values, their ways of looking at the world. For the Open Source movement, the issue of whether software should be open source is a practical question, not an ethical one. As one person put it, "Open source is a development methodology; free software is a social movement." For the Open Source movement, non-free software is a suboptimal solution. For the Free Software movement, non-free software is a social

²⁴ see <http://opensource.org/docs/osd>

problem and free software is the solution. According to the Free Software Foundation²⁵, free software is a matter of the users' freedom to run, copy, distribute, study, change and improve the software. More precisely, it means that the program's users have the four essential freedoms:

- The freedom to run the program, for any purpose (freedom 0).
- The freedom to study how the program works, and change it to make it do what you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help your neighbor (freedom 2).
- The freedom to distribute copies of your modified versions to others (freedom 3). By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

From a practical point of view of using the software, there is no big difference. In particular in the GIS community, the term FOSSGIS is used for Free and Open Source Software GIS, to show the close relation of the two terms.

The use and distribution of software is based on different licensing models. An **open source license** is a copyright license for computer software that makes the source code available under terms that allow for modification and redistribution without having to pay the original author. Such licenses may have additional restrictions such as a requirement to preserve the name of the authors and the copyright statement within the code. Most OSGIS are using either GPL or LGPL as licensing policy.

The **GNU General Public License** (GNU GPL or simply GPL) is a widely used free software license. The GPL is the most popular and well-known example of the type of strong copyleft license that requires derived works to be available under the same copyleft. Under this philosophy, the GPL is said to grant the recipients of a computer program the rights of the free software definition and uses copyleft to ensure the freedoms are preserved, even when the work is changed or added to. This is in distinction to permissive free software licences, of which the BSD licences are the standard examples. The distribution rights granted by the GPL for modified versions of the work are not unconditional. When someone distributes a GPL'd work plus their own modifications, the requirements for distributing the whole work cannot be any greater than the requirements that are in the GPL. This requirement is known as copyleft.

The **GNU Lesser General Public License** (LGPL) is a modified, more permissive, version of the GPL, originally intended for some software libraries.

There exist quite a number of other license models, a discussion of pros and cons can be found on <http://www.gnu.org/licenses/license-list.html>.

2 Free and Open Source Desktop GIS Projects

2.1 The Jump Family

The Java Unified Mapping Platform (JUMP) is a GUI-based application for viewing and processing spatial data. It includes many functions common to other popular GIS products for the analysis and manipulation of geospatial data. The JUMP also provides a highly extensible framework for the development and execution of custom spatial data processing applications.

In 2002, as a project for the British Columbia Ministry of Sustainable Resource Management, Vivid Solutions Inc. created a software system to do automated matching ("conflation") of roads and rivers from different digital maps. The software team wisely made the program flexible enough to be used not just for roads and rivers, but almost any kind of spatial data: provincial boundaries, power-station locations, satellite images, and so on. The program was named JUMP (JAVA Unified Mapping Platform).

Based on the JUMP-framework several different applications have been developed, like OpenJump, deeJump, and Kosmo. All these application have been developed on Java and offer the possibility to add more functionality using plug-ins.

2.1.1 OpenJump

After the initial creation and deployment of JUMP, regular development of the program by Vivid Solutions stopped. However, the company continued offering support to the user community that had grown around JUMP,

²⁵ <http://www.gnu.org/philosophy/free-sw.html>

and provided information to developers that had begun to improve JUMP in small ways, or who had customized it to fit their needs.

It soon became evident that both the users and developers would benefit from a "unified" JUMP platform. This central or core platform would eliminate the compatibility issues that plagued the JUMP user community, and would give developers a platform on which to focus and coordinate their efforts. A number of the lead members from each team working with JUMP formed the JPP Development Committee, whose purpose was to guide and oversee this new unified platform. A name was chosen for this open source GIS program to be based on JUMP, "OpenJUMP"

So most of the features of JUMP are available in OpenJump, and meanwhile a lot of additional features have been added.

Open Jump

- reads and writes ESRI Shapefile, GML files, DXF and PostGIS
- reads raster files like TIFF, JPG, PNG and ECW
- saves view to georeferenced rasters like JPG and PNG
- offers full geometry and attribute editing
- is OpenGIS SFS (simple feature specification) compliant
- supports WMS and SLD as OGC-standards
- uses Geometry algorithms based on Java Topology Suite (JTS)
- is an easy extensible GIS programming environment for own GIS-applications

On the very well organized OpenJump homepage[1] additional information can be found, including documentation and links to plug-ins. The focus of the plug-ins is on data loading and viewing of data (e.g connecting to PostGIS Oracle database or ArcSDE, GPS data, print, reproject vectors, quality check, charts, map generalization, WFS, etc.). For spatial analysis the Sextante extension (see below) can be used. The language for the graphical user interface has to be set in the start up configuration file; up to now Spanish is only partly supported.

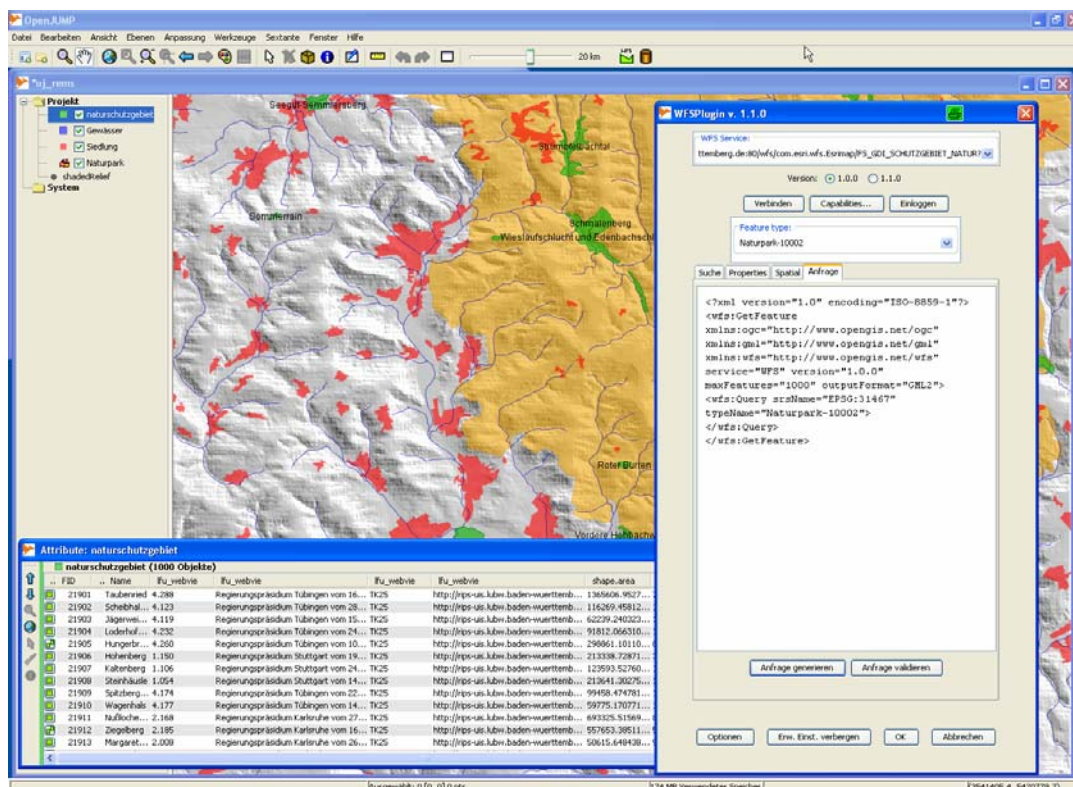


Figure 1: openJump user interface with attribute table and dialog to configure a WFS

2.1.2 deeJump

deeJump [2] is a project of the German company lat/lon. The goal of this project is to add OGC compliant web functionality to Jump. It has some functions of OpenJUMP and has some further functionality with respect to web standards. Main parts of the implementation have been taken from the deegree server project of lat/lon, so the named it deeJump.

2.1.3 Kosmo

The idea of the **Kosmo** [3] project is to develop a whole suite of GIS software including desktop, server, OGC client, and mobile. Kosmo design and architecture is focused on management and analysis of territorial information through Spatial Databases, so providing it with corporative nature. The ambitious project is still under development supported by the Spanish company SAIG with the first of its components -Kosmo Desktop- in continuous evolution and available for whom requires advanced functionality in a powerful desktop GIS. Kosmo-Desktop was implemented using the Java programming language and is developed from the JUMP platform and several other open source libraries, e.g. JTS, Geotools, GDAL, etc.

2.2 gvSIG

gvSIG is a GIS that can handle both vector and raster data. It features basic editing tools for the creation and maintenance of vector or raster spatial data on a variety of file formats, including remote data sources. gvSIG is being actively developed in Java by IVER Tecnologías under the GNU General Public License (GPL). Its name is an abbreviation that stands for *Generalitat Valenciana, Sistema d'Informació Geogràfica*.

gvSIG was started in 2003 when the *Conselleria d'Infraestructures i Transports* (CIT) of Valencia in Spain proposed the development of software for the management of geographic information. The private enterprise IVER Tecnologías won the proposal and is now developing the software together with the Generalitat Valenciana and the Jaume I University of Castellón. So it is a Spanish driven project, which still can be noticed from the GUI, the documentation, and the homepage [4].

It is targeted at professional users of geographic information from the public sector, private industry sciences and education. It is developed as an Open Source project and currently undergoes the OSGeo Incubation process. The software provides internationalization modules and a lot of languages are already supported.

As the development process is mainly driven by the requirements of the customers of IVER Tecnologías and not but the user community, there was a branching of gvSIG into the OADE gvSIG. OA Digital is a private company with offices in UK and France. Though it is planned to merge the two branches of gvSIG in the future, OADE gvSIG [5] offers at this moment the most user-friendly installation including most of the available plug-ins in addition to several bug-fixes. The plug-ins for gvSIG include language (118n) extension, network, database (Oracle and ArcSDE), Topology and SEXTANTE just to name a few.

gvSIG has been developed with INSPIRE principles (Infrastructure for Spatial Information in the European Community) in mind. It Provides the most common GIS tools like data loading, map navigation, query map information like alphanumeric information, distance measurement, thematic cartography, legend edition using the most common legend types, labelling, feature selection by many selection types, data tables with statistics, ordering, table relations, table linking, layout manager, geoprocessing tools, CAD, raster processing, etc.

It is highly interoperable, i.e. it is able to work with most of the known data formats like ecw, ENVI hdr, ERDAS img, (Geo)TIFF, GRASS for raster; shapefile, GML, KML, DGN, DXF, DWG for vector, and databases like PostGIS, MySQL, Oracle, ArcSDE. Its SDI client condition permits the connection to the OGC standards WMS, WFS (and WFS-T), WCS, and WMC.

Its functionality includes Geoprocessing tools like proximity (buffer, spatial join), overlay (clip, difference, intersect, union), advanced raster tools (georeferencing images, set image transparency, adjust bright and contrast, highlight, etc.) and advanced functionalities like Scripting support, a re-projection engine (PROJ4 wrapper), 3D visualization, network analysis, raster analysis features like classification or rectification, and gvSIG for mobile devices.

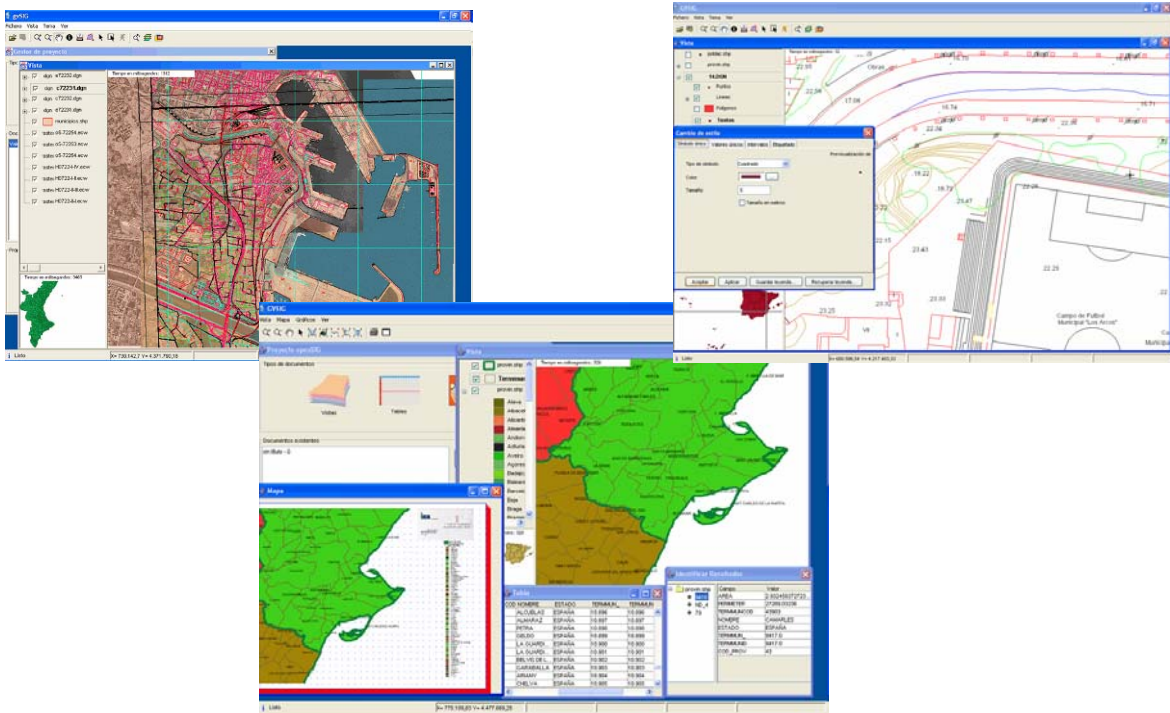


Figure 2: gvSIG user interface

2.3 Quantum GIS

Quantum GIS (QGIS) [6] is a user-friendly Desktop-GIS, released under the GNU General Public License (GPL) and available for Linux, Unix, MacOS X and Windows operating systems in several interface languages. Quantum GIS also acts as frontend user interface to the GRASS program since 2005 and thus benefits from the huge number of features of this package. The project itself started 2002.

Quantum GIS is implemented in C++ and includes Plug-In interfaces to extend the initial functionality. You can also use the software to configure UMN MapServer MAP-files, which reduces the effort needed to set up UMN MapServer map files.

Without the GRASS plug-in (and of course the installation of GRASS itself), QGIS mainly offers functionality as a GIS viewer and as a digitizing/editing tool. As GIS formats, it supports GDAL/OGR. Beside the GRASS extension, plug-ins mainly focus on simple vector layer processing (buffer, convex hull, intersection, difference, dissolve, union), geometry calculations, voronoi polygons, sampling, basic raster algebra. There exist also an extension for loosely coupling QGIS to R. It allows upload of QGIS layers directly into R, and the ability to perform R operations on the data directly from within QGIS.

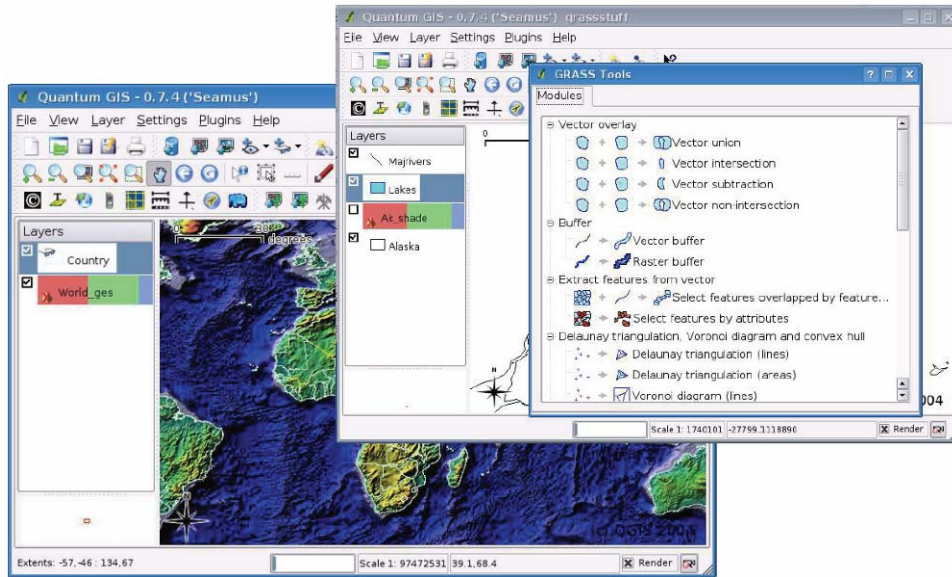


Figure 3: QGIS user interface

2.4 uDig

The User-friendly Desktop Internet GIS (uDig) [7] is a GIS application that serves as a viewer and editor for geospatial data that is served via OGC compliant WMS and OGC WFS. uDig is implemented in Java and based on GeoTools.

uDig is not only used as an 'off the shelf product', but provides an environment to implement solutions for domain specific applications. The software has been developed by Refrations Inc., Canada and is available under the GNU Lesser General Public License (LGPL). Its main focus is on viewing and digitizing/editing GIS data.

uDig also serves as a base framework for other applications like forest management, hydrological modelling, route planning, etc.

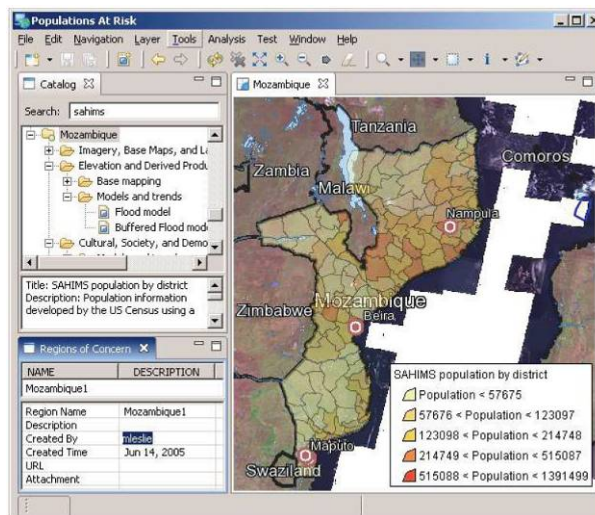


Figure 4: uDig user interface

2.5 SAGA

SAGA - short hand for "System for Automated Geoscientific Analyses" - is a free, hybrid, cross-platform GIS software [8]. The heart of SAGA is its C++ and thus object oriented *Application Programming Interface* (API), providing data object definitions and computational methods for raster, vector and tabular data. As a normal user,

you will not get into touch with the API. But as an interested scientist or coder you will soon discover its great flexibility.

As user, you will most likely focus on using the steady growing availability of geo-scientific methods. These are implemented in various SAGA modules which are bundled in so-called module libraries. These module libraries are accessible to you in different ways: By a *Graphical User Interface* (SAGA GUI) or by one of the scripting methods. The scripting methods include a *Command Line Interface* (SAGA CMD) which allows for batch/bash scripting (i.e. coupling of different modules to automate tasks) and a python interface which gives you also direct access to the SAGA API and is thus more flexible.

There are a lot of spatial analysis tools available, e.g.

- geostatistics (kriging, regression analysis)
- fuzzy logic
- grid analysis
- point pattern
- interpolation
- terrain analysis
- vector analysis

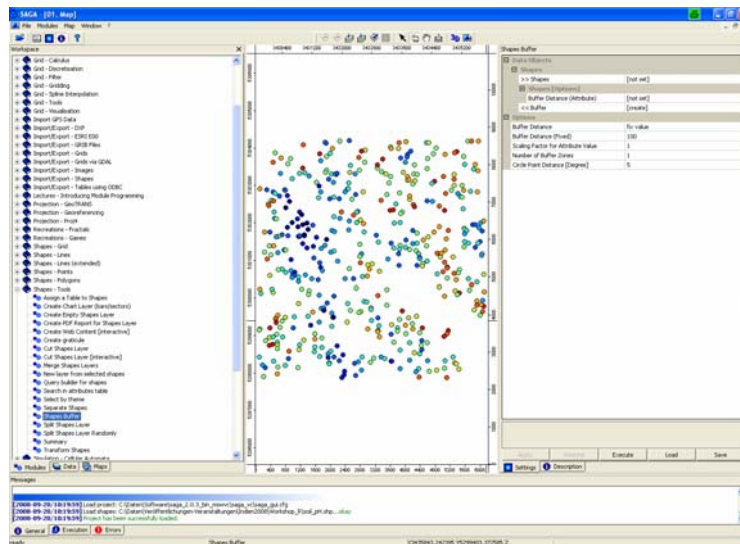


Figure 5: SAGA user interface

2.6 GRASS

The Geographic Resources Analysis Support System (GRASS) [9] is the most complex and professional Open Source based desktop GIS and is supplied under the GNU General Public License (GPL). Supported platforms include Linux, Unix, MacOS X and Windows. The strength of this GIS lies in its modular structure and its utility for raster processing and modelling.

Users can interface with the software features through a graphical user interface (GUI) by or by "plugging into" GRASS via Quantum GIS. They can also interface with the modules directly through the modified version of the shell that the application launches or by calling individual modules directly from a preferred shell (this latter method requires the setting of several environment variables).

The recent GRASS 6 release introduces a new topological 2D/3D vector engine and support for vector network analysis. Attributes are managed in .dbf files or SQL-based DBMS such as MySQL, PostgreSQL/PostGIS, and SQLite. The system is capable of visualizing 3D vector graphics data and voxel volumes. GRASS supports an extensive range of raster and vector formats through the binding to GDAL/OGR libraries, including OGC-conformal (Open Geospatial Consortium) Simple Features for interoperability with other GIS. It also supports Linear Reference System.

The GRASS Development Team is a multi-national group consisting of developers at numerous locations. GRASS is one of the eight initial Software Projects of the Open Source Geospatial Foundation.

GRASS is designed as an environment in which tools that perform specific GIS computations are executed. Unlike typical application software, upon starting GRASS, the user is presented with a UNIX shell containing a

modified environment that supports the execution of GRASS commands (known as modules). The environment has a state that includes such parameters as the geographic region covered and the map projection in use. All GRASS modules read this state and additionally are given specific parameters (such as input and output maps, or values to use in a computation) when executed. The majority of GRASS modules and capabilities can be operated via a graphical user interface (provided by a GRASS module), as an alternative to manipulating geographic data in shell. There are over 200 core GRASS modules included in the GRASS distribution, and over 100 add-on modules created by users and offered on the GRASS web site. The GRASS libraries and core modules are written in C; other modules are written in C, UNIX shell, Tcl, or other scripting languages. The GRASS modules are designed under the Unix philosophy and hence can be combined using shell scripting to create more complex or specialized modules by a user without knowledge of C programming.

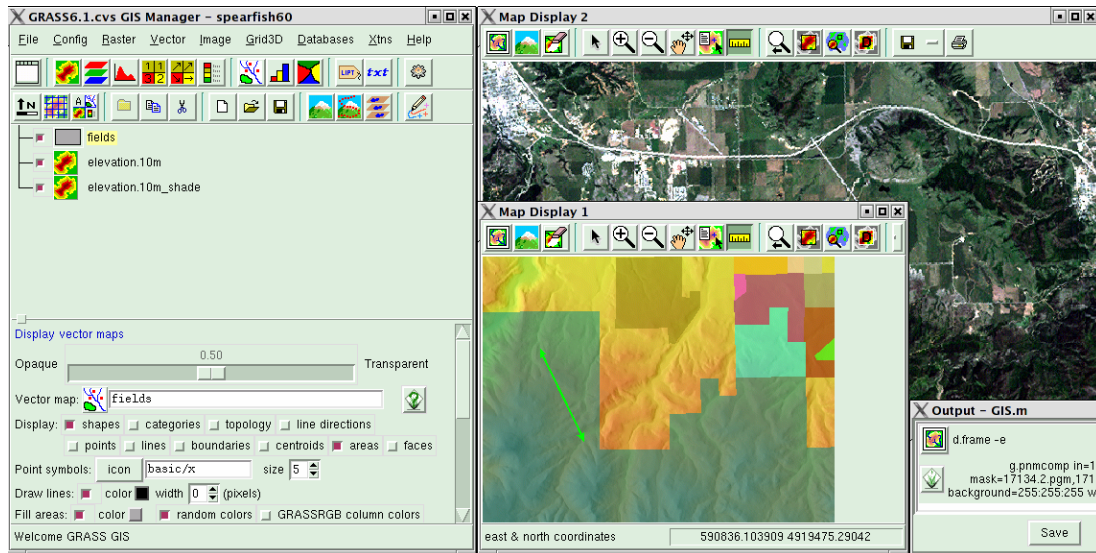


Figure 6: GRASS user interface

In the next Sextante-Release, the GRASS-tools will be made available as well. That means, that some additional 300 tools can be used in the future by gvSIG-Sextante or openJump-Sextante as well.

2.7 ILWIS Open

The Integrated Land and Water Information System (ILWIS) is a PC-based GIS & Remote Sensing software, developed by ITC (International Institute for Geo-Information Science and Earth Observation, The Netherlands) from 1988 up to its last release (version 3.3) in 2005. ILWIS comprises a complete package of image processing, spatial analysis and digital mapping. It is easy to learn and use; it has full on-line help, extensive tutorials for direct use in courses and 25 case studies of various disciplines. It has been originally programmed in C, the open source version has switched to an MS Visual Studio (.NET) project.

As per July 1st, 2007, ILWIS software is freely available ('as-is' and free of charge) as open source software (binaries and source code) under the 52°North initiative (GPL license). This software version is called **ILWIS Open** [10].

Some of the key features are:

- Integrated raster and vector design
- Import and export of widely used data formats
- On-screen and tablet digitizing
- Comprehensive set of image processing tools
- Orthophoto, image georeferencing, transformation and mosaicing
- Advanced modeling and spatial data analysis
- 3D visualization with interactive editing for optimal view findings
- Rich projection and coordinate system library
- Geo-statistical analyses, with Kriging for improved interpolation
- Production and visualization of stereo image pairs
- Spatial Multiple Criteria Evaluation

- WMS (new in V3.5)
- Surface Energy Balances (new in V3.5)

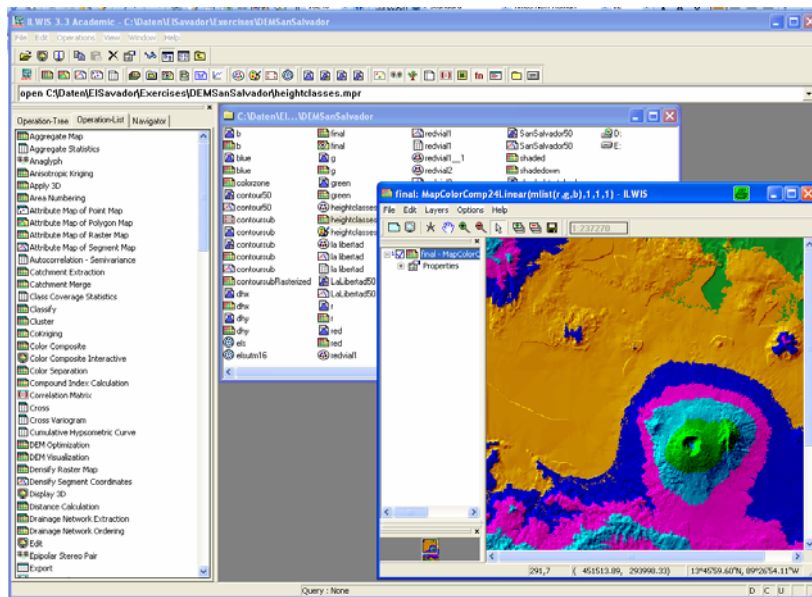


Figure 7: ILWIS user interface

2.8 Sextante

Sextante [11] is not a desktop GIS, but a toolbox for GIS analyses, which can be added to several FOSSGIS. It is an analysis suite of more than 220 tools that has been developed by the Regional Government of Extremadura.

Meanwhile Sextante is supported by several FOSSGIS, e.g. openJump and uDig, but the best integration will be found with gvSIG. Sextante has provided gvSIG with both raster and vector geographical analysis capabilities, geo-statistics, vegetation indexes, profiles and hydrological analysis, fuzzy logic, point pattern analysis to name but a few of the implemented functions. All the extensions are based on a set of foundational classes, specially conceived to ease the implementation of algorithms for geographical analysis, thus making them useful for other developers. These classes have been inspired by the internal architecture of SAGA, overcoming some of its main limitations, such as the lack of flexibility in the design of graphical interfaces or the difficult combination of raster layers from different sources. While retaining its main advantages, the user experience has been improved. Most SAGA modules have been ported to this platform (those developed by the saga core team and also those developed by the Sextante team, which up to this date has been the main contributor to this project), excluding those already implemented somehow in gvSIG, such as file I/O modules. New modules have also been added, and others have been improved.

In the newest version, into Sextante the GRASS toolbox has been integrated as well.

3 Outlook and Discussion

It looks like a big diversity of applications. But, except the two “dinosaurian” GRASS and ILWIS open, which are on the one hand based on a rather old technology, on the other hand mature and well documented, there are many links between the new projects, all emerged in the last 5 to 7 years.

Concerning the supported data formats, most of them are based on the same libraries: OGR [12] for vector data and GDAL [13] for raster format. Both libraries can also be used as a toolkit without a GIS for transformation between different data formats. For coordinate systems and re-projections, most of the OS GIS are using the library proj4 [14].

The main difference between the developments is mainly the programming language used. Still it is argued that Java programs are not as fast as C++ programs working on big datasets. As operating platform, LINUX is loosing some importance; the use of Windows is increasing. MacOS users form a much smaller, but stable group.

To synchronize and promote the development of OpenSource GIS the OSGeo (Open Source Geospatial Foundation) has been founded. The Open Source Geospatial Foundation has been created to support and build

the highest-quality open source geospatial software. The foundation's goal is to encourage the use and collaborative development of community-led projects. For desktop GIS, OSGeo promotes GRASS, QGIS and gvSIG as projects.

As for many other Open Source project one of the major drawbacks is the lack of a good documentation. Free software development is driven by the (developer) needs, not really by the market. The demands fixed in a road-map for further development seems sometimes not very realistic. The focus on further development described inn most of the road-maps is to integrate more geoprocessing functionality and to use the software as elements of a geospatial infrastructure based on the OGC specifications.

Overview of free and open source desktop GIS

		language/ technology	OS	licence	community	dissemination	features	Application focus
jump (Java Unified Mapping Platform) 2002	deeJump				small group of developers, few users	Germany/Europe	WMS, WFS	OGC viewer for WMS, WFS
	openJump 2004				small group of developers, few users	Europe/America	WMS, WFS, PostGIS, digitizing/editing, plug-ins	Viewing, editing
	Kosmo 2006				small group of developers, few users	Spain/Europe	under development	Corporate GIS Platform (desktop, server, web clients, PDA clients,...)
gvsig** (Generalitat Valencia Sistema de Información Geográfica) 2004		Java	Windows, Mac, Linux, ...	GPL	small group of developers, many plug- ins developer, many user	Spain/Europe	WMS, WFS-T, GML3, digitizing/editing, analysis (vector), many plug-ins (SEXTANTE),	Viewing, analysis
uDig (user-friendly Desktop GIS 2004		Java, Eclipse	Windows, Mac, Linux, ...	LGPL	small group of developers, but active many users	North America	visualization, WMS, WFS- T, GML3, Shape, PostGIS, digitizing/editing, (GeoTools)	Viewing, editing
Quantum GIS* (2002)		C++, Qt, Python API	Windows, Mac, Linux, ...	GPL	about 20 core developers, community large and active	worldwide	WMS, WFS-T, OGR-vector formats, GDAL-raster formats, PostGIS, many plug-ins (vector analysis)	Viewing, GRASS-Graphical User Interface
SAGA (System for Automated GeoScientific Analysis) (2001)		C++	Windows, Linux	GPL/LGPL (GUI)	small, academic	Germany/Europe	GDAL-raster, Spatial Analysis (hydrology, kriging, fuzzy, filter, simulation)	Analysis, modeling, scientific visualization
GRASS* Geographic Resources Analysis Support System (1982)		C	Windows, Linux, Mac	GPL	large dev, many users	worldwide	WMS, WFS, OGR-vector	Analysis and scientific visualization, cartography, simulation
ILWIS (1988/2007)		Visual C/ .NET	Windows	GPL	small dev, many users	worldwide	Raster analysis	Analysis

*osGEO project

** in incubation process for osGEO project

SEXTANTE is based on SAGA tools

Links & References:

[1] <http://openjump.org>

[2] <http://www.deejump.com/>

[3] <http://www.opengis.es/>

[4] <http://www.gvsig.gva.es>

[5] <http://oadigital.net>

[6] <http://www.qgis.org/>

[7] <http://udig.refrains.net>

[8] <http://saga-gis.wiki.sourceforge.net/>

[9] <http://grass.osgeo.org/>

[10] http://52north.org/index.php?option=com_content&task=view&id=131&Itemid=319

[11] <http://www.sextantgis.com/en/index.htm#>

[12] <http://www.gdal.org/ogr/>

[13] <http://www.gdal.org/>

[14] <http://trac.osgeo.org/proj/>

Paul Ramsey: The state of Open Source GIS. Sribd, 2007 <http://www.scribd.com/doc/510902/Survey-on-Open-Source-GIS>

Gary E. Sherman: Desktop GIS. Mapping the Planet with Open Source Tools The Pragmatic Programmers.
First Edition Oktober 2008

Workshop on “R” for Geospatial Analysis

Dietrich Schröder

Department of Geomatics, Computer Science and Mathematics,
University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
dietrich.schroeder@hft-stuttgart.de

KEYWORDS:

R, geo-statistics

ABSTRACT

R is an open-source environment for statistical computing and visualisation. It is based on the S language developed at Bell Laboratories in the 1980's, and is the product of an active movement among statisticians for a powerful, programmable, portable, and open computing environment, applicable to the most complex and sophisticated problems, as well as “routine” analysis, without any restrictions on access or use.

There are a number of packages for R to analyze spatial data. In the workshop a short general introduction into R is given. It is shown how geo data – vector as well raster – is loaded into R. In a first example kriging is used to interpolate point data using the gstat package. In a second example the spatstat package is used for point pattern analysis.

Internet Mapping – An Introduction

Franz-Josef Behr

Department of Geomatics, Computer Science and Mathematics,
University of Applied Sciences Stuttgart
Schellingstraße 24, D-70174 Stuttgart, Germany
franz-josef.behr@hft-stuttgart.de

KEYWORDS:

Internet, Web Mapping, Services, Geodatabase, HTTP

EXTENDED ABSTRACT

Besides hardware, local networks, and transatlantic backbones there are just a few pillars which assure that the Internet and the servers provides the functionality and services we are accustomed to, namely the HyperText Transfer Protocol (HTTP), the Multipurpose Internet Mail Extensions (or content types), the identification of resources by IP addresses and host names, the Common Gateway Interface (CG), and the Structured Query Language (SQL) for accessing the databases. In many cases information is transferred in the form of the eXtensible Markup Language (XML).

The interaction between clients and servers through the Web will be introduced and architectural principles explained. You will understand the multi-tier (multi-layer) concept, its components, and its advantages. It will be unrolled that the GeoWeb is the combination and interplay of distributed resources offering geodata and services.

Relevant software packages in this field will be briefly introduced, so that the audience members get a helpful personal overview over the software scene.

In the practical work we will learn to install our own client server and development environment. We will see how easily a Web server (HTTP server) can be installed and administered. We will publish static HTML pages with CSS and JavaScript and then move forward to dynamic provision of resources. On this way we will get in contact with a Database Management System and its databases, touch SQL and load our own geodata.

With a keen step to server side programming we will see how we can extract the geodata in order to share them as a service through the web.

On client side you will learn how a mapping client, in our case based on the OpenLayers application programming interface (API), can be implemented. We will perceive that questions of licensing and availability have to be taken into account if we plan the outroö of our own geodata and geoservice.

In the next part, you will get in touch with Web 2.0, the current state of contemporary web technology if we understand AJAX, i.e. Asynchronous Javascript and XML and its use in conjunction with your mapping client.

In a final step you will conflate client side and server side technology and applications. You will see how base data can be overlaid and enhanced with your geodata, extracted from the database, restructured in the application layer, transferred through the web, and finally visualized in the HTTP browser.

The workshop aims to give a clear understanding of the architecture, protocols, and technologies and how they can be applied in your own geospatial Internet solution, ypir contribution to the GeoWeb.

References

- Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., Yergeau, F., Cowan, J. 2006: Extensible Markup Language (XML) 1.1 (Second Edition). <http://www.w3.org/TR/2006/REC-xml11-20060816/> [2010-06-01]
- Garrett, J. J., 2005. Ajax: a new approach to web application. Adaptive path. <http://www.adaptivepath.com/publications/essays/archives/000385.php> [2009-05-05]
- Fielding, Roy Thomas (2000): Architectural Styles and the Design of Network-based Software Architectures. Doctoral dissertation, University of California, Irvine, <http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm> [2008-02-12]

O'Reilly, Tim (2005): What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software. <http://www.oreilly.de/artikel/web20.html> [2010-07-13]

Poster

Using GIS and Remote Sensing for the Assessment of the Potential Impacts of Flooding on Roads Network in Makkah "Mecca", Saudi Arabia

Meraj N. Mirza and Ramze A. Elzahrany

Center of Research Excellence in Hajj and Ummrah

Umm Al-Qura University, Makkah, Saudi Arabia

KEYWORDS:

Makkah, Saudi Arabia, flooding, road networks, GIS, RS

ABSTRACT:

Makkah is the holiest site for all Muslims. Millions of Muslims arrive to Makkah from all over the World annually. They perform Hajj "pilgrimage", the famous annual worship for those who could afford it once in their life time. The number has increased significantly during the last few years, due to the increase of those who arrive for Umrah (visit to Al-Masjed Al-Haram "the Holy Mosque"). The number of arrivals for pilgrimage and visit has to be taken into consideration with the number of the city population, which exceeds one million inhabitants. The city of Makkah is located in the western part of Saudi Arabia, which is known for its arid and semi-arid climate. Therefore, massive flooding is a rare phenomena in Makkah, but it could be highly devastating when it occurs. It might cause considerable destruction on various local infrastructures. The objective of this research is to assess the potential impacts of flooding on road network in Makkah, based on various data; Statistical data on rainfall and roads, maps, aerial photos, images, etc.

This research argues that new techniques, such as Geographic Information Systems and Remote Sensing provide better alternatives for more accurate spatial planning of road networks. Hence, ArcGis 9.3 and ERDAS Imagine 9.1 have been the principal tools for the study. They were used for modeling the impacts of this natural phenomenon on road networks, as man-made features. Layers of roads network, drainage system network of Makkah, its catchment area, contours, and Digital Elevation Model (DEM) were overlaid. Velocity, speed of running water and height of flow were calculated. Final assessment was made based on the results.

It is hoped that the outcomes of this study will contribute to the enhancement of the decision making process in the area, based on a solid scientific rationale. Consequently, different levels of modifications of road network structures and other related infrastructure accordingly are proposed. This research might be considered as a new classic model for further research in other cities within Saudi Arabia and elsewhere with similar physical conditions.

Plataforma de Procesamiento de Imágenes Satelitales para la Detección de Contaminación Minera.

Juan Carlos Gutierrez Cáceres¹ and César Beltrán Castañón^{1,2}

¹ Universidad Nacional de San Agustín – Arequipa

² Universidad Católica de Santa María

ABSTRACT:

El presente trabajo propone desarrollar una plataforma de procesamiento de imágenes satelitales que permita detectar regiones contaminadas con relaves mineros permitiendo crear una base de datos temporal que permita realizar un seguimiento del problema de contaminación. Una característica fundamental del proyecto es su interactividad basada en la tecnología de Cliente/Servidor, buscando promover el desarrollo de software como servicio. Aunque el presente proyecto está orientado a analizar la contaminación minera, al desarrollar una plataforma de análisis y procesamiento de imágenes satelitales, ello permitirá expandir la plataforma hacia otras aplicaciones de similar interés (cultivos, caudales, georeferenciamento, entre otros).

Sistema de Seguimiento de Cuidado de la Salud, basado en tecnología Geoinformática: Caso madres gestantes.

César Beltrán Castañón and Agueda Muñoz del Carpio Toia

Universidad Católica de Santa María

ABSTRACT:

En nuestro medio no existe una política adecuada de prevención de la salud, especialmente por el hecho que no se cuenta con las herramientas tecnológicas adecuadas para respaldar dicha política. El presente trabajo detalla el desarrollo de una plataforma computacional vía web, que permita desarrollar un seguimiento del estado de salud de pacientes, con el objeto de realizar prevención, para ello se propone utilizar tecnología geoinformática, georeferenciamiento, que de forma gráfica permita tener un control sobre pacientes que son atendidos por un centro de salud. La aplicación está basada en el control de madres gestantes, pero también se propone el uso de la herramienta en pacientes de VIH y tuberculosis de la Región Arequipa. La aplicación de la plataforma permitiría que el cuerpo médico se anticipe a la ocurrencia de males y/o realizar un chequeo permanente, lo cual deberá incidir en la mejora de salud y/o disminución de riesgos de salud.

Developing a Community Forum using Ushahidi

Emmanuel Bobb Anya Lansana, M. Abdullah Sohl, Franz-Josef Behr

Department of Geomatics, Computer Science and Mathematics,

University of Applied Sciences Stuttgart

Schellingstraße 24, D-70174 Stuttgart, Germany

KEYWORDS:

Community Mapping, Ushahidi, Crisis Mapping, Social Networks

ABSTRACT:

Ushahidi, a Swahili word meaning "Testimony", was originally developed in 2008 by a group of developers in Kenya. The purpose was to map eye-witness reports of post-election violence after Kenya's 2007 Presidential elections. Today, Ushahidi has gained dramatic attention for various forms of crisis mapping, election monitoring and several other web mapping applications. It provides a flexible way of information collection, visualization and interactive mapping. With Ushahidi, information collection is not limited to internet users; it can be synchronized with Frontline SMS to collect information from ordinary mobile phone users which can then be mapped. It can also be linked to other social networking sites like Facebook, Twitter, etc for larger crowd sourcing. Ushahidi is free and open source. The Haiti Crisis Map, Sudan Vote Monitor and Chile Crisis Map are just a few of recent implementations of Ushahidi.

In this study project, the primary objective of implementing Ushahidi was to connect past and present members of the Stuttgart University of Applied Sciences.

A Framework Experiences of a Distance Education Program for GIS in Latin America. UNIGIS 2000 – 2010

Carla M. Valdospinos Navas

UNIGIS, Universidad San Francisco de Quito

unigis.carla@gmail.com

KEYWORDS:

ABSTRACT:

The presented study describes how GIS education takes importance in the context of a distance post graduated program in Latin America. The program structure and curriculum can be read at www.unigis.edu.ec. Our objective reflects based upon real numbers the characteristics of professionals, study activities and the applications where SIG is being used. In this context an interesting compilation of information to future learning models, actual programs, regarding GIS and the management of geospatial information with geo-technologies in Latin America is presented.

1 Introduction

UNIGIS²⁶ sees one of their main targets in transmitting a basic known of GIS, and also cartographic techniques according a rapidly growing community of professionals that come from disciplines and fields of application traditionally quite unrelated to spatial concepts and their means of cartographic representation.

2 Theoretical foundation and results

GIS Education demand in Latin America has grown enormously in the last years. At 2000 when the distance program of education in SIG UNIGIS started, there was a group of 24 students. However in the current year just for the first of two UNIGIS cycles for Latin America there are 48 enrolled students and a greater number of expecting for the second cycle.

The demand by sectors of the studies in GIS has a great variety; there are enrolled professionals from the fields of biology, civil engineer, engineering in systems, agronomists, environmental engineering, administration and accounting, teachers; like the big groups but there are still different kind of professionals in small groups.

Another characteristic in the group of GIS students and applicants is the gender. At the beginning (first 5 years), there were 19 (22%) women and 69 (78%) men, and it has been a small growth of women the next 5 years, with 70 (35%) women and 132 (65%) men, but there is still a bigger number of men enrolled.

GIS application demands has also grown; with a new demand of applications the specializations UNIGIS L.A. offers have had to also increase. A new offer will be applied since 2011 and it would include the following applications:

- GIS for Public health
- GIS Management of Natural Risks
- GIS for Territorial development
- GIS for communal services
- GIS for precision agriculture
- GIS for Open Source applications

²⁶ Worldwide operating network of universities offering regionally specific postgraduate distance learning programs with a common core curriculum since 1991. See www.unigis.net.

In another context the structure of applications may differ for the next years, but the cartographic representation is what makes GIS communication the same language and enables an informed dialog between the experts making “common” user applying spatial information for his daily tasks in conserving, managing and designing space and spatial interactions.

3 Results and Conclusions.

As the spatial domain of information goes main stream, and new user groups emerge out of different disciplines and interests, a common learning experience is difficult to establish. Students ask for a very flexible syllabus, meeting just their needs both in content and modality. GIS in Education stands in a constant conflict between technology and content driven approaches. It is not easy to satisfy all student’s needs at the right time. Common concepts never become really obsolete, but nevertheless they might not reflect the demands to effectively apply learned methodologies to new applications.

The biggest challenge is to create an own network for Latin America, with a constantly building out of the student’s very specific own needs. These will be considering particular cultural contexts of all regions strengthening both the local and regional identity and not falling behind on a rapidly growing global one.