Virtual globes and geospatial health: the potential of new tools in the management and control of vector-borne diseases

Anna-Sofie Stensgaard[1,2], Christopher F.L. Saarnak[2], Jürg Utzinger[3], Penelope Vounatsou[3], Christopher Simoonga[4], Gabriel Mushinge[4], Carsten Rahbek[1], Flemming Møhlenberg[5], Thomas K. Kristensen[2]

[1]*Center for Macroecology and Evolution, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark;* [2]*DBL, Department of Veterinary Disease Biology, University of Copenhagen, Thorvaldsensvej 57, 1871 Frederiksberg C, Denmark;* [3]*Department of Public Health and Epidemiology, Swiss Tropical Institute, P.O. Box, CH-4002 Basel, Switzerland;* [4]*Department of Community Medicine, University of Zambia, P.O. Box 32372, Lusaka, Zambia;* [5]*DHI, Agern Allé 5, DK-2974 Hørsholm, Denmark*

Abstract. The rapidly growing field of three-dimensional software modeling of the Earth holds promise for applications in the geospatial health sciences. Easy-to-use, intuitive virtual globe technologies such as Google Earth™ enable scientists around the world to share their data and research results in a visually attractive and readily understandable fashion without the need for highly sophisticated geographical information systems (GIS) or much technical assistance. This paper discusses the utility of the rapid and simultaneous visualization of how the agents of parasitic diseases are distributed, as well as that of their vectors and/or intermediate hosts together with other spatially-explicit information. The resulting better understanding of the epidemiology of infectious diseases, and the multidimensional environment in which they occur, are highlighted. In particular, the value of Google Earth™, and its web-based pendant Google Maps™, are reviewed from a public health view point, combining results from literature searches and experiences gained thus far from a multidisciplinary project aimed at optimizing schistosomiasis control and transmission surveillance in sub-Saharan Africa. Although the basic analytical capabilities of virtual globe applications are limited, we conclude that they have considerable potential in the support and promotion of the geospatial health sciences as a user-friendly, straightforward GIS tool for the improvement of data collation, visualization and exploration. The potential of these systems for data sharing and broad dissemination of scientific research and results is emphasized.

Keywords: geospatial health, virtual globes, Google Earth™, geographical information system, vector-borne disease, schistosomiasis, sub-Saharan Africa, CONTRAST.

Introduction

An important aspect of epidemiology is the study of risk factors for an infection and disease-related morbidity and mortality. The identification of

Corresponding author:
Anna-Sofie Stensgaard
Center for Macroecology and Evolution
Department of Biology, University of Copenhagen
Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark
Tel. +45 3532 1211; Fax +45 3532 2321
E-mail: asstensgaard@bio.ku.dk

causal pathways between risk factors and disease in turn allows the design and implementation of preventive and control measures which aim to reduce disease burdens. Infectious diseases are often considered "environmental" diseases because a considerable fraction of their burden can be attributed to environmental factors (Listorti and Doumani, 2001; Prüss-Üstün and Corvalan, 2007). Hence, distribution patterns of infectious diseases are strongly associated with the spatially heterogeneous environment in which they are entranced (Woolhouse et al., 1997; Brooker and Clements, 2009).

Visualization of this heterogeneity, at different spatial scales, is paramount in revealing new insights into the patterns of disease. The simultaneous visualization of health data with environmental data obtained from a diversity of sources thus holds promise to further our understanding of environmental-health linkages and can generate new hypotheses to be tested in future research. Geographical information systems (GIS) are designed for this purpose and contain an increasing number of sophisticated capabilities for data display and analysis. The use of GIS in spatial epidemiology in human and animal health and the geospatial health sciences in general have been firmly established as a useful tool for collating, exploring, visualizing and analyzing health data in a spatially explicit manner (Yang et al., 2005; Rinaldi et al., 2006; Brooker, 2007). The integration of remotely-sensed environmental data into a GIS platform can assist in a better understanding of the spatio-temporal dynamics of a wide range of disease systems, especially those with environmental correlates. However, high-resolution, geo-rectified imagery in a digitized format is difficult to visualize and explore without the expertise and availability of sophisticated (and often expensive) GIS software. This remains a hurdle for fully harnessing GIS for elucidation of spatial epidemiology and geospatial health, especially in a developing world context.

Recent advances in the development of virtual globe technology has provided an opportunity for a cheap and accessible method to communicate epidemiological data more effectively to non-specialists, as well as among scientists. Virtual globe technologies, such as Google Earth™ are essentially web-based GIS tools which brings some of the functionality of applied GIS to the non-specialist. It is thus suitable for the display and dissemination of research results where location and spatial variation are critical components. While virtual globe applications have limited analytic functions and are not designed to replace professional GIS software, they may be a useful complement to traditional GIS technologies and geostatistical analyses tools. In our opinion, the virtual globe technology

holds a large and hitherto under-explored potential for applications in the health sciences, especially in low-income countries.

This article reviews recent advances in the application of the Google Earth™ virtual globe (and its web-page pendant, Google Maps™) for epidemiological purposes. The specific objectives are: (i) to review existing scientific literature and available Internet resources (web-based research projects), and (ii) to illustrate its applicability within the context of an EU-funded multi-partner research project to improve the management and transmission control of schistosomiasis in sub-Saharan Africa. Finally, the potential of virtual globe technology for surveillance and control of vector-borne and other infectious diseases that depend on intermediate hosts are discussed and a set of conclusions drawn from the examples presented.

Virtual globes and Google Earth™: brief background

A virtual globe is essentially a three-dimensional (3-D) representation of the Earth, usually based on satellite imagery, upon which various types of information with a spatial character can be superimposed. A virtual globe provides the users with the ability to add their own data, to share the added data layer with other users, and to freely move around in the virtual environment by zooming and changing the position and viewing angle. The first widely publicized virtual globe was Google Earth™ and there are a growing number of publications using Google Earth™ for display purposes and a host of applications as discussed later. There are also several other virtual globe technologies currently available online such as NASA's WorldWind (NASA Ames Research Center, 2004, http://worldwind.arc.nasa.gov/), ESRIs ArcGIS explorer (ESRI, http://www.esri.com/software/arcgis/explorer/index.html), Microsoft's Virtual Earth (http://www.microsoft.com/virtualearth/) and Free Earth by the Poly9 Group (http://freeearth.poly9.com/). A thorough review of

available online geo-informatics services and other virtual globes has been published by Boulos (2005). The current paper emphasizes Google Earth™ due to its dominance among the virtual globe technologies and it is not within the scope of this paper to describe all currently available virtual globe technologies or to discuss their strengths and limitations in a comparative way.

Launched without much publicity, the prominence of Google Earth™ is now manifested by peaks of 12 million hits per hour (Blamont, 2008). Perhaps Google Earth's biggest impact is the opening of the exploration of spatially-explicit data to a large audience of mainly laymen. Google Earth™ was originally called Earth Viewer and was developed by Keyhole Inc., a company acquired by Google in 2004 which launched it as a virtual globe programme in 2005. Essentially, Google Earth maps the world by superimposing satellite raster imagery, aerial photography, vector maps and other layers, in a single and integrated tool, allowing users to interactively "fly" in three-dimensional space, zooming from global to regional and local levels. Although Google Earth™ is primarily aimed at the general public, and mainly as a search tool, it has also attracted a large community using the application for a wider array of applications and purposes. Google Earth™ is a stand-alone application, and the user is required to install the software before it can be used. The web-based pendant to Google Earth™, Google Maps™, is a complete webpage with no need for the users to install any additional software to be able to browse the maps displayed by the provider. Like Google Earth™, Google Maps™ can include satellite imagery, but are limited to a two-dimensional (2-D) view – without Google Earth's 3-D "tilt" feature. Thus, Google Maps™ is essentially not a virtual globe; however, it is included in this review as it is often used in parallel with Google Earth™ on research project web pages to maximize outreach and target a wider audience.

The primary method for visualizing data in Google Earth™ is through the creation of Keyhole markup language (KML) files for managing and displaying 3-D geospatial data, an approach developed by Keyhole Inc. The most recent version of KML (version 2.1) contains many features relevant to scientific data, such as large data support and the ability to timestamp features and to create animations. To our knowledge, Google Earth™ is currently the only virtual globe technology that fully supports all of these features, but KML is supported by a number of other virtual globe applications and GIS packages and is therefore already becoming a *de facto* standard. There are numerous ways to produce KML files for Google Earth™, with detailed documentation and tutorials freely available online (http://earth.google.com/userguide/v4/ug_kml.html). New tools and add-ons to existing software packages are increasingly being developed to make this task as quick and easy as possible. Users of ESRI's ArcGIS products (version 9.2 and onwards) can now export their vector and raster layers or complete map projects directly into KML file format. The users of the Matrix Laboratory (MATLAB) platform can export results of geostatistical analysis to Google Earth™ by using the Google Earth™ Toolbox. In addition, other large software developers are now integrating their own software with Google Earth™. Of particular relevance for flood-related disease management, DHI Water and Environment (http://www.dhigroup.com), an independent research and consulting organization, has developed a tool that facilitates the installation of flood maps from their MIKE FLOOD software (DHI, 2005) in Google Earth™ to visualize and animate models of different flooding scenarios.

Google Earth™ and the scientific community

Virtual globe applications in general, and Google Earth™ in particular, are increasingly being recognized by the scientific community as far more than just a handy map or a fun toy (Butler, 2006b). They are becoming a meeting place for scientists and non-scientists alike, offering a way to connect and share data with each other, and exchange scientific ideas. To obtain an overview of the growing use of Google Earth™ in the peer-reviewed literature, we searched

two widely used electronic databases, namely (i) ISI Web of Science (http://apps.isiknowledge.com) and (ii) US Medical Library of Medicine and the National Institutes of Health's PubMed/MEDLINE and Pubmed Central databases (http://www.ncbi.nlm.nih.gov/pubmed/) by employing the keyword "Google Earth". Since we are also interested in research employing Google Maps™, we did a cross search with keyword "Google Maps". Our search was conducted on February 1, 2009, being restricted to the years 2005 to 2008. Since the launch of Google Earth™ in mid-2005, a steady increase of publications has been noted. In total, 123 publications with reference to the application of Google Earth™ and/or Google Maps™ were identified (Fig. 1). Of these most pertained to the use of Google Earth™ only, while about 15% deal with Google Maps™ and about 7% explicitly makes use of both Google Earth™ and Google Maps™. The first publications with explicit reference to Google Earth™ mainly discussed this new tool in its own right. Since then, there has been a steady increase in publications, with more than 70 publications alone in 2008 pertained to actual application of Google Earth™ for scientific purposes.
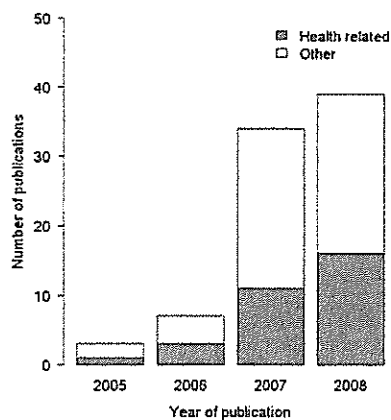

Fig. 1. The number of published articles since 2005 with reference to Google Earth™ and/or Google Maps™ is steadily growing (ISI Web of Science and PubMed/MEDLINE). Health-related work (filled sections) has contributed around 30% of the total number of published papers (total bar) since the initiation of Google Earth™ in mid-2005.

The application pertain to a wide variety of scientific disciplines, including geology, paleontology (Conroy et al., 2008), environmental management (Pearce et al., 2007; Nethery et al., 2008), conservation (Bruno et al., 2007; Aburto-Oropeza et al., 2008) and medicine (Mikula et al., 2007; Weibel, 2008). Not only the field of application, but also the tasks show a wide range, from the visualization of earthquakes and tsunamis (Yuan et al., 2008), oasis detection in deserts (Luedeling and Buerkert, 2008), estimating the size of irrigated agriculture (Thenkabail et al., 2007), as an educational tool (Doering and Veletsianos, 2007; Patterson, 2007) for global biodiversity assessments (Guralnick et al., 2007), community mapping (Lefer et al., 2008) and to tracking polio virus down the Congo River (Kamadjeu, 2009). The field where Google Earth™ has seen the most extensive use thus far is in the environmental sciences. However, also within the health sciences and public health-related projects in particular, the use of Google Earth™ is growing, as witnessed by the increasing number of articles published in the peer-reviewed literature (Sundvall et al., 2007; Boulos et al., 2008; East et al., 2008; Freifeld et al., 2008; Gao et al., 2008; Lefer et al., 2008; Lozano-Fuentes et al., 2008; Yi et al., 2008; Kamadjeu, 2009). One of the earliest references made to the potential of virtual globe technology for disease mapping was highlighted by Malone in 2005 (Malone, 2005). In the literature searched, the use of Google Earth™ in the health sciences accounted for approximately 30% of all published work.

In addition to its direct application in scientific publications, Google Earth™ is also being used retrospectively by creating and publishing Google Earth™ KML files of key findings to supplement the scientific publications and broaden the dissemination of knowledge. An example of particular relevance to the topic of this paper is a collaborative project between epidemiologists looking at the impact of climate change on vector-borne diseases and Google. Based on the findings from several separate studies (Hales et al., 2002; Tanser et al., 2003;

Ezzati et al., 2004; Ebi et al., 2005), these research groups created a KML animation named "Global Warming and Disease" (http://blog.google.org/2008/04/global-warming-heats-up-infectious.html) that shows the projections for the changes in the global transmission of dengue fever and of malaria in Africa. Furthermore, large amounts of data are becoming available, often in real time, in formats that can be displayed by virtual globe applications, and often in parallel with custom website Google Maps™. For example, a complete database dealing with the distribution of African vertebrate species from the Zoological Museum of Copenhagen (Galster et al., 2007; Hansen et al., 2007a,b; Rasmussen et al., 2007) can be browsed and visualized through a Google Maps™ interface on the Museum's web-page (http://130.225.211.158/subsaharanafrica/subsaharan.htm). Likewise, millions of animal and plant species distributions from the Global Biodiversity Information Facility, headquartered in Copenhagen, are available for visualization in Google Earth™ (http://ge.gbif.net/).

Web-based health projects

A search through the Google Internet browser for web-based projects using the search terms "Google Earth", "spatial", "health" and "epidemiology" revealed several official websites currently using either Google Earth™ or Google Maps™ to display health-related data. We identified at least 17 projects relevant to the topic of this paper that currently have health maps or data available through project or institutional websites. The projects deal with a variety of epidemiological subjects and diseases, from malaria data display in the Malaria Atlas Project (MAP) (http://www.map.ox.ac.uk), to the 'mashing' together genetic and epidemiological data on pathogenic micro-organisms in the Spatialepidemiology.net facility (http://www.spatialepidemiology.net). The latter not only provides a Google Earth™ map-based interface for the display and analysis of epidemiological data, but also allows the user to create their own maps easily through the Google Maps™ application programming interface (API). A list of links to a selection of projects has been summarized in Table 1.

Table 1. List of official web-based projects where Google Earth™ and/or Google Maps™ is used to display infectious disease data, e.g. as part of decision support systems or to disseminate the results of research projects.

Organization/project	Subject/disease(s)	Link to project web-page
CONTRAST	Schistosomiasis in sub-Saharan Africa	http://eu-contrast.eu
Epispider.org	Global disease maps	http://www.epispider.org/index.php
Harvard-MIT Division of Health Sciences & Technology and Children's Hospital Informatics Program	Global disease alert map	http://www.healthmap.org/en
Imperial College London	Pathogenic micro-organisms, Batrachochytrium dendrobatidis	http://www.spatialepidemiology.net/
Kenya Medical Research Institute (KEMRI) and the University of Oxford	The Malaria Atlas Project (MAP)	http://www.map.ox.ac.uk/MAP_data.html
National Cancer Institute, US National Institutes of Health	Cancer	http://li-gis.cancer.gov/maps/
Nature	Avian flu	http://www.nature.com/news/2006/060105/full/news060105-1.html
RAMS-AID, Colorado State University	Dengue Decision Support System Project	http://www.rams-aid.org/DDSS/ddss_GE.php
RAMS-AID, Colorado State University	West Nile Virus Decision Support System	http://www.ivcc-ddss.org/mapserver/rmwtLarimermap.php
Wildlife Disease Information Node	Global Wildlife Disease	http://wildlifedisease.nbii.gov/wdinNewsDigestMap.jsp
Google.org (Google Earth outreach initiave)	Various	http://earth.google.com/outreach/p_health.html

A detailed analysis of the individual projects is not the scope of this paper. For this, the reader is referred to the Google Earth Outreach initiative website (http://earth.google.com/outreach/index.html) that contains an updated online overview of public health related showcases (http://earth.google.com/ outreach/p_health.html). One example worth highlighting, however, is the scientific journal Nature's use of Google Earth™ to track the spread of the H5N1 avian influenza virus around the globe (Butler, 2006a), a project that won the Association of Online Publishers (AOP) Use of a New Digital Platform Award 2006. This Google Earth™ presentation compiles data on outbreaks of avian flu in birds from 2003 and onwards, confirmed human cases of infection, as well as other relevant spatial data layers, in order to map cases and outbreaks by location and time, with links to relevant web resources from the United Nation's Food and Agriculture Organization (FAO), the World Health Organization (WHO), and other organizations (Butler, 2006a). This example demonstrates the potential of the so-called "mashups" in Google Earth™, a term originally used to describe the mixing together of musical tracks, but now referring to websites that weave data from different sources into a new service (Boulos et al., 2008). A link to the project web-page and the KML file can be found in Table 1.

Using Google Earth™ to map and monitor schistosomiasis in sub-Saharan Africa

The CONTRAST project

Google Earth™ and Google Maps™ have been applied within the framework of the EU-funded project which carries the acronym CONTRAST (http://www.eu-contrast.eu/). The project aims to build a multi-disciplinary research platform to better understand which interventions to be used to control the snail-borne parasitic disease schistosomiasis and how to tailor interventions at the local level. The morbidity of this disease is predominantly controlled by chemotherapy campaigns, but in order to achieve enduring success in reducing transmission, a thorough consideration of the environmental components is essential since the disease is completely dependent on its freshwater-dwelling intermediate host snail. The CONTRAST project is multidisciplinary, bringing together key skills and expertise to generate new knowledge on biological, environmental and socioeconomic factors relating to schistosomiasis in different parts of Africa. It consists of a strong research node network across Africa. These nodes are working on establishing innovative molecular tools to characterize both snails and schistosomes, defining the importance of host-parasite dynamics across different eco-epidemiological settings and developing new spatial models for disease risk mapping and prediction. An additional aim of CONTRAST is to encourage and assess novel local control interventions using a social science approach, while ensuring widespread dispersal and access to information.

Virtual globe technology lends itself particularly well to the activities within the CONTRAST project that are concerned with the spatial epidemiology for schistosomiasis risk mapping and prediction at non-sampled locations. At the moment, Google's virtual globe technologies are used for a variety of purposes, ranging from partner communication, data visualization and validation and to the dissemination of project research and results. The conceptual diagram in Figure 2, illustrates where and how Google Earth™ and Google Maps™ are used in the project work flows and stages.

The overall communication and assembly of project-related data takes place via the CONTRAST web-page (Fig. 2; stage 1), where Google Earth™ and Google Maps™ are used to communicate in a spatially-explicit manner. New data collected by the project partners during epidemiological and malacological surveys can be entered online using a standardized format managed by the FireFlower data management system (http://www.fireflower.ca). Historic data on schistosome parasites and intermediate host snails for all of sub-Saharan Africa are continuously being collated and digitized using a
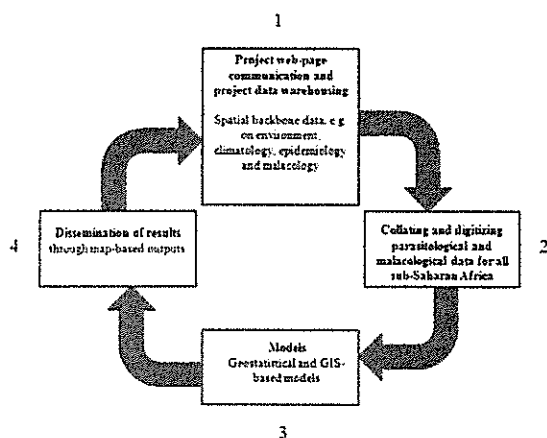
Fig. 2. Conceptual diagram showing the work-flow and stages in the EU-funded CONTRAST project, relating specifically to the spatial epidemiology for schistosomiasis risk mapping and prediction at non-sampled locations. The red boxes indicate processes where Google's virtual globe technologies are applied.

systematic approach to review extant literature and managed in a relational database with direct KML export capabilities (Fig. 2; stage 2). Here, the Google Earth™ display of data can be used for data location geo-referencing, validation and gap identification. The development of predictive models of disease distribution takes place via spatial and geo-statistical analysis (Fig. 2; stage 3). As Google Earth™ has no analytical capabilities, this process essentially takes place through statistical software packages (STATA, Stata Corporation, College Station, TX, USA, version 9.2; WinBUGS, Imperial College & Medical Research Council, London, UK, version 1.4.2; and R, http://www.R-project.org) as well as more conventional GIS packages. Finally, in addition to the common dissemination of results through scientific literature, models and other research results can be readily illustrated via Google Earth™ (Fig. 2; stage 4). The extensive public interest in such visualizations makes scientific information widely accessible to general users. Personal data and results can be "mashed up" with other spatial data and thus stimulate scientific debate and the articulation of new research questions and hypotheses. Data visualization is also useful for decision

support, particularly the spatial targeting of disease control measures, but also the identification of knowledge gaps and the stimulation of further data collection.

Enhancing partner communication and data sharing

Three examples are provided here that illustrate how Google's virtual globe technologies are used in the CONTRAST project. They relate to the specific work-flows and stages depicted in the diagram of Figure 2. The first example pertains to project management and information exchange. Of note, the CONTRAST project brings together not less than 14 partner institutions, four of them located in Europe (Belgium, Denmark, Switzerland and UK), and the remaining 10 based across sub-Saharan Africa (Cameroon, Kenya, Niger, Senegal, Tanzania, Uganda, Zambia and Zanzibar). The management of such a large body of institutions and in variety of settings is not only a challenge from a logistics point of view, but also comprises a challenge with regard to communication. To enhance information exchange among the CONTRAST partners and the scientific community at large, a website was created at the onset of the project. This website is hosted by the University of Copenhagen, and managed by the project coordinator unit there. On the website, a mapping service is available which has been developed using the Google Maps™. This gives CONTRAST partners and the public an opportunity to follow the progress and development of the ongoing research. On the CONTRAST web-page Google Maps™ is used to show partner and research node locations and associated information. In this way, partners and other stakeholders can gain quick updates with regard to the status of the fieldwork that takes place throughout the year at the various study locations. This is possible because all partners in the CONTRAST project have been equipped with suitable hand-held global positioning systems (GPS) devices ensuring that the recording of geographical coordinates is a standard part of any fieldwork
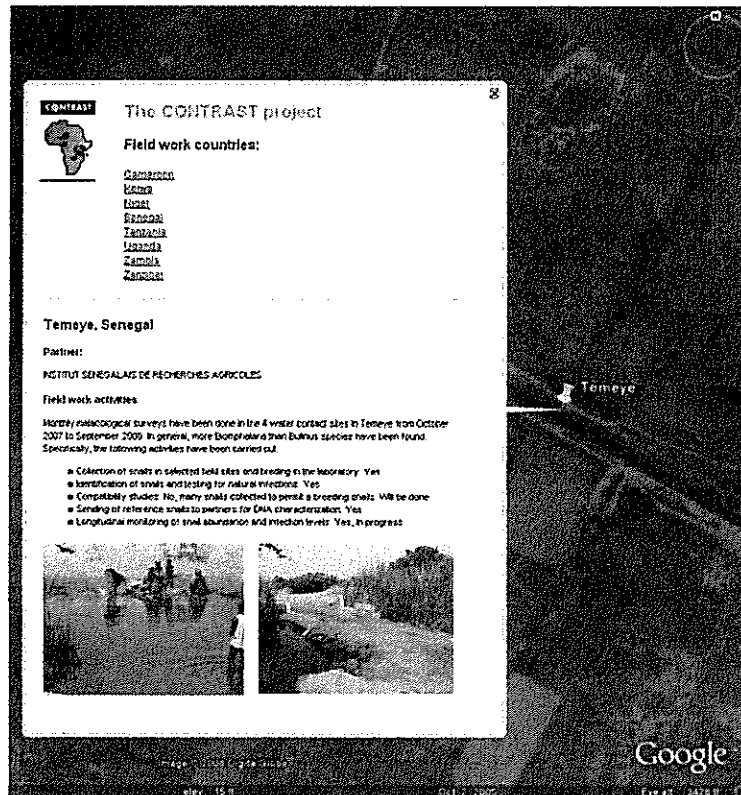
Fig. 3. An example of the use of Google Earth™ to "fly" the user to fieldwork locations of interest, and illustrate the most important findings from the given fieldtrips by means of attached info-balloons. The user can either download a full Google Earth™ presentation of the on-going CONTRAST-related field work as a KML file or view directly on the project web-page through Google Maps™.

exercise. An example of a spatial illustration of fieldwork in progress can be seen in Figure 3.

Developing a digital, interactive atlas of the distribution of schistosomiasis and host snails

The second example pertains to the role of Google Earth™ in the efforts to develop a digital, interactive atlas of the distribution of schistosomiasis and intermediate host snails. It is thus a logical continuation of tremendous efforts made by the University of Bordeaux (France) and the WHO in the 1980s that resulted in the publication of the first global, printed atlas of schistosomiasis (Doumenge et al.,

1987). A major shortcoming of this atlas, which is due to its appearance before the "age of the Internet", is that it cannot be readily updated as new data become available. However, with the advent of more advanced GIS facilities in the mid-1990s, the WHO has launched a new effort to develop a global digital helminth atlas (Brooker et al., 2000).

The value of disease atlases as a means of visually revealing spatial patterns of disease occurrence has been emphasized (Lawson and Williams, 2001), and since the inception of the CONTRAST project in October 2006, considerable efforts have gone into digitizing and geo-referencing historical
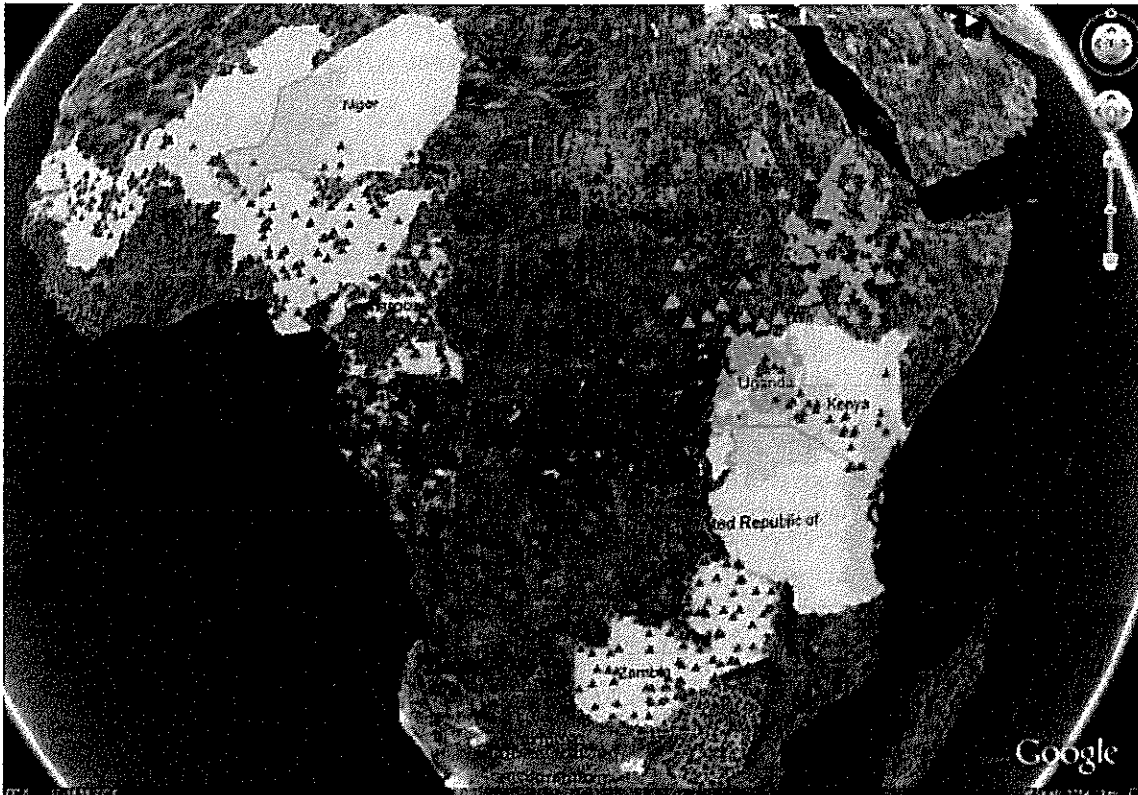
Fig. 4. A Google Earth™ illustration of the locations where data records of *Schistosoma haematobium* (red placemarks) and *Schistosoma mansoni* (green marks) are currently available in the CONTRAST database. The points are overlaid on the CONTRAST African partner countries and the lake and river basins covered by the project. The information is contained in a single KML file.

records of schistosomiasis prevalence and intermediate host snail species. These data are being assembled in the open source MySQL relational database management system, and continuously updated. Currently the database consists of more than 4,000 geo-referenced locations with associated parasite prevalence data along with information on the distribution of the intermediate host snail species associated with the transmission of schistosomiasis. Geo-referencing, the process of converting text descriptions of locations to computer-readable geographical locations (e.g. latitudes and longitudes), is carried out using gazetteers such as the BioGeoMancer (http://www.biogeomancer.org/)

(Guralnick et al., 2006), which uses the Google Maps™ API. Google Earth™ can be used to capture point and polygon coordinates, and also provides a measure tool that helps to find precise locations based on location descriptions (for instructions how to use these tools, the reader is referred to the Global Biodiversity Information Facility's Biodiversity Data Portal http://ge.gbif.net/gbifcap-co.php). The CONTRAST database is planned to be made publicly available, using Google's virtual globe technology as part of the database interface. A snapshot of some of the database contents can be viewed in Figure 4.

The advantages with this approach is that any

person working with the database can directly export any chosen data selection directly from the database to a KML file for instant viewing in Google Earth™. The files can be e-mailed to colleagues (who may have no knowledge of or access to GIS), and then simply "dragged and dropped" by the recipient onto their own desktop Google Earth™ display, where the layer is immediately draped over the Google Earth™ landscape. This simple and instant mapping in Google Earth™ serves several purposes. It has made it possible to quickly identify spatial data gaps and been helpful in identifying areas where further epidemiological and malacological surveys are warranted. It has also proven instrumental in validating the global position of the thousands of localities (small rural settlements, towns, etc.) that have been retrospectively geo-referenced: partners in various geographical regions can quickly indentify geo-coding errors, and hence assist with improving the precision of the geographical coordinates.

Visualizing, exploring and disseminating project information and results

The third example highlights how mapping exercises of relevance for geospatial health applications can be "mashed up" and presented. Suppose that data have been collected and analyzed and predictions generated, the next obvious step – besides scientific publication – is the preparation of the data and results for web-sharing. In the CONTRAST project this indeed is one of the main objectives. We conjecture that one eminently appealing option for sharing the results of such a mapping project to a wider community is via Google Earth™.

In CONTRAST, a KML mashup of relevant project information together with data, maps and models on parasite and intermediate host snail distributions has been put together for Google Earth™ virtual globe presentation (Fig. 5). Any information and layers of choice can be written into the KML, making it possible to illustrate and disseminate mul-
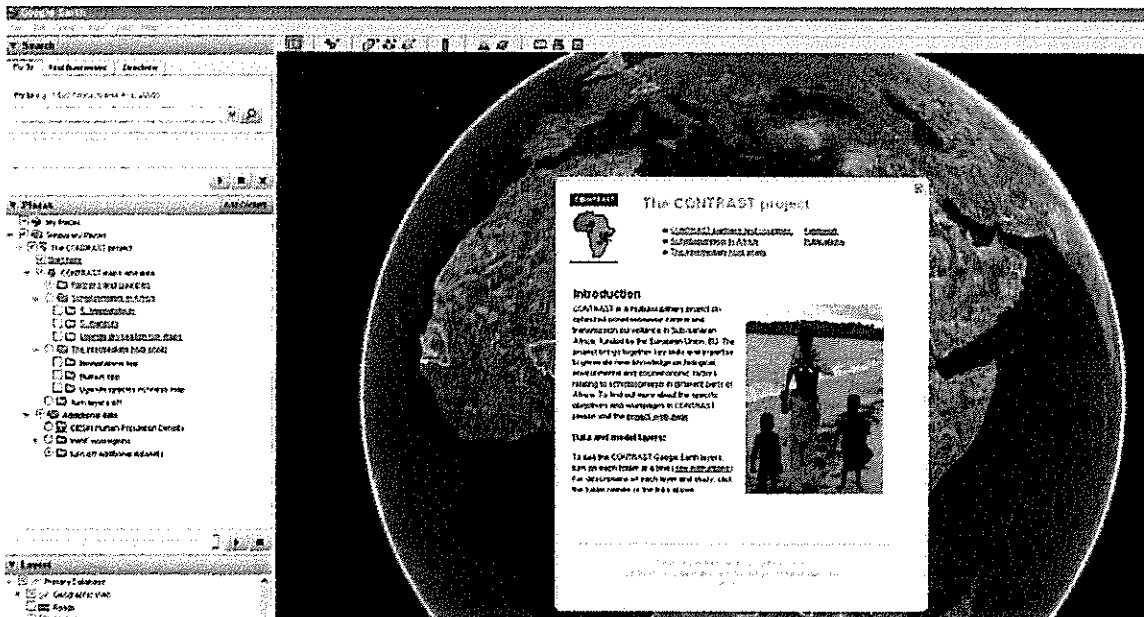

Fig. 5. A screenshot showing a snapshot of the CONTRAST KML mashup. The user is guided through project-related information, data and maps through a system of information balloons with links, and can decide which data layers to view through the side pane to the left. A link to the full KML can be found on the CONTRAST web-page (http://www.eu-contrast.eu/).

tiple information and results within a single file. In the CONTRAST KML presentation, the user is first guided through some introductory information on the CONTRAST project and background information on schistosomiasis and the intermediate host snail species. Through spatially referenced information "balloons", an introduction is also given to each project partner and research node. The next part of the presentation gives an overview of the fieldwork that has been undertaken by "flying" the user directly to the fieldwork locations – and providing additional information on findings, relevant links to publications, etc.

The central part of the KML pertains to outputs from the CONTRAST databases on parasitology and malacology, along with various other spatially-explicit data layers (e.g. environmental data). These can be CONTRAST-generated layers, or spatial layers from other sources that could be of relevance for viewing together with the CONTRAST data. Finally, various map outputs from the geospatial analysis that has been undertaken can be viewed, along with short descriptions of the main findings and links to any available published papers. The user is guided through all the information, data and results contained in the KML through the information balloons that can be linked to geographical locations, relevant web links or other KML files. It is also possible to navigate using the side pane in the Google Earth™ user interface, to switch on and off the individual layers. Layers of interest to the user can be viewed individually or together with other layers in the CONTRAST KML.

The key strength of the virtual globe technologies is the ease with which data can be incorporated from a number of different providers and visualized simultaneously to indentify relationships for subsequent quantitative investigation. Accordingly the layers comprising the CONTRAST mashup can easily be viewed in conjunction with other spatially distributed data to further explore data and model results. In addition to the data already in the CONTRAST KML file, large amounts of environmental data are now available in KML format from a number of projects and data centres, including Kings College Geodata portal (http://www.kcl.ac.uk/schools/sspp/geography/resea rch/emm/geodata/), the Reading e-Science Centre (http://www.resc.rdg.ac.uk/), among others.

Discussion and conclusions

The ease, efficiency, and speed of data communication and analyses are paramount to, and characteristic of, any mature science. The key strengths of virtual globe applications are their simple, intuitive nature and ability to incorporate new data in a straightforward manner. GIS software is already an important tool for understanding spatial and temporal factors in a wide range of disciplines, increasingly so in geospatial health research investigating links between diseases and the multidimensional environment in which they exist. However, commercial GIS tools have traditionally been an expensive and complex solution, especially in the developing world. They are often not mutually compatible, making it difficult to combine data from different sources in a smooth manner. Google Earth™ and other virtual globe applications offer researchers a simpler alternative to GIS software and it is probable that this will lead to increased data sharing (beyond static images) while enabling the implementation of a new and exciting science. Thus, Google Earth™ has the potential of making mapping accessible to a new set of public health users in developing countries. The availability and quality of satellite imagery, combined with features such as KML or image overlay provide a flexible but yet powerful platform that set it apart from traditional mapping tools. It should, however, be emphasized that virtual globe applications are engineered to do only a small portion of what a full GIS technology does, and should be viewed as a complement to, rather than a full replacement of more sophisticated GIS technologies.

The following points are offered for discussion regarding some of the merits and limitations of Google Earth™ as compared to other virtual globe

technologies and GIS. First and foremost, Google Earth™ allows easy simultaneous visualization of point data together with many of types of auxiliary environmental data making it well-suited for "exploratory" phases of scientific work. Users can upload their own geo-referenced data (in point, polygon and raster format) and share it with selected users or, alternatively, the whole impressive network which is the Google Earth™ community today. Although Google Earth™ does not offer traditional GIS functionality, it can be used also to add content, such as points or lines to the existing maps, measure areas and distances, derive coordinates and ultimately load GPS data.

Second, with regard to web-based data sharing, the data is readily located on Internet servers so users do not have to download or install any data locally. However, the need to access the Internet, e.g. to capture satellite images of areas of interest is a limitation, that will impede the use of Google Earth™ for this purpose in some parts of the world.

Third, the base maps in Google Earth™ (remote sensing images, roads, administrative units, topography, etc.) are extensive and above all, frequently updated and with constantly improving quality. Though still far from suitable for all kinds of epidemiological studies, this brings new promise for the improved use of remote sensing applied to epidemiology, which has been criticized for not fulfilling its potential (Herbreteau et al., 2007).

Presently, high-resolution IKONOS images (spatial resolution of 2 m or better) cover about 20-30% of the world. The limitation here is that non-urban areas in the developing parts of the world suffer from poor coverage of high-resolution satellite imagery. This is a limitation which is a particular problem with regard to rural, local studies of many diseases, e.g. schistosomiasis and other neglected tropical diseases, but also malaria. Another limitation with acquiring high resolution imagery through Google Earth™ is that satellite image data are intended for optimized ground cover presentation, which makes it weak for analysis purposes as the time of image acquisition cannot be chosen

(Monkkonen, 2008; Kamadjeu, 2009).

The strength of the Google Earth™ is that it uses a single coordinate system, i.e. the world geodetic system (WGS-84) and that the geodata are visualized using a 3-D model rather than a projected 2-D system. This means that the user avoids having to deal with the complexity of understanding and merging of maps and layers from different projection systems. The fact that more and more people and organizations are now producing KML versions of their spatial data (as opposed to other spatial data formats that require different types of GIS software) means that it has become easy for scientists, and the general public alike, to quickly prepare mashups and explore data obtained by different research groups.

The use of Google Earth™ and Google Maps™ in the CONTRAST project, which is generally applicable to other research projects with geospatial health components, can be summarized into a few overall categories such as:
(i)   visualization;
(ii) communication;
(iii) data-exploration (e.g. the identification of spatial and temporal disease clustering);
(iv) validation;
(v)   dissemination of research result for a wider audience; and
(vi) support for decision-making.

A remaining problem, still hindering the full potential of virtual globe applications for these purposes, is how to attain proper geo-referencing of the disease-related data, not least for the vectors and the intermediate hosts carrying the infectious agents. Failure to include spatial information may eliminate potentially highly productive routes to analysis, including those not yet foreseen. But these data are frequently inadequate or absent. This remains one of the main obstacles for the direct mapping of vector-borne diseases and the exploration of relationships with the heterogeneous environment in which they exist. Other important issues, not dealt with in this paper, pertain to the ethics of displaying traceable health information in public space and concerns with

confidentiality of the data (Curtis et al., 2006). According to some, the type of visualization exemplified by the virtual globes crosses several key thresholds in communicating scientific and environmental information (Sheppard and Cizek, 2009). Thus, while the appeal of these techniques are evident, with unprecedented opportunities for public access to data and collaborative engagement over the web, there are nonetheless risks that need to be considered before applying these techniques in areas of public interest, such as planning and policy-making.

At present, Google Earth™ appears to be primarily used as a geo-browser for exploring spatially referenced data. However, its functionality can be integrated with various analytical tools for spatial analysis (e.g. GIS and open source statistical packages such as R (Bivand, 2006; Yi et al., 2008) while facilitating the sharing of spatially-referenced data between international research groups, and agencies (Wood et al., 2007). We hope this paper will stimulate further exploration of virtual globe applications for spatial epidemiologists, and that this could also introduce to the broader research community the potential of recording and making accessible spatial data in appropriate formats. We will continue to use Google Earth™ and Google Maps™ as an integral part of the ongoing CONTRAST project and we invite readers to check our website, as well as the web-sites of the many other research projects mentioned in this paper, to witness how virtual globe applications can be used for the display and sharing of data and research relevant to the management and control of vector-borne and other environmental diseases.

Acknowledgements

References

Aburto-Oropeza O, Ezcurra E, Danemann G, Valdez VÃ, Murray J, Sala E, 2008. Mangroves in the Gulf of California increase fishery yields. Proc Natl Acad Sci USA 105, 10456-10459.

Bivand R, 2006. Implementing spatial data analysis software tools in R. Geogr Anal 38, 23-40.

Blamont J, 2008. We the people: consequences of the revolution in the management of space applications. Space Policy 24, 13-21.

Boulos M, 2005. Web GIS in practice III: creating a simple interactive map of England's strategic health authorities using Google Maps API, Google Earth KML, and MSN Virtual Earth Map Control. Int J Health Geogr 4, 22.

Boulos M, Scotch M, Cheung KH, Burden D, 2008. Web GIS in practice VI: a demo playlist of geo-mashups for public health neogeographers. Int J Health Geogr 7, 38.

Brooker S, 2007. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. Trans R Soc Trop Med Hyg 101, 1-8.

Brooker S, Clements ACA, 2009. Spatial heterogeneity of parasite co-infection: determinants and geostatistical prediction at regional scales. Int J Parasitol 39, 591-597.

Brooker S, Rowlands M, Haller L, Savioli L, Bundy DAP, 2000. Towards an atlas of human helminth infection in sub-Saharan Africa: the use of geographical information systems (GIS). Parasitol Today 16, 303-307.

Bruno JF, Selig ER, 2007. Regional decline of coral cover in the Indo-Pacific: timing, extent, and subregional comparisons. PLoS ONE 2, 8.

Butler D, 2006a. Mashups mix data into global service. Nature 439, 6-7.

Butler D, 2006b. Virtual globes: the web-wide world. Nature 439, 776-778.

Conroy GC, Anemone RL, Van Regenmorter J, Addison A, 2008. Google Earth, GIS, and the great divide: a new and simple method for sharing paleontological data. J Hum Evol 55, 751-755.

Curtis AJ, Mills JW, Leitner M, 2006. Spatial confidentiality and GIS: re-engineering mortality locations from published maps about Hurricane Katrina. Int J Health Geogr 5, 44.

DHI, 2005. MIKE FLOOD User Manual. DHI Water and Environment, Denmark.

Doering A, Veletsianos G, 2007. An investigation of the use of real-time, authentic geospatial data in the K-12 classroom. J Geogr 106, 217-225.

Doumenge JP, Mott KE, Cheung C, Villenave D, Chapuis O, Perrin MF, Reaud-Thomas G, 1987. Atlas of the global distribution of schistosomiasis. World Health Organization, Parasitic Diseases Programme, Presses Universitaires de Bordeaux, Bordeaux, France.

East IJ, Hamilton S, Garner G, 2008. Identifying areas of Australia at risk of H5N1 avian influenza infection from exposure to migratory birds: a spatial analysis. Geospat Health 2, 203-213.

Ebi K, Hartman J, Chan N, Mcconnell J, Schlesinger M, Weyant J, 2005. Climate suitability for stable malaria transmission in Zimbabwe under different climate change scenarios. Climatic Change 73, 375-393.

Ezzati, M, Lopez AD, Rodgers A, Murray CJL, 2004. Comparative quantification of health risks: global and regional burden of diseases attributable to selected major risk factors. World Health Organization, Geneva, Switzerland.

Freifeld CC, Mandl KD, Reis BY, Brownstein JS, 2008. HealthMap: global Infectious disease monitoring through automated classification and visualization of Internet media reports. J Am Med Inform Assoc 15, 150-157.

Galster S, Burgess ND, Fjeldså J, Hansen LA, Rahbek C, 2007. One degree resolution databases of the distribution of 1085 species of mammals in Sub-Saharan Africa. On-line data source-Version 1.00. Zoological Museum, University of Copenhagen, Denmark.

Gao S, Mioc D, Anton F, Yi X, Coleman DJ, 2008. Online GIS services for mapping and sharing disease information. Int J Health Geogr 7, 8.

Guralnick RP, Hill AW, Lane M, 2007. Towards a collaborative, global infrastructure for biodiversity assessment. Ecol Lett 10, 663-672.

Guralnick RP, Wieczorek J, Beaman R, Hijmans RJ, 2006. BioGeomancer: automated georeferencing to map the worlds biodiversity data. PLoS Biol 4, 1908-1909.

Hales S, de Wet N, Maindonald J, Woodward A, 2002. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. Lancet 360, 830-834.

Hansen LA, Burgess ND, Fjeldså J, Rahbek C, 2007a. One degree resolution databases of the distribution of 739 species of amphibians in Sub-Saharan Africa. On-line data source-Version 1.00. Zoological Museum, University of Copenhagen, Denmark.

Hansen LA, Fjeldså J, Burgess ND, Rahbek C, 2007b. One degree resolution databases of the distribution of 1789 resident birds in Sub-Saharan Africa. On-line data source-Version 1.00. Zoological Museum, University of Copenhagen, Denmark.

Herbreteau V, Salem G, Souris M, Hugot JP, Gonzalez JP, 2007. Thirty years of use and improvement of remote sensing, applied to epidemiology: from early promises to lasting frustration. Health Place 13, 400-403.

Kamadjeu R, 2009. Tracking the polio virus down the Congo River: a case study on the use of Google Earth in public health planning and mapping. Int J Health Geogr 8, 4.

Lawson AB, Williams FL, 2001. An introductory guide to disease mapping. John Wiley & Sons, Inc, New York, NY, USA.

Lefer TB, Anderson MR, Fornari A, Lambert A, Fletcher J, Baquero M, 2008. Using Google Earth as an innovative tool for community mapping. Public health reports (Washington, DC: 1974) 123, 474-480.

Listorti JA, Doumani FM, 2001. Environmental Health. Bridging the Gaps, 2001. World Bank. World Bank - Discussion Papers.

Lozano-Fuentes S, Elizondo-Quiroga D, Farfan-Ale JA, Loroño-Pino MA, Garcia-Rejon J, Gomez-Carro S, Lira-Zumbardo V, Najera-Vazquez R, Fernandez-Salas I, Calderon-Martinez J, Dominguez-Galera M, Mis-Avila P, Morris N, Coleman M, Moore CG, Beaty BJ, Eisen L, 2008. Use of Google Earth™ to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. Bull World Health Organ 86, 718-725.

Luedeling E, Buerkert A, 2008. Typology of oases in northern Oman based on Landsat and SRTM imagery and geological survey data. Remote Sens Environ 112, 1181-1195.

Malone JB, 2005. Biology-based mapping of vector-borne parasites by geographic information systems and remote

sensing. Parassitologia 47, 27-50.

Mikula S, Trotts I, Stone JM, Jones EG, 2007. Internet-enabled high-resolution brain mapping and virtual microscopy. Neuroimage 35, 9-15.

Monkkonen P, 2008. Using online satellite imagery as a research tool mapping changing patterns of urbanization in Mexico. J Plan Educ Res 28, 225-236.

Nethery E, Teschke K, Brauer M, 2008. Predicting personal exposure of pregnant women to traffic-related air pollutants. Sci Tot Environ 395, 11-22.

Patterson TC, 2007. Google Earth as a (not just) geography education tool. J Geogr 106, 145-152.

Pearce JM, Johnson SJ, Grant GB, 2007. 3D-mapping optimization of embodied energy of transportation. Resour Conserv Recy 51, 435-453.

Prüss-Üstün A, Corvalan C, 2007. How much disease burden can be prevented by environmental interventions? Epidemiology 18, 167-178.

Rasmussen JB, Hansen LA, Burgess ND, Fjeldså J, Rahbek C, 2007. One degree resolution databases of the distribution of 467 species of snakes in Sub-Saharan Africa. On-line data source-Version 1.00. Zoological Museum, University of Copenhagen, Denmark.

Rinaldi L, Musella V, Biggeri A, Cringoli G, 2006. New insights into the application of geographical information systems and remote sensing in veterinary parasitology. Geospat Health 1, 33-47.

Sheppard SRJ, Cizek P, 2009. The ethics of Google Earth: crossing thresholds from spatial data to landscape visualisation. J Environ Manage, doi: 10.1016/j.jenvman.2007.09.012.

Sundvall E, Nystrom M, Forss M, Chen R, Petersson H, Ahlfeldt H, 2007. Graphical overview and navigation of electronic health records in a prototyping environment using Google Earth and openEHR archetypes. Medinfo 12, 2-7.

Tanser FC, Sharp B, le Sueur D, 2003. Potential effect of climate change on malaria transmission in Africa. Lancet 362, 1792-1798.

Thenkabail PS, Biradar CM, Noojipady P, Cai XL, Dheeravath V, Li YJ, Velpuri M, Gumma M, Pandey S, 2007. Sub-pixel area calculation methods for estimating irrigated areas. Sensors 7, 2519-2538.

Weibel D, Schelling E, Bonfoh B, Utzinger J, Hattendorf J, Abdoulaye M, Madjiade T, Zinsstag J, 2008. Demographic and health surveillance of mobile pastoralists in Chad: integration of biometric fingerprint identification into a geographical information system. Geospat Health 3, 113-124.

Wood J, Dykes J, Slingsby A, Clarke K, 2007. Interactive visual exploration of a large spatio-temporal dataset: reflections on a geovisualization mashup. IEEE Transactions on Visualization and Computer Graphics.

Woolhouse MEJ, Dye C, Etard JF, Smith T, Charlwood JD, Garnett GP, Hagan P, Hii JLK, Ndhlovu PD, Quinnell RJ, Watts CH, Chandiwana SK, Anderson RM, 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc Natl Acad Sci USA 94, 338-342.

Yang GJ, Vounatsou P, Xiao-Nong Z, Utzinger J, Tanner M, 2005. A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in China. Acta Trop 96, 117-129.

Yi Q, Hoskins RE, Hillringhouse EA, Hillringhouse EA, Sorensen SS, Sorensen SS, Oberle MW, Oberle MW, Fuller SS, Fuller SS, Wallace JC, Wallace JC, 2008. Integrating open-source technologies to build low-cost information systems for improved access to public health data. Int J Health Geogr 7, 29.

Yuan X, Liu Y, Yuen D, Chen B, Pergler T, Shi Y, 2008. An efficient system for creating synthetic InSAR images from simulations. Pure Appl Geophys 165, 671-691.