A Digital Observatory for Protected Areas - DOPA, a GEO-BON contribution to the monitoring of African biodiversity

DUBOIS Grégoire, CLERICI Marco, PEEDELL Stephen, MAYAUX Philippe, BARTHOLOME Etienne, GREGOIRE Jean-Marie

Joint Research Centre - European Commission
Global Environment Monitoring Unit
Via Fermi 2749, TP 440
I-21027 Ispra (VA), ITALY

Tel : +39 0332 786360
Fax : +39 0332-789960
Email: gregoire.dubois@jrc.ec.europa.eu

Biodiversity data are particularly difficult to manage at the global scale as the volume of information is huge and at the same time very fragmented. Biodiversity is one of the 9 key themes by which the activities of the intergovernmental Group on Earth Observations (GEO) are organised. GEO-BON, GEO's Biodiversity Observation Network, is therefore aiming to integrate biodiversity data by bringing together the large variety of sensors, databases and systems. The purpose of this paper is to present DOPA, a possible contribution to GEO-BON. DOPA is a Digital Observatory for Protected Areas currently developed at the Joint Research Centre (JRC) of the European Commission in collaboration with other international organizations, including GBIF, UNEP-WCMC and BirdLife International. DOPA is conceived as a set of distributed databases and open, interoperable web services for the assessment of the state and pressure of protected areas. Already benefiting from existing data for environmental assessments, DOPA is being extended to become an operational monitoring and modelling service capable of capturing the dynamics of spatio-temporal changes in habitats in and around protected areas. To achieve these goals, there is a need for collecting and processing remote sensing data automatically. Such a collecting and processing system has been developed by the JRC in the frame of its support to the AMESD (African Monitoring of Environment for Sustainable Development) project. Called, eStation, this component of AMESD is the software dealing with the reception, processing, analysis and dissemination of key environmental parameters. Remotely sensed data, derived from the Earth Observation measurement platforms such as SPOT/VGT, SEVIRI/MSG and TERRA-AQUA/MODIS, are processed to compute environmental trends and detect anomalies. In contrast to the other information processed by DOPA, the eStation receives the data mainly through the EumetCAST broadcasting system, providing DOPA with some means to be partly independent from an internet based infrastructure. Ultimately, DOPA could become a significant contribution to the local implementation of operational environmental information services that can help to improve environmental surveillance, management and policy making. Further developments of DOPA will focus on providing means to submit ground based information back into the system for validations.

I. Introduction

Biodiversity data are particularly difficult to manage at the global scale as the volume of information is huge and at the same time very fragmented. Biodiversity is one of the 9 key themes by which the activities of the intergovernmental Group on Earth Observations (GEO) are organised. GEO-BON, GEO's Biodiversity Observation Network, is therefore aiming to integrate biodiversity data by bringing together the large variety of sensors, databases and systems. DOPA, a Digital Observatory for Protected Areas in development at the Joint Research Centre (JRC) of the European Commission in collaboration with other international organizations, including GBIF, UNEP-WCMC and BirdLife International, is an example of such a system. Conceived as a set of distributed databases and open, interoperable web services, DOPA is designed to provide decision makers and researchers with means to assess and monitor the state and pressure of protected areas at the global scale. Data on protected areas, species distributions, socio-economic indicators are therefore combined with remote sensing information in order to generate the necessary environmental indicators, maps and alerts.

As for most environmental information systems, the main challenge in setting up the DOPA comes from the handling a large variety of data types and sources. Hartley et al. (2007) provide a description of a first prototype of DOPA focusing on Africa - the African Protected Areas Assessment Tool[1] (APAAT). APAAT was designed to provide decision makers with a tool to assess the state of protected areas in Africa and to prioritize them according to biodiversity values and threats, so as to support decision making and fund allocation processes. In order to evolve towards an operational environmental monitoring and forecasting service, matching the needs from a broader community of end-users, the system needed to be able to capture more regularly the spatio-temporal environmental changes, implying that new means had to be put in place to ensure the automatic update of the information presented.

In particular, the development of tools for biodiversity monitoring and forecasting requires some significant changes in the way data are collected and processed. Figure 1 summarises a typical process in which data are collected from various sources, processed for further integration with other data before the publication of new information.

II. The GEOSS philosophy: One for All, All for One

This relatively simple schema of data exchange can easily become extremely complex in the absence of syntactic and semantic interoperability of the data and the systems serving them. This is an issue that is well known and addressed in a number of international initiatives, the largest one being GEOSS, the Global Earth Observation System of Systems[2]. The Biodiversity Observation Network of the intergovernmental Group on Earth Observations (GEO-BON) is

[1] http://bioval.jrc.ec.europa.eu/PA

[2] http://www.earthobservations.org/

leading the integration process of biodiversity data by bringing together the large variety of sensors, databases and systems. GEO-BON's objectives to improve the collection, management, sharing, and analysis of data on the status and trends of the world's biodiversity rely heavily on information and communication technologies. This generates numerous requirements, from the use of common standards, to the development of means to discover and chain services, to the development of portals and client applications to allow end-users to access these services. Hence, beyond the simple sharing and exchange of information, GEO-BON conceives open-ended systems of interoperable computer models and databases communicating via Service Oriented Architectures (SOA).

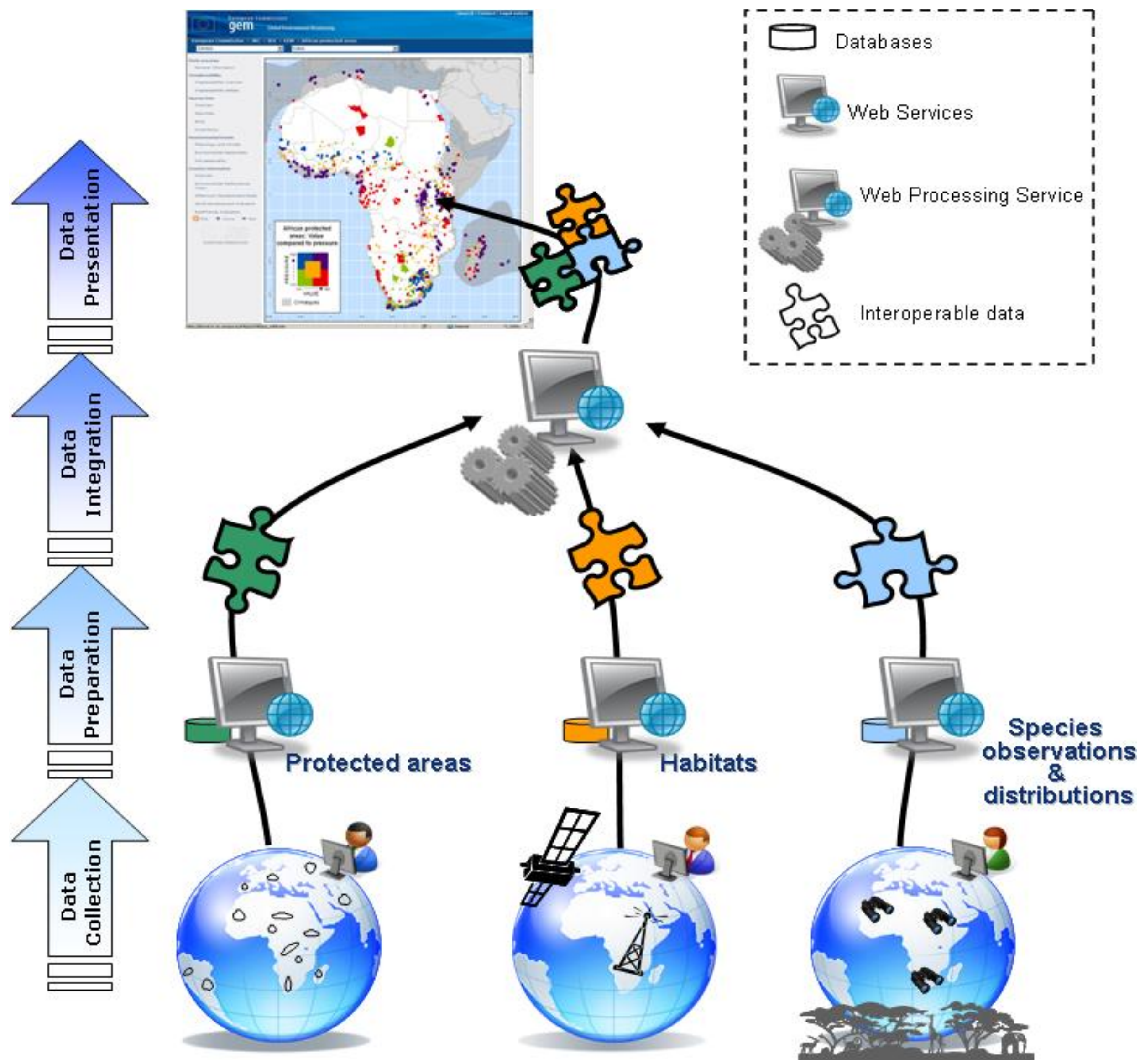

Figure 1. From ground based and remote sensing observations to environmental indicators: data need to be collected, prepared to allow their combined use, and integrated for the preparation of indicators.

The development of DOPA (Dubois et al., 2009) takes into full account the key recommendations of these international initiatives to promote the effective exchange, processing and modelling of spatial data, ensuring that the data used are

1) managed as close as possible to its source;
2) collected once and documented to allow their use for many purposes;
3) easily retrievable and accessible by others;
4) interoperable at the syntactic and semantic level to allow their combination for multiple purposes;
5) scalable, when applicable, to be meaningful at different levels of geographic detail;
6) shared and, where possible, processed through common, free open-source software tools.

Developing such a working environment requires significant coordination between data providers and users to reduce as much as possible the redundancy in the collection, storage and processing of the information. A well-coordinated management in the development of environmental data repositories, catalogues, registries and associated web services are at least as important as the growth in accessible and updated information.

III. From assessing to monitoring, the need for automating data collection and processing

Because DOPA is targeting mainly continental and global biodiversity issues, spatio-temporal changes in habitats and anthropogenic pressure on protected areas and other areas of ecological importance need to be captured as much as possible automatically and on a large scale. This requirement is primarily addressed via the widespread use of remote sensing data, which is a key component of the system.

DOPA will partly rely on the eStation, a collecting and processing system developed by the JRC in the frame of its support to the AMESD (African Monitoring of Environment for Sustainable Development) initiative. AMESD is a continuation of the PUMA project (see Counet and de Backer, 2005). The eStation is the component of AMESD dealing with the reception, processing, analysis and dissemination of key environmental parameters. Remotely sensed data, derived from the measurements done by the SPOT/VGT, SEVIRI/MSG and TERRA-AQUA/MODIS EO systems, are processed to compute environmental trends and detect anomalies (Figure 2).

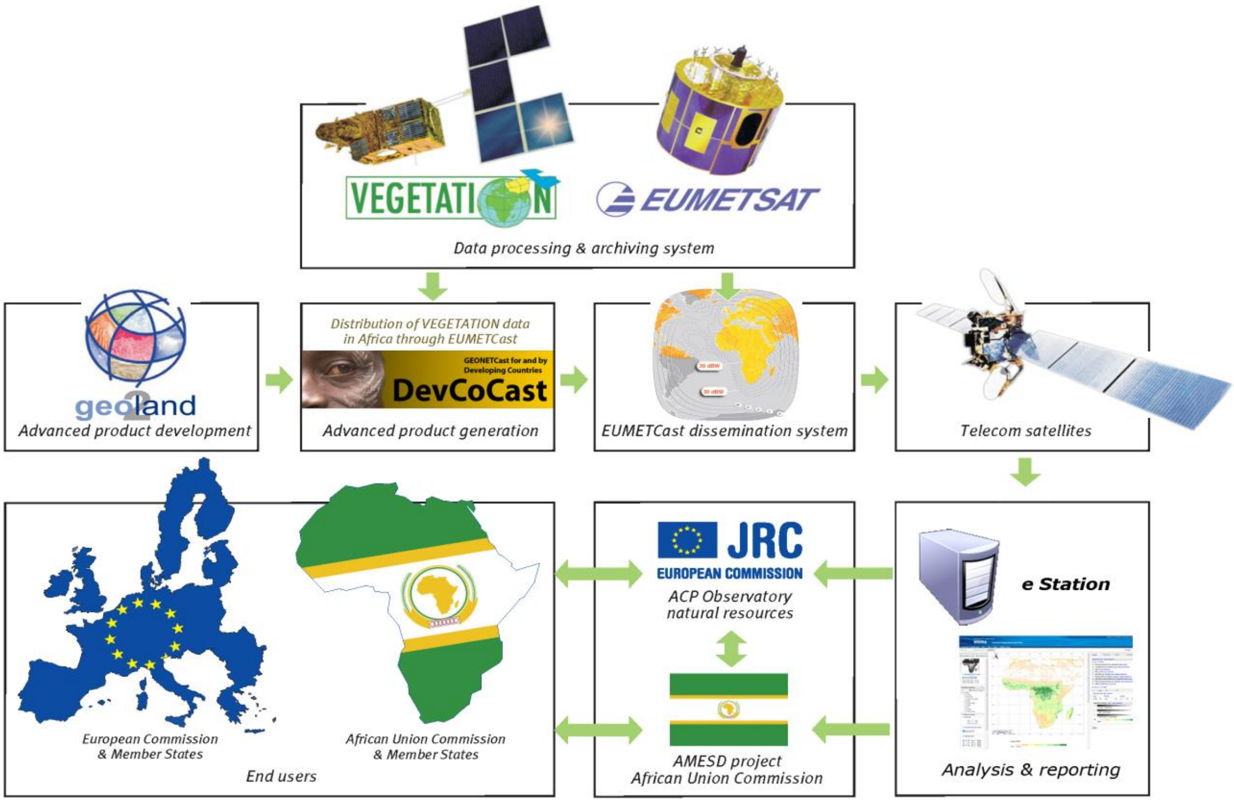

Figure 2. Data flow in the framework of the African Monitoring of Environment for Sustainable Development (AMESD) project. The e-Station is the component used in the DOPA for the processing of the remote sensing data

The eStation is a modular system providing three main functionalities with a high degree of automation. The first component, PS (Processing Station), is the data management and processing component. It automatically retrieves Earth Observation data and computes ad-hoc thematic products and environmental indicators, according to the end-user needs. All processing steps are easily configurable in order to allow the user to modify the generated environmental indicators and to implement new ones. In the DOPA, environmental anomalies are detected in protected areas by contrasting every 10 days environmental data against historical records. The parameters analysed include and rainfall, active fires, small water body presence, a water index (NDWI), and a vegetation index (NDVI). These anomalies can be characterised by their strength, their duration and their deviation from their expected occurrence in time, something typical of seasonal changes. See Hartley et al. (2007) and Carrara et al. (2008) for more details on this alert system.

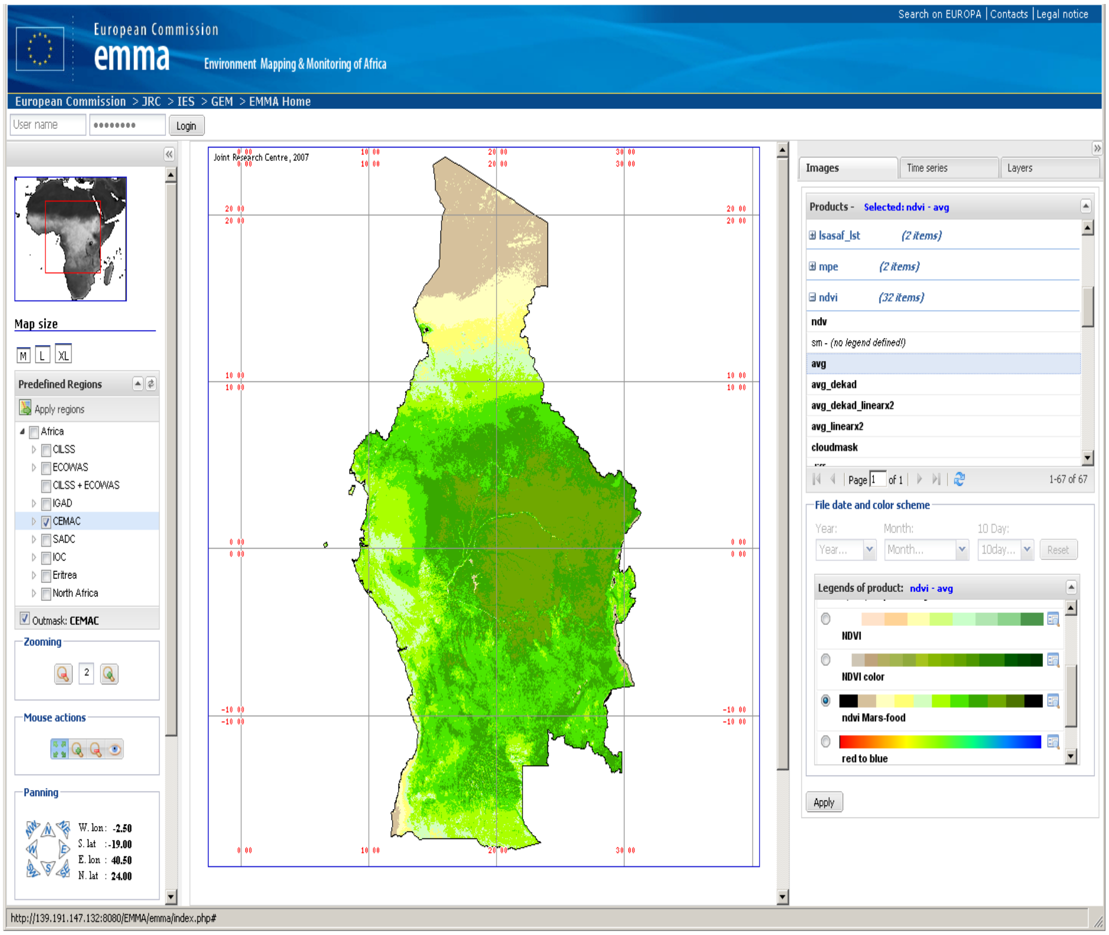

Figure 3. Snapshot of EMMA (Environmental Mapping and Monitoring of Africa), the eStation's graphical user interface developed for facilitating the exploration of the variables processed by the eStation

For what concerns the monitoring of fire activity and burned areas in sub-Saharan protected areas, the information is derived from the MODIS fire products available through the NASA funded Fire Information for Resource Management System (FIRMS). Information on fire occurrence is derived four times a day by the MODIS sensors onboard the TERRA and AQUA satellites, at a 1km spatial resolution, while the burned area information has a 500x500 m resolution. From all the datasets, active fires and burned areas, we derive statistics with a 10-day time step (decade) for each protected area in sub-Saharan Africa over a time series which spans from year 2000 to present (Grégoire and Simonetti, 2010). The fire-related data and information, updated continuously as soon as new satellite data become available, are provided in support to park managers as well as to researchers exploring land-use and climatic changes. Data and fire statistics are accessible through the JRC-funded system[3].

[3] http://bioval.jrc.ec.europa.eu/PA/

Anthropogenic threats are also currently mainly captured through the processing of remote sensing information used to assess agricultural pressure, population growth (combined also with statistical information) and habitat fragmentation around the protected areas.

IV. Ecological forecasting with DOPA's Web Processing Service: eHabitat

In striving to improve the predictive capabilities of ecological forecasting we face three basic choices – develop new models, improve existing ones or increase the connectivity of models so they can work together. The latter approach of chaining different interoperable models is of particular interest, as technical developments have made it increasingly viable to combine models that can answer more questions than the individual models alone, allowing users to address complex questions, often of a multi-disciplinary nature. This concept of a Model Web (Geller and Turner, 2007) encourages the setting up of a dynamic network of interoperating models, communicating with each other using standardized web services.

DOPA relies on a model called eHabitat (Figure 4) for performing ecological modelling, which computes the likelihood of finding ecosystems with equal properties and therefore is the basis for defining an index of *Habitat Irreplaceability* (see e.g. Rotenberry, Preston and Knick, 2006) (Figure 4).

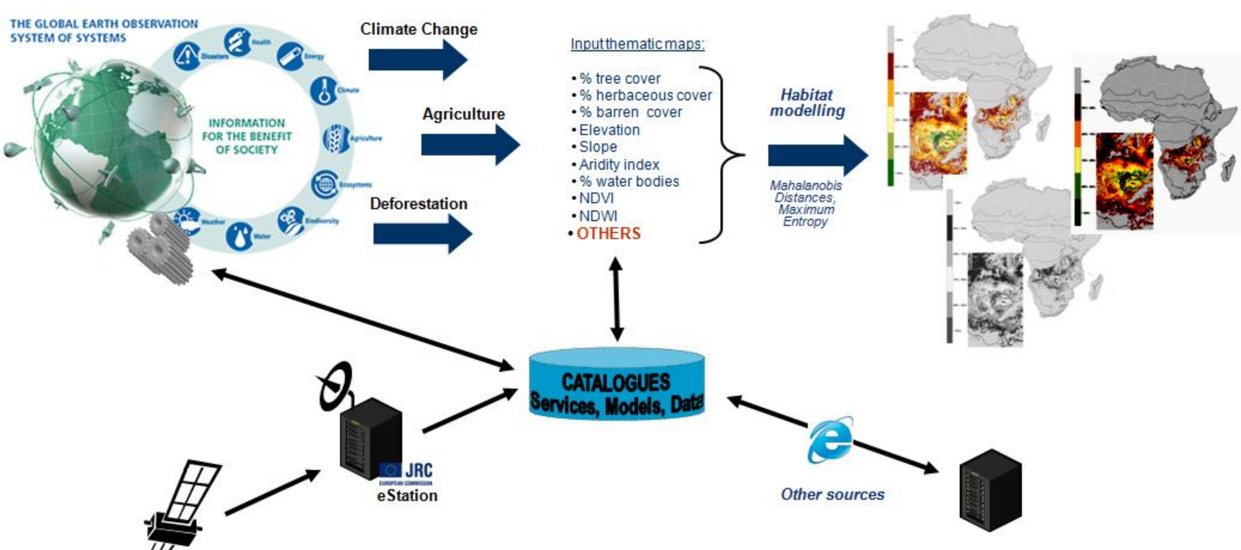

Figure 4. Conceptual design of eHabitat, a Web Processing Service (WPS) designed to compute the likelihood to find similar ecosystems using thematic maps selected by the end-users. eHabitat is conceived to view to potentially integrate a more complex processing chain as needed for forecasting, for example, the potential impact of climate change, deforestation scenarios or changes in agricultural policies on these habitats.

In previous implementations, eHabitat has been a stand-alone model, developed in proprietary desktop software. In DOPA, eHabitat is conceived as a Web Processing Service (WPS). By developing eHabitat according to Model Web principles, end-users can not only define the thematic layers for input to the model from various sources using standards-based catalogues, but also chain the modelling service with others, typically those used for climate change forecasting. Because eHabitat is scalable and can integrate data ranging from remote sensing data to socio-economical indicators, the service offers a huge potential for multi-disciplinary modelling.

V. DOPA's potential architecture: a hybrid system using the internet and broadcasting services

Both the DOPA and the eStation are completely based on OpenSource software and on an architecture of open standards and specifications. It is therefore a flexible and cost-effective platform for the implementation of information services needed for environmental and natural resources policy orientation, management and assessment, in different thematic areas, like the monitoring of natural vegetation, agricultural production, water resources, coastal and marine regions. In contrast to the other information processed by DOPA that is exchanged over the Internet, the eStation receives the data mainly through EUMETCast (EUM/OPS/DOC/06/0118), a satellite-based data dissemination system, providing DOPA with some means to be partly independent from an internet-based infrastructure.

The eStation is being distributed to the National Centres in all sub-Saharan African countries involved in the AMESD project (African Monitoring of Environment for Sustainable Development), implemented by the African Union Commission in partnership with the ACP Secretariat, IGAD, IOC, SADC, CEMAC and ECOWAS, five Regional Economic Groups that are supported by the European Development Fund.

EUMETCast is a regional component, covering Europe, Africa and part of the Americas, and is integrated with other regional dissemination systems under the GEOSS GEONETCast[3] umbrella. GEONETCast is a near real-time, global network of satellite-based data dissemination systems designed to distribute space-based, air-borne and in-situ data, metadata and products to diverse communities. Ultimately, benefiting from a hybrid architecture, using the Internet and broadcasting services like GEONETCast, DOPA could become a significant contribution in the future to the local implementation of operational environmental information services that can help to improve environmental surveillance, management and policy making. This presents many challenges, not least the effort required to store, maintain and efficiently access the vast archives of historical records in order to calculate trends when assessing climate change impact for example.

VI. The data uncertainties paradigm in the Model Web context: the increased freedom calls for more control

The GEO Model Web concept envisages the integration of complex information and processing resources, to construct sophisticated models, chained together as web services. This offers exciting opportunities for model development in a more loosely-coupled, component-oriented manner, encouraging sharing, re-use and easy access. However, when combining services of limited, or unknown, quality it is necessary to account for uncertainty if the outputs of the Model Web are to be used for rational decision making. This is a fundamental challenge for DOPA which, whilst encouraging a multi-scale, cross-disciplinary approach to biodiversity, also potentially generates additional problems because of the broad range of different types of data and their associated, usually undocumented, levels of uncertainties. These uncertainties can have serious consequences if not properly considered. Increased access of information and tools to users with very different backgrounds and aims, can potentially also lead to conflicting results. As certain inter-disciplinary issues are resolved, so the risk of misuse of data and analytical models also increases dramatically, as is the risk to see errors propagate between models.

Chained to other services to perform ecological forecasting, eHabitat would be an additional component of the Model Web further propagating uncertainties from a potentially long chain of model services. To address this issue, which can impact on any attempts to perform distributed modelling of heterogeneous data, we are engaged in the recently launched UncertWeb[4] project, which aims to develop mechanisms, standards and tools to enable uncertainty management in an interoperable model web context

In addition to the issue of uncertainty, we should not forget fundamental challenges of data management - with the steady growth of environmental web services, distributed databases and more open data, the risk of data redundancy also increases. In our case, it should be ensured, for example, that borders of protected areas are maintained and updated by a single organisation only, thus adhering to the earlier stated principle of GEOSS, that data should be collected once and documented to allow their use for many purposes

With the increasing freedom of setting up chained services grows the need for more clearly identifiable, reference spatial data infrastructures that would partly alleviate the misuse of data. By relying mainly on reference data centres, managed by organisations with clear and well established mandates, when generating products that are derived from disparate resources, DOPA will stimulate a culture of "quality control" for robust science in the field of biodiversity. This control will apply and throughout the entire data flow: from the harvesting of the data to their combination with other sources by different experts to generate new information products and services.

[4] http://www.uncertweb.org

Acknowledgements

This work is partly funded by the European Commission, under the 7th Framework Programme, by the EuroGEOSS project funded by DG RTD and by the UncertWEB project funded by DG INFSO. The views expressed herein are those of the authors and are not necessarily those of the European Commission. More information about the DOPA can be found on the Internet, see http://dopa.jrc.ec.europa.eu/.

References

P. Carrara, G. Bordogna, M. Boschetti, P.A. Brivio, D. Stroppiana and A. Nelson (2008) A flexible multi-source spatial data fusion system for environmental status assessment at continental scale. *International Journal of Geographical Information Science*, 22(7): 781-799.

P. Counet and H. De Backer (2005). The PUMA Project, an exemplar African story'. *eCourier* no 5, November 2005, p.3.

G. Dubois, A. Hartley, A. Nelson, P. Mayaux, and J.-M. Grégoire (2009). Towards an inter-operable web service for the monitoring of African protected areas. In: *Proceedings of the 33rd International Symposium on Remote Sensing of Environment*, 33rd ISRSE, Stresa, Italy.

EUM/OPS/DOC/06/0118: EUMETCast – EUMETSAT's Broadcast System for Environmental Data, v5b, February 2010.

G.N. Geller and W. Turner (2007). The model web: a concept for ecological forecasting, In: *Geo-science and Remote Sensing Symposium, IGARSS 2007*, IEEE International, pp. 2469 – 2472.

J.-M. Grégoire and D. Simonetti (2010). Interannual changes of fire activity in the protected areas of the SUN network and other parks and reserves of the West and Central Africa region derived from MODIS observations. *Remote Sensing*, 2: 446-463.

A. Hartley, A. Nelson, P. Mayaux, and J.-M. Grégoire (2007). The Assessment of African Protected Areas, EUR 22780 EN, Office for Official Publications of the European Communities, Luxembourg.

J.T. Rotenberry, K. L. Preston and S. T. Knick (2006). GIS-based niche modeling for mapping species' habitat. *Ecology*, 87, 1458-64.

Biographies

 Grégoire Dubois is a Senior Scientific Officer at the Joint Research Centre (JRC) of the European Commission. After more than 10 years of work in the field of radioecology at the JRC, he joined in 2008 the Global Environment Monitoring Unit where he coordinates the biodiversity related activities, in particular the development of the Digital Observatory for Protected Areas (DOPA). His qualifications include a BSc in Zoology, an MSc in Radiobiology and a PhD in Applied Geostatistics.

 Marco Clerici received the B.S. degree in electronic engineering, specializing in microelectronics and instrumentation, from the University Politecnico di Milano, Milan, Italy, in 1993. He has been with EUMETSAT from 2000 to 2004 providing software development support for the MSG/SEVIRI vicarious calibration and for the development of radiative transfer models. He is currently with the Institute for Environment and Sustainability (IES) in Ispra. His work focuses on the implementation of operational algorithms and systems for the land and vegetation characterization from remote sensing, applied to the global and continental scale, with a focus on applications for the African continent.

 Stephen Peedell is a Senior Scientific Officer at the Joint Research Centre (JRC) of the European Commission. His qualifications include a Masters degree in Natural Resource Management and, since joining JRC in 1996, he has been involved in many European initiatives combining geospatial technology and environmental decision making. In particular he has worked on the development of the INSPIRE Directive and designing and building geographic information systems supporting Natura2000 and the Water Framework Directive. In 2009 he joined the Global Environment Monitoring Unit where he is working on developing spatial databases, catalogues and interoperable web services for geoinformation.

 Philippe Mayaux is a Senior Scientific Officer at the Joint Research Centre (JRC) of the European Commission. He has developed for the last 20 years skills in land-cover mapping in Africa from high and coarse resolution satellite data in the frame of the TREES-2 and GLC2000 projects. His current research interests include the development of information systems, based on a combination of in-situ and EO data for monitoring forest and biodiversity resources in ACP countries and the assessment of the implications of forest cover changes on the global carbon budget and the biodiversity trends. Philippe Mayaux has a MSc in Forestry, a MSc in Urban planning and a PhD in Forestry.

 Jean-Marie Grégoire is a Senior Scientific Officer at the Joint Research Centre (JRC) of the European Commission and Deputy Unit Head within the Global Environment Monitoring Unit. Jean-Marie Grégoire has a degree in Geomorphology applied to environmental management and a PhD on the use of remote sensing for modelling water and energy transfer in the soil-plant-atmosphere system. His current research interests are in developing remote sensing approaches to environmental monitoring in tropical regions of the globe, with particular emphasis on vegetation fires and fire ecology in Africa.

 Etienne Bartholomé is a Senior Scientific Officer at the Joint Research Centre (JRC) of the European Commission. He joined in 1984 the group that would become later the Global Environment Monitoring Unit of the JRC where he coordinates R&D activities related to the use of low resolution EO data for environmental monitoring in Africa in the GMES framework. His qualifications include a MSc in Physical Geography and a PhD in Geography related to crop production monitoring in Sahelian countries using EO data.