Processing on Demand of EO value added vegetation products

Erwin Goor¹, Dirk Daems¹, Lieve Van den Heuvel¹, Steven Smolders²

Abstract

VITO is responsible for the production, archiving and dissemination of products from the SPOT-VEGETATION mission and also serves many derived vegetation products from this sensor, complementary sensors and higher resolution sensors. To facilitate the dissemination, VITO deploys a web based data access portal and standardized web services which allow authorized users and the general public to discover, view, download and order available products, possibly customized towards their needs. The next step is to deploy processing on demand services via a web based portal and WPS interfaces. This will allow the user to generate tailored data and retrieve information based on the extensive data archive at VITO on-the-fly. This paper describes the current status and future steps with regard to processing on demand services at VITO.

1. VITO data access infrastructure

The remote sensing department of VITO, the Flemish Institute for Technological Research, is hosting the ground segment for the SPOT-VEGETATION sensors (http://www.spot-vegetation.com/) which allow a daily monitoring of terrestrial vegetation cover at a 1km spatial resolution. Currently VITO is also developing the user segment for the PROBA-V mission, which will be launched end of 2012. This will guarantee the continuity of data from SPOT-VEGETATION and offers additionally products at 1/3km spatial resolution.

Based on these data and data from complementary sensors (e.g. from MODIS, MERIS, AATSR, AVHRR), VITO produces several derived Earth Observation (EO) products, e.g. bio-geophysical parameters or higher level vegetation indices. In addition, VITO processes high resolution and hyper-spectral data from airborne sensors (e.g. from the APEX instrument). These EO sensors allow the detection and mapping of short and long term vegetation changes and thus contribute to the monitoring of the status of vegetation on a global and local scale. The VITO ground segment is responsible for the processing, archiving and dissemination of these sensor data and derived products to the users.

To facilitate the dissemination, VITO currently hosts a number of web based portals that allow authorized users and the general public to discover and order available products, possibly customized towards their needs. Early 2012, VITO will deploy one centralized data access portal, which provides one unique access point to all sensor- and derived data offered by VITO. Users will be able to discover products from the catalogues and registered users can order and download these products. A screenshot of the catalogue interface is shown in figure 1.

¹ VITO nv, Flemisch Institute for Technological Research, Boeretang 200, 2400 Mol, Belgium, email: erwin.goor@vito.be, Internet: www.vito.be.
² GIM nv, Researchpark Haasrode, Interleuvenlaan 5, 3001 Leuven, Belgium, Internet: steven.smolders@gim.be.
When designing this system, VITO employs a Service Oriented Architecture implemented using web services based on open standards, to ensure interoperability in the frame of INSPIRE and GMES (Global Monitoring for Environment and Security). Currently VITO provides the following interfaces for standardized machine-to-machine communication:

- **Discovery services**: These services are accessible via the OGC 07-045 CSW ISO Application Profile, which allows integration with INSPIRE compliant clients and applications. VITO serves e.g. bio-geophysical products in the frame of the geoland2 project ([http://www.geoland2.eu](http://www.geoland2.eu)). To serve the EO-data from the SPOT-VEGETATION and PROBA-V sensors, VITO provides additionally discovery services via the OGC 07-045 CSW ebRIM Application Profile, to allow integration within GSCDA (GMES Space Component Data Access) from the European Space Agency.

- **View and download services**: WMS viewing services are provided. For download FTP(s) is used, as well as HTTP and WCS. VITO also uses Eumetcast services to broadcast specific data over satellite communication to users in Africa and Latin America.

- **Ordering services**: The available ordering interface is compliant with OGC 06-141, the best practices document which describes a profile to order Earth Observation data products. This interface allows users to order existing products from the catalogue or order products which will become available in the future (subscription).
2. Processing on demand

2.1 The needs

VITO hosts a time series of nearly 15 years of SPOT-VEGETATION data and derived vegetation parameters. This is an extremely valuable data source for research on climate change, deforestation, bio-diversity, etc. Users can access these data as explained in the previous paragraph, but in addition, there is a need to offer more user-oriented services. These services allow on-demand production of customized products meeting individual user requirements as opposed to the standard products that are typically available nowadays. Therefore VITO is building an image processing and information extraction system which provides the user with a means to:

- **Analyse time-series of data** in a convenient way and extraction of agricultural parameters. Global earth observation system (GEO) themes as climate (IPCCC), biodiversity (EAA), agriculture (FAO, CAP) and ecosystems (UNEP) require long time series and integration of other data sources to analyze changes and trends to better predict the future and adapt their policies. A huge amount of data and algorithms are required both from an EO and modeling perspective.

- **Prototype algorithms** by including new components into EO-processing chains with immediate access to the data; this provides a more efficient way to design EO-applications from the research phase into an operational service. The results can be published automatically using state-of-the-art GIS techniques for immediate use by the user.

- **Design and deploy multi-mission applications**, invoking data at different archiving centers as efficiently as possible, by connecting to infrastructures of other providers via open standards.

To implement the two latter use cases (‘prototype algorithms’ and ‘design multi-mission applications’), VITO is involved in the ESA funded SSE-Grid project to design a platform for EO-specialists which allows them to design EO-applications as workflows, involving processing services and data from VITO, powerful Open Source EO processing libraries from third-parties, as well as algorithms and data from the user themselves. Hence we target a research platform to experienced users with near real-time access to the extensive data archive at VITO.

The next paragraphs explain how VITO implements the first use case by establishing a processing on demand environment, which offers VITO the possibility to easily implement value adding processing services that target a user’s specific needs. This way, non-expert users (i.e. non EO-specialists) who don’t have the knowledge or infrastructure to process the data provided by VITO are able to process the data into user-tailored information. In line with VITO’s existing architecture, the system is implemented via web service interfaces based on open standards, namely the OGC Web Processing Service (WPS) specification.

2.2 VITO Processing on demand prototype

A system is implemented and deployed that employs 52°North WPS service as WPS framework and a geographically enabled content management system as web client application. The 52°North WPS service is used as WPS enabler for both open source libraries (e.g. GDAL) and a set of VITO legacy EO processing libraries and database applications.

The WPS Client application, which allows the easy instantiation of WPS Service Clients, is built based on an enhanced geographically enabled Content Management System that in itself is built on JBOSS SEAM and Richfaces which employs recent Java Standards like Java Server Faces, Java Persis-
tence API, JAXB and JAXWS. A Content Management System is a system that facilitates the creation of and publication of “documents” and provides facilities for document presentation, document discovery, version control, access control using users, groups and document permissions and monitoring of statistics. The CMS system that is used allows the definition of custom document types, using a model driven approach. This permits that new document types can be modeled and on this model, both persistence and graphical user interfaces are generated on the fly.

For the VITO Processing On Demand WPS Client Framework, the most important document types are:

- Process Descriptions: these documents contain all metadata required to be able to call the process and interact with the user.
- Process Instances: process runs that contain information on the input and output values, the start and end time, the status and the user who launched the process.
- Map Configurations: documents that contain all information that is required to represent the state of a Web Map Instance with the end user functionality it should contain, the coordinate Reference System and map extent, the services and layers with their specific parameters. Each WPS Process description has associations to two map configurations: one for capturing input parameters and the other for displaying the output data of a process.

Currently a number of prototype WPS services and associated WPS clients are being deployed that allow the user to visualize e.g. aggregated vegetation indices as Normalized Difference Vegetation Index (NDVI), Fraction of Absorbed Photo synthetically Active Radiation (FAPAR), ... as time series graphs and maps. The user is able to work with predefined region sets whereby he can select regions on a map or from a set of hierarchically linked dropdown lists or by drawing a Region Of Interest. In addition the user is able to select a number of other parameters as the vegetation variable to present, the land cover class, the begin time and end time and the periodicity of the measurements. Figure 2 shows a screenshot of an input panel.

![Input panel for a prototype processing on demand service](image_url)
When this set of parameters is passed to the Web Processing Service that will use the Region Of Interest (if present) to perform a spatial overlay with the predefined regions and then use the identifiers of those regions together with the other defined parameters to perform a database query and subsequently generate a time series graph. As all data is pre-calculated, processing happens very fast and the process has been implemented in a synchronous manner. Figure 3 shows the output of a processing run, where the user gets a time series graph and a map on which the chosen regions are highlighted.

Users who want to be able to derive the same vegetation indices for their own regions are given the possibility to upload their region data in the form of shape files. These shape files are then transformed into a raster data file, which is subsequently used as input for the process that calculates the average vegetation indices for the custom regions; data is provided in data files and in addition this data is published to a Sensor Observation Service (SOS) for interactive visualization within the geographically enabled WPS client. Since this processing and data publication can be rather time intensive, this process is implemented in an asynchronous manner with the WPS client regularly polling the WPS service for status updates.

Figure 4 shows the architecture of the system which is currently deployed at VITO. The WPS client invokes a WPS request, based on the input provided by the user in the web-based client. The WPS server handles this WPS request by invoking the appropriate module at the processing backend. The module can be written in C, C++, JAVA or a script language and can be deployed on a Linux or Windows server. The response is formatted as a WPS response message and returned to the WPS client. For asynchronous requests (which can take several hours/days), the user is notified by e-mail when the response is available.
The e-mail contains a URL which allows the user to immediately access the output of the on-demand service in a web browser. When the output consists of raster- or vector data, the data can be published with mapserver to allow a user to view or download the data in the corresponding map client.

2.3 Outline to the future

In the near future VITO will gradually deploy more and more services based on this model; more processing algorithms will be included to provide additional services and more data will be integrated, including data from the future PROBA-V mission in 2012.

However the underlying ICT infrastructure will also evolve to offer a richer set of features to the users. E.g. VITO will soon provide more dynamic graphs to the user; today a time series graph is presented to the user as a static image and as a CSV-file for further offline analysis. In the next months the graphs will be presented in a more dynamic way, which will allow a user to perform operations on the graph in the web browser, without invoking another WPS-request. A user will be able to zoom and pan on the graph, (un-)hide specific lines or customize the formatting of the graph.

It is also important to integrate these processing on demand services within the current data access web portal. Users will be able to discover products from the VITO catalogues and use these products as input to invoke further on-demand processing. Or the results of an on-demand processing can be ingested automatically into the product catalogue. This gives the user the possibility to use the complete VITO portfolio to access these products, possibly via ordering and invoking further product customizations. In other words the processing on demand capabilities shall be fully integrated within the current data access capabilities of VITO.

This integration will provide a rich set of processing on demand capabilities and services to the users in the future, which is important to give non-expert users the means to retrieve relevant information from the extensive archive of EO-data and derived products at VITO.