Load testing of HELIDEM geo-portal: an OGC open standards interoperability example integrating WMS, WFS, WCS and WPS

Massimiliano Cannata
IST-DACD-SUPSI
Campus Trevano
CH-6952 Canobbio, Svizzera
+41(0)586666200
massimiliano.cannata@supsi.c

Milan Antonovic
IST-DACD-SUPSI
Campus Trevano
CH-6952 Canobbio, Svizzera
+41(0)58666620
milan.antonovic@supsi.ch

Monia Elisa Molinari
Politecnico di Milano, DICA,
Laboratorio di Geomatica,
Como Campus, Como, Italy
+390313327564
monia.molinari@polimi.it

ABSTRACT

This paper presents a load testing of the HELIDEM geo-portal, which is an example of interoperability between a numbers of standard geospatial services as defined by the Open Geospatial Consortium. The portal was developed within the European project HELIDEM (www.helidem.eu) with the aim of valorizing the main project output which is a cross-border digital terrain model. The portal aims at fostering its diffusion and usage trough easily accessible tools. The DTM covers the alpine area located between Southern Switzerland (Canton Ticino) and Northern Italy (Lombardy and Piedmont Regions). From a technological point of view, the server-side component of the portal is based on a Service Oriented Architecture implemented using the open source software Zoo-Project, GRASS GIS and Geoserver; the client-side component is a Web interface based on CSS3 and HTML5 trough the usage of the ExtJS framework and the OpenLayers software. The presented solution is a mix of technologies and software, some of which are considered, within the open source for geospatial community, mature and robust while others are considered promising but not sufficiently tested yet. For this reason this research conducted a load test over concurrent users in order to verify the robustness, quality and performance of the system and to identify eventual bottlenecks. Test results didn't register any exception confirming the good quality of the implemented system and underlying software. Nevertheless, performance and response time exponentially degrades with increasing number of concurrent users, area of analysis and process complexity. Finally, the test confirms that system is robust, in fact no system failure was recorded during the analysis.

Categories and Subject Descriptors

D.2.12 [**Software**]: Interoperability – *data mapping*

General Terms

Standardization, Verification, Performance, Algorithms, Design, Experimentation.

Keywords

OGC, WMS, WCS, WPS.

1. INTRODUCTION

Digital Terrain Models (DTMs) are a representation of the natural terrain surface by means of terrain altitude expressed in orthometric heights. In practice, these DTM are widely used to extrapolate derived information which is useful in supporting a large number of activities, including land use planning and engineering design [1], [3]. The derivation of contour lines or of profiles are certainly some of the most basic and common operations, but also the evaluation of terrain surface derivatives like slopes, aspects, curvature are very common. Watershed analysis from DTM is a more specific task which is generally performed by hydrologist, but also used for land planning and infrastructure design. These kinds of processing are common features of modern Geographical Information System software (GIS) which, although nowadays available to a larger public if compared with 10 years ago, still remain accessible to specialized personnel only.

The main scope of the HELIDEM project (Helvetia-Italy Digital Elevation Model) was the creation of a unified digital terrain model and geoid, for the Alpine and Sub-Alpine area on the border between Italy and Switzerland, correctly geo-referenced in the three dimensions. The project was funded in the frame of the European Regional Development Fund within the Italy-Switzerland cooperation program and run from September 2010 to September 2013.

The project leads to two new datasets: a DTM and a geoid of the cross-border region which includes northern Lombardy, northern Piedmont, Canton Ticino and southern parts of the Canton of Graubünden (Switzerland)and the Canton of Valais (Switzerland). The DTM was calculated in the ETRF2000 reference frame in geographic coordinate system with a resolution of $4*10^{-4}$ degrees (about 22 m in latitude and 15 m in longitude). The geoid was derived on the same region combining the Italian geoid ITALGEO2005 and the Swiss geoid CHGeo2004 with the inclusion of GOCE [2] data for solving the height datum problem. Both of them are distributed for free. Nevertheless, it has to be noted that they are not official datasets neither for Italy nor for Switzerland

A secondary, but not less important, goal of the project is the experimentation of the DTM, its diffusion and valorization trough easily accessible tools. In view of this least aim, the authors conducted research and development to the creation of a geo-spatial portal which could make the DTM and some of the commonly used derived data easily accessible. The requirement analysis, conducted internally within the project team, leads to the identification of the following requisites:

- Accessing at data and derived data without need of any software or components (plug-in) except of a Web Browser;

- Flexibility in the selection of the area of interest to be elaborates;
- Capability of download or visualize the results of an elaboration;
- Capability to derive data in the desired coordinate reference system;
- Accessing at contour, profiles, watershed, derivatives analysis and data extraction tools.

2. OWS: OPEN WEB SERVICES

In the last years, thanks to the large diffusion of Internet and the growing development of open standards related to geospatial sector, today we can access to a number of technologies and standard services which are well established and tested in productive environments. Yet, some of the most recent standards and related software are less verified and applied.

Among the available options, those defined by the Open Geospatial Consortium (OGC, www.opengeospatial.org) are certainly among the most used solutions for Web mapping services. Those services use the HTTP protocol, and in particular the GET and POST methods, to communicate with servers and specific XML schemas to encode exchanged information. In few cases, OGC standard services contemplate the usage the SOAP [4] protocol and of binary data (also in streaming) as a response.

In the next sections a brief description of a series of standards which has been selected to fulfill the HELIDEM geo-portal requirements are reported.

Web Mapping Service (WMS) [5] enables the access to geospatial data in the form of images throughout Internet. Received images represent cartographic elements rendered according to specific requested setting parameters. WMS is widely used and supported by almost all of the available GIS software and is a mature technology.

Web Feature Service (WFS) [6] defines methods and formats to request, also using spatial and semantic filters, and transmit geospatial vector data trough the Internet. With respect to the WMS, this standard permits to receive the actual data, and not only its graphical representation on an image. Moreover, this standard enables the capability of the user to remotely modify the data by means of transactional requests. Responses are always in geospatial vector formats like, for example, GML, KML or SHP.

Web Coverage Service (WCS) [7] can be considered similar to WFS with the difference that it handles distributed spatial data only. The standard defines rules to distribute, according to user's requests, raster data. Data could eventually be cropped, reprojected, resampled or converted in one of the supported raster formats like, for example, GeoTiff or ESRI ASCII grid.

Web Processing Service (WPS) [8] permits to provide geospatial processing capabilities over the Web. Thanks to this standard predefined processes or algorithms are exposed to Web users similarly to typical desktop GIS modules. The *GetCapabilities* and *DescribeProcess* requests allow to identify offered processes and their inputs and outputs names and types; the *Execute* requests allows to run a selected process.

3. THE HELIDEM SYSTEM

The HELIDEM system implements a Service Oriented Architecture (SOA) based on the OWSs presented in the previous chapter 2. The schematization of the implemented architecture, which is represented in Figure 1, is based on four OGC services that are deployed in the Cloud. Among them, the WPS only interacts directly with the others three by using them as input data provider (WCS) or as output data dispatching services (WMS, WFS). As well as the HELIDEM portal all the services are freely accessible on the Web, and can be consumed as stand-alone service to be integrated in other contexts: for example desktop GIS or specific apps for different devices (tablets, smartphones,, etc.).

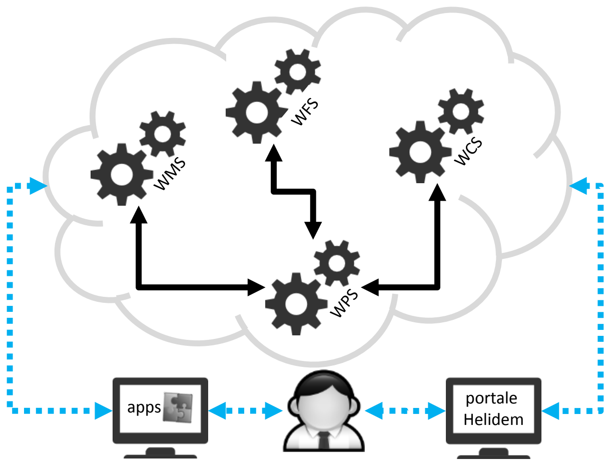


Figure 1. Schematization of the HELIDEM system architecture.

The implemented OWSs have been deployed using different servers which are physically located at the Institute of Earth Science at SUPSI in Lugano and at the Geomatics Laboratory at the Polytechnics of Milan in the Como Campus. Figure 2 and Figure 3 represent the technological stacks that have been used respectively for the implementation of the OSWs component (server side applications) and the HELIDEM geo-portal (client side application).

The server side component of the system is based on Linux OS (Ubuntu server) and Apache 2 Web server [9].

The data services (WMS, WFS and WCS) rely on Geoserver [10] deployed in the Tomcat 7 application server [11] and served by Apache HTTP server. Geoserver is a mature solution word widely used in large scale projects and in production systems, for example by NOAA or Ordinance Survey. It mostly relies on GeoTools libraries [12] and is developed in JAVA language [13].

The reliability of Geoserver is confirmed by the fact that it has the status of official Open Source Geospatial Foundation (OSGeo) project. One of OSGeo objectives is to group under its umbrella a number of Free and Open Source Software for Geospatial (FOSS4G) projects which have shown a collaborative development community and a high quality code: such a kind of software is promoted as OSGeo project. This status can only be achieved passing the incubation process which carefully reviews code and community [14].

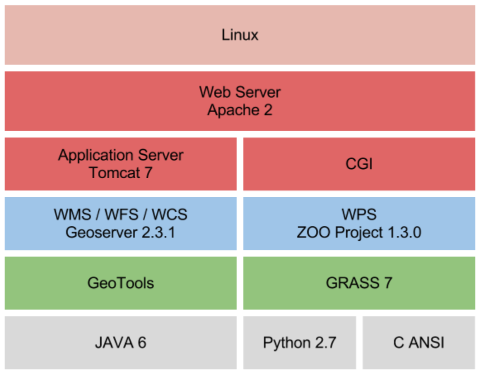


Figure 2. Software and languages stack used from the server sides of the HELIDEM system.

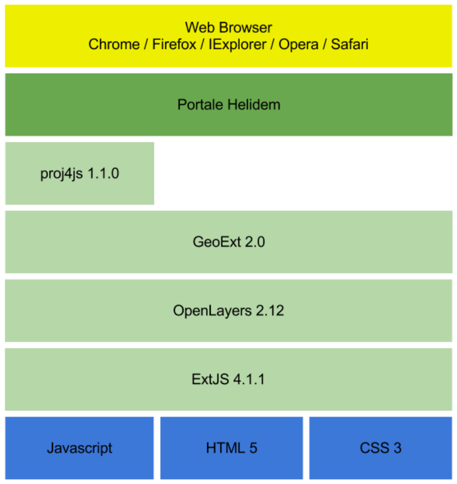


Figure 3. Software, libraries and languages stack used for the client implementation (HELIDEM portal).

The processing servers (WPS) take advantage of the Zoo-project software [15] which relies on CGI. The selection of this software to implement the WPS component of the system is because it has already successfully been used in conjunction with GRASS GIS version 7 and because it is with respect to other WPS solution less investigated. In fact, Zoo-project is quite recent and, at the time of writing, is under the OSGeo incubation process.

ZOO-project is a framework to create WPS compliant services; it is composed by three components (see Figure 4): (*i*) the ZOO kernel that is the engine, written in C [16], that enable the managing and chaining of different processes written in different programming languages; (*ii*) the ZOO services that implements the processes by means of a configuration file and function implementing the process algorithms in one of the supported languages (C, Python, Perl, PhP, JavaScript, Java, Fortran); and (*iii*) the ZOO API that is a JavaScript library to easy service chaining and interface development.

In HELIDEM, the WPS processes have been coded in Python [17] taking advantages of the python-requests library and of the modules of the GIS GRASS 7 [18] which are generally coded in C. The processes results are always returned as *simple output* type (link to a resource), so that they are of simple integration in third parties.

The HELIDEM geo-portal, as described in Figure 3 is based on the most modern technologies such as *CSS 3 and HTML 5* [19] and JavaScript [20] trough a number of libraries. ExtJS [21] was used to graphically design the portal thanks to the ability of this library to easy access advanced graphical elements guaranteeing cross browser compatibility; OpenLayers [22] to create the map viewer which allows for dynamically navigation of geospatial data; GeoExt [23] which combining OpenLayers and ExtJS provides additional features and controls; proj4js [24] used to provide on-the-fly reprojection of coordinates.

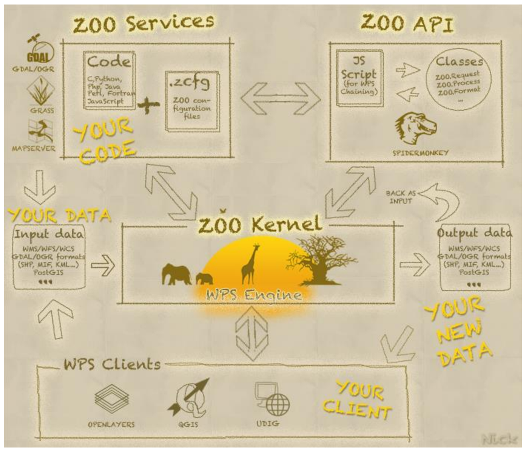
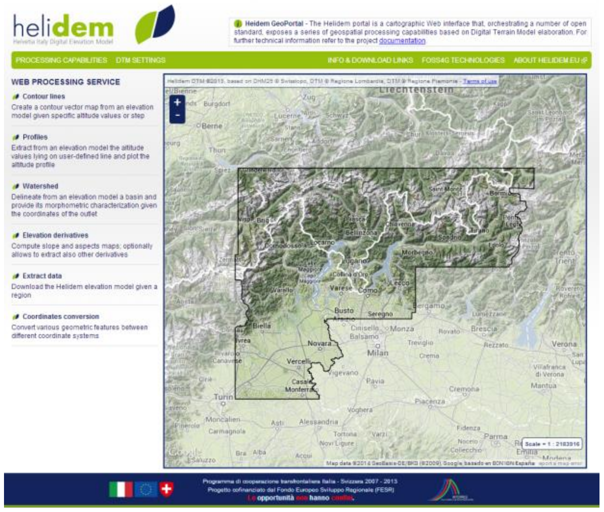


Figure 4. Zoo-project components and interations. Source: www.zoo-project.com, © Nicolas Bolzon

Thanks to the combination of those libraries and technologies the HELIDEM portal was designed and its graphical interface realized as shown in Figure 5. The interface is separated in three sections: a left-side panel which presents all the offered processing capabilities, a right-side map panel to represent contextual base map and processing results and an up-side panel which shows general information on the selected process, data accessibility and copyrights.



Heidem GeoPortal - The Helidem portal is a cartographic Web interface that, orchestrating a number of open standard, exposes a series of geospatial processing capabilities based on Digital Terrain Model elaboration. For further technical information refer to the project documentation

WEB PROCESSING SERVICE

Contour lines
Create a contour vector map from an elevation model given specific altitude values or step

Profiles
Extract from an elevation model the altitude values lying on user-defined line and plot the altitude profile

Watershed
Delineate from an elevation model a basin and provide its morphometric characterization given the coordinates of the outlet

Elevation derivatives
Compute slope and aspects maps; optionally allows to extract also other derivatives

Extract data
Download the Helidem elevation model given a region

Coordinates conversion
Convert various geometric features between different coordinate systems

Figure 5. HELIDEM geoportal. Base maps from Google.

4. HELIDEM processes

In Figure 5 the left-side panel shows a number of processing capabilities that are offered by the HELIDEM portal. Those processes have been implemented in the Zoo-project and offered by means of the WPS standard. The implemented Python code take advantage of specially implemented classes which enables the interaction with the GIS GRASS 7 and Geoserver.

The processes are thus orchestrated by the zoo-project process core while the core of the processing is designated to GRASS and then the results are pushed into Geoserver for easy access and map navigation by means of WMS and WFS capabilities. To successfully connect Python, or other languages, with GRASS the environment variable of GRASS and the integrity of concurrent operation shall be guaranteed [25]. To publish the geospatial data resulting from the process we took advantage of the RESTful API [26] of Geoserver and of the python-requests [27] library.

In order to achieve the automatic publishing of results, each implemented process includes a sub-process that:

- Export the output data in a specific formats (shapefiles for vectors and GeoTiff for raster) in a directory accessible by Geoserver.
- Add the data to Geoserver creating the store and the coverage with the RESTful API.
- Create the Style Layer Descriptor (SLD) file using specifically developed commands to export the GRASS styles in SLD.
- Upload and assign the style to the coverage with Geoserver RESTful API.

In the next sections the processes accessible trough the portal are shortly described indicating modules, inputs and outputs.

4.1 Data extraction

This feature has been implemented taking direct advantage of the WCS capability, in fact, once defined the bounding box and the coordinate system, the JavaScript compose a *getCoverage* request to the WCS server hosting the DTM. The result is available to the user as download.

4.2 Coordinate conversion

This capability is offered using the proj4js library and is mainly intended to allow defining indicative location and bounding box areas expressed in different coordinate system. In fact coordinates conversion uses predefined parameters which are not suitable for high precision transformations. No server processing is involved in this feature and all of the job is performed at browser level.

4.3 Contour lines

This feature is actually implemented in two WPS processes: given a defined elevation model the first allows for the extraction of contour lines at a predefined list of altitudes (Table 12) while the second at defined intervals among two extremes. The processes are written in Python and takes advantage of *r.contour* module of GRASS.

Table 1. Contour levels at defined altitudes: inputs and outputs type: i/s = input type string, i/d = input type decimal, o/U = output type URI, o/s = output type string, o/dt = output type datetime; M = mandatory, N = multiplicity).

Parameter	Type	Description	M	N
covermap	i/s	Link to a DTM in geoTiff	Y	1
coverSRS	i/s	EPSG code (EPSG:XXXX)	Y	1
bbox	i/s	BBOX (EWKT)	Y	1
levels	i/d	CSV list of values	Y	n
oSRS	i/s	EPSG code	Y	1
oformat	i/s	output format: GML, KML or Shapefile (default)	N	1
odata	o/U	Link to the contour level zipped file	Y	1
WMS_URL	o/U	WMS address to access the results	Y	1
WFS_URL	o/U	WFS address to access the results	Y	1
layerName	o/s	Laye name to be used in WMS e WFS	Y	1
Expiration Time	o/dt	Exipration date of the results	Y	1
message	o/s	Additional process information	Y	1

Table 2. Contour levels at defined intervals: inputs and outputs type: i/s = input type string, i/d = input type decimal, o/U = output type URI, o/s = output type string, o/dt = output type datetime; M = mandatory, N = multiplicity).

Parameter	Type	Description	M	N
covermap	i/s	Link to a DTM in geoTiff	Y	1
coverSRS	i/s	EPSG code (EPSG:XXXX)	Y	1
bbox	i/s	BBOX (EWKT)	Y	1
step	i/d	interval	Y	n
minLevel	i/d	Minimum altitude	N	1
maxLevel	i/d	Maximus altitude	N	1
oSRS	i/s	EPSG code	Y	1
oformat	i/s	output format: GML, KML or Shapefile (default)	N	1
odata	o/U	Link to the contour level zipped file	Y	1
WMS_URL	o/U	WMS address to access the results	Y	1
WFS_URL	o/U	WFS address to access the results	Y	1
layerName	o/s	Layer name to be used in WMS e WFS	Y	1
Expiration Time	o/dt	Expiration date of the results	Y	1
message	o/s	Additional process information	Y	1

4.4 Profiles

Provided a linestring this process allows to extract the altimetric profile along the path as CSV file of 3D coordinates and PNG image. This feature is based on the *r.profile* module of GRASS combined with the Python library matplotlib [28].

Table 3. Profile extraction process: inputs and outputs. (type: i/s = input type string, i/d = input type decimal, o/U = output type URI, o/s = output type string, o/dt = output type datetime; M = mandatory, N = multiplicity).

Parameter	Type	Description	M	N
covermap	i/s	Link to a DTM in geoTiff	Y	1
coverSRS	i/s	EPSG code (EPSG:XXXX)	Y	1
bbox	i/s	BBOX (EWKT)	Y	1
coord	i/d	Linestring in EWKT format	Y	n
oSRS	i/s	EPSG code	Y	1
oformat	i/s	output format: GML, KML or Shapefile (default)	N	1
outImage	o/U	PNG of the profile	Y	1
outText	o/U	CSV of 3D coordinates	Y	1
Expiration Time	o/dt	Expiration date of the results	Y	1
Message	o/s	Additional information	Y	1

4.5 Elevation derivatives

Using the module *r.slope.aspect* this process generates raster maps of slope, aspect, curvatures and partial derivatives from an elevation raster map. Aspect is calculated counterclockwise from east.

Table 4. Elevation derivatives process: inputs and outputs. (type: i/s = input type string, i/d = input type decimal, o/U = output type URI, o/s = output type string, o/dt = output type datetime; M = mandatory, N = multiplicity).

Parameter	Type	Description	M	N
covermap	i/s	Link to a DTM in geoTiff	Y	1
coverSRS	i/s	EPSG code (EPSG:XXXX)	Y	1
bbox	i/s	BBOX (EWKT)	Y	1
oSRS	i/s	EPSG code	Y	1
slope	i/s	Create slope map (yes/no)	N	1
format	i/s	Slope format (degrees/perc.)	N	1
aspect	i/s	Create slope map (yes/no)	N	1
pcurv	i/s	Create slope map (yes/no)	N	1
tcurv	i/s	Create tangential map (yes/no)	N	1
dx, dy, dxx, dyy, dxy	i/s	Create derivative map (yes/no)	N	1
odata	o/U	Zip of output maps	Y	1
slopeLayer	o/U	Layer name	Y	1
aspectLayer	o/U	Layer name	Y	1
pcurvLayer	o/U	Layer name	Y	1
tcurvLayer	o/U	Layer name	Y	1
dxLayer,	o/U	Layer name	Y	1
dyLayer,	o/U	Layer name	Y	1
dxxLayer	o/U	Layer name	Y	1
dyyLayer	o/U	Layer name	Y	1
dxylayer	o/U	Layer name	Y	1
WMS_URL	o/U	WMS address to results	Y	1
WFS_URL	o/U	WFS address to results	Y	1
Expiration Time	o/dt	Expiration date of the results	Y	1
message	o/s	Additional process information	Y	1

4.6 Watershed analysis

This process allows to conduct a number of analysis of a basin taking advantage of the *r.basin* command [29] specifically ported during this work at the version 7 of GRASS and integrated with a the newly developed modules named *r.nearest.coord*, which find the coordinates of the nearest cell having value higher of a threshold, and *r.out.sld*, which produce valid style descriptor.

Given a DTM the process evaluate the flow accumulation and flow direction maps [30] using one of the possible approaches: the Single Flow Direction (SFD) which allows the water to flow in one cell only using the maximum slope criterion or the Multiple Flow Direction (MFD) which allows the water to flow in all the neighbor lower cells at the same time proportionally to the slope. SFD approach is further subdivided in two algorithms named D8 if all the neighbor cells are considered or D4 if only the 4 neighbors cells along the cardinal directions are considered.

Combining the calculated flow accumulation and flow direction maps with the coordinate of the closing section of the watershed and a threshold value of the accumulation, the process calculates the basin and the hydrographic network. These maps are further used in the process to perform a number of analyses and elaboration that allows the morphological and hydrological characterization of the watershed producing a number of outputs including raster maps, plots and reports.

The produced raster maps are: hierarchical classification of the hydrographic network according to Horton, Strahler, Hack and Shreve, distance to outlet and the length of the slopes. The process, using the *r.out.sld* module, produces for each map a Styled Layer Descriptor (SLD) [31] which defines symbolization and coloring of layers.

The created plots are: hypsographic and hypsometric curves and width function [32]. The reports are: CSV file and PDF report with a number of morphometric parameters, including center of gravity, area, perimeter, mean slope, length of the directing vector, prevalent orientation, characteristic altitudes, shape factors, topological diameter, magnitude, Horton ratios and concentration time, drainage density. The PDF have been produced using the pyUNO library [33] which allows to produce documents using the libreoffice API.

Table 5. Watershed analysis process: inputs and type: i/s = input type string, i/d = input type decimal, o/U = output type URI, o/s = output type string, o/dt = output type datetime; M = mandatory, N = multiplicity).

Parameter	Type	Description	M	N
covermap	i/s	Link to a DTM in geoTiff	Y	1
coverSRS	i/s	EPSG code (EPSG:XXXX)	Y	1
bbox	i/s	BBOX (EWKT)	Y	1
coord	i/d	Point in EWKT format	Y	1
threshold	i/s	Flow accumulation threshold	Y	1
method	i/s	Flow direction algorithm	Y	1
flat_area	i/s	Beatify of flat area	Y	1
oSRS	i/s	EPSG code	Y	1
parameters	o/U	CSV report	Y	1
pdf_report	o/U	PDF report	Y	1
outmaps	o/U	Zip of output raster maps	Y	1
ipsographic_curve	o/U	PNG of hypsographic curve	Y	1
ipsometric	o/U	PNG of hypsometric curve	Y	1

_curve				
width_function	o/U	PNG of width function	Y	1
WMS_URL	o/U	WMS address to access the results	Y	1
WFS_URL	o/U	WFS address to access the results	Y	1
networkLayer	o/s	WMS/WFS River network layer name	Y	1
dist2outLayer	o/s	WMS/WFS distance to outlet layer name	Y	1
accumulationLayer	o/s	WMS/WFS flow accumulation layer name	Y	1
hillslope_distanceLayer	o/s	WMS/WFS hillslope length layer name	Y	1
hackLayer	o/s	Hack classification layer name	Y	1
hortonLayer	o/s	Horton classification layer name	Y	1
shreveLayer	o/s	Shreve classification layer name	Y	1
strahlerLayer	o/s	Strahler classification layer name	Y	1
Expiration Time	o/dt	Expiration date of the results	Y	1
message	o/s	Additional process information	Y	1

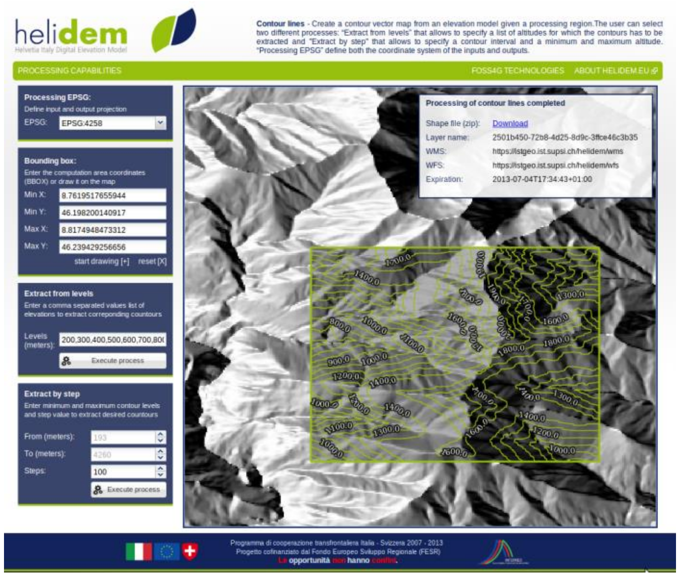
5. HELIDEM GEOPORTAL

The HELIDEM geo-portal is the official access point of the project to the produced data. Moreover, it gives access to the previously described processes, which even if available as stand-alone services, are best fitted to be chained by means of the portal business logic.

When the user selects a desired process (see Figure 6), the left-side panel allows to set the required inputs parameters by means of selecting available options, inserting values and text, or interacting with the map to derive geospatial features like bounding boxes, point or polylines; it worth to be noted that all the geospatial features are also editable as text box, so that expert users could define these parameters by means of known coordinates.

When the process is executed, it runs asynchronously and the server keep updated the execute response document indicating the current status of the process; the client periodically check for process status description, so that the user is aware if the process is running, what is the approximate percentage of execution and eventual fails.

When the process execution ends, results are made available within a collapsible frame on top right of the map: link to zipped layers, layer names, WMS and WFS URLS, images and expiration date are opportunely reported. At the same time, derived geospatial data are represented in the map as additional WMS layer; in case of multiple layers in the outputs (see Figure 7), the user has the ability to select one of them for its visualization.


helidem
Helvetia Italy Digital Elevation Model
PROCESSING CAPABILITIES
FOSS4G TECHNOLOGIES
ABOUT HELIDEM.EU
start drawing [+]
reset [X]
Execute process
Programma di cooperazione transfrontaliera Italia - Svizzera 2007 - 2013
Progetto cofinanziato dal Fondo Europeo Sviluppo Regionale (FESR)
La opportunità non hanno confini.

Contour lines - Create a contour vector map from an elevation model given a processing region.The user can select two different processes: "Extract from levels" that allows to specify a list of altitudes for which the contours has to be extracted and "Extract by step" that allows to specify a contour interval and a minimum and maximum altitude. "Processing EPSG" define both the coordinate system of the inputs and outputs.

Processing EPSG:
Define input and output projection
EPSG: EPSG:4258

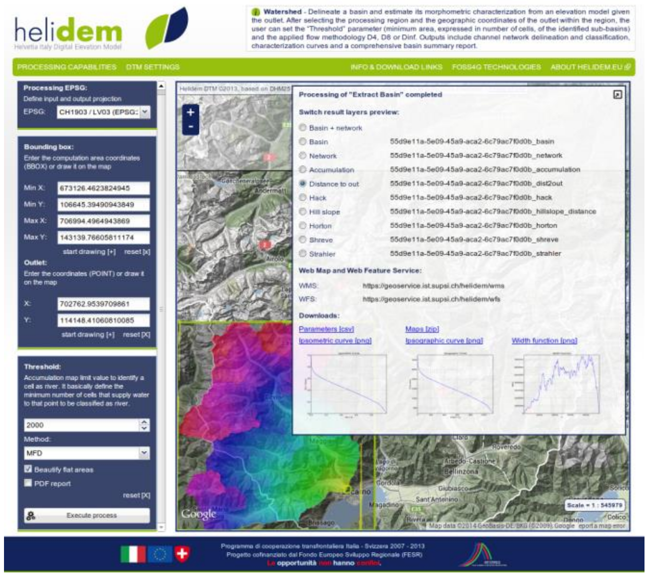
Bounding box:
Enter the computation area coordinates (BBOX) or draw it on the map
Min X: 8.7619517655944
Min Y: 46.198200140917
Max X: 8.8174948473312
Max Y: 46.239429256656

Extract from levels
Enter a comma separated values list of elevations to extract correponding countours
Levels (meters): 200,300,400,500,600,700,80(

Extract by step
Enter minimum and maximum contour levels and step value to extract desired countours
From (meters): 193
To (meters): 4260
Steps: 100

Processing of contour lines completed
Shape file (zip): Download
Layer name: 2501b450-72b8-4d25-8d9c-3ffce46c3b35
WMS: https://istgeo.ist.supsi.ch/helidem/wms
WFS: https://istgeo.ist.supsi.ch/helidem/wfs
Expiration: 2013-07-04T17:34:43+01:00

Figure 6. Example of processing interface with user: input on the left-side panel, results in the top-right frame and on map.


helidem
Helvetia Italy Digital Elevation Model
PROCESSING CAPABILITIES
DTM SETTINGS
INFO & DOWNLOAD LINKS
FOSS4G TECHNOLOGIES
ABOUT HELIDEM.EU
start drawing [+]
reset [x]
Execute process
Google
Scale = 1 : 545979
Programma di cooperazione transfrontaliera Italia - Svizzera 2007 - 2013
Progetto cofinanziato dal Fondo Europeo Sviluppo Regionale (FESR)
La opportunità non hanno confini.

Watershed - Delineate a basin and estimate its morphometric characterization from an elevation model given the outlet. After selecting the processing region and the geographic coordinates of the outlet within the region, the user can set the "Threshold" parameter (minimum area, expressed in number of cells, of the identified sub-basins) and the applied flow methodology D4, D8 or Dinf. Outputs include channel network delineation and classification, characterization curves and a comprehensive basin summary report.

Processing EPSG:
Define input and output projection
EPSG: CH1903 / LV03 (EPSG:

Bounding box:
Enter the computation area coordinates (BBOX) or draw it on the map
Min X: 673126.4623824945
Min Y: 106645.39490943849
Max X: 706994.4964943869
Max Y: 143139.76605811174

Outlet:
Enter the coordinates (POINT) or draw it on the map
X: 702762.9539709861
Y: 114148.41060810085

Threshold:
Accumulation map limit value to identify a cell as river. It basically define the minimum number of cells that supply water to that point to be classified as river.
2000
Method: MFD
[✓] Beautify flat areas
[ ] PDF report

Processing of "Extract Basin" completed
Switch result layers preview:

Basin + network	
Basin	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_basin
Network	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_network
Accumulation	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_accumulation
● Distance to out	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_dist2out
Hack	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_hack
Hill slope	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_hillslope_distance
Horton	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_horton
Shreve	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_shreve
Strahler	55d9e11a-5e09-45a9-aca2-6c79ac7f0d0b_strahler

Web Map and Web Feature Service:
WMS: https://geoservice.ist.supsi.ch/helidem/wms
WFS: https://geoservice.ist.supsi.ch/helidem/wfs

Downloads:
Parameters [csv] Maps [zip]
Ipsometric curve [png] Ipsographic curve [png] Width function [png]

Figure 7. Example of processing interface with multiple output layers and different output types.

6. LOAD TESTING

Load testing is performed to determine a system's behavior under both normal and anticipated peak load conditions. In the next paragraphs system configuration, test settings and results are presented.

As previously described the HELIDEM geo-portal relies on a server-side component composed of several OWS services which are located on two different servers: a server for data processing and results dispatch implementing the WPS, WMS and WFS (herein after referred to as "WPS server") and a server for DTM serving implementing the WCS (herein after referred to as "WCS server"). Both are virtual machines set up using Oracle VM

VirtualBox [34] whose characteristics are summarized in Table 67.

Table 6. Servers characteristics.

parameter	WPS server	WCS server
OS	Ubuntu server version 12.04, 32 bit	Ubuntu server version 12.04, 32 bit
RAM	4GB	4GB
CPU	Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz	Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz
N° of processors	6	4
Disk size	100GB	100GB

While the Geoserver instance has 2 GB of RAM assigned the Zoo-project does not have any pre-allocated memory and thus the limit is that of the operating system (4.5 GB including SWAP memory).

The load testing has been conducted using an open Source framework named Locust [35]. Locust is a scalable and distributed framework developed in Python and available under the MIT-license. This framework enable the set up of a load test under different scenarios identified by the number of concurrent users and the *hatch-ratio* which indicate the user spawned for seconds.

Locust test requires a "locustfile" which is a normal python file that overrides the *TaskSet* and *Locust* classes. The extension of the *TaskSet* class defines the user behaviors by mean of a series of tasks the simulated user would perform: generally those tasks are HTTP POST and GET requests. The extension of the *Locust* class represents one user which is defined trough few attributes: the *task_set* which specifies the operations to be considered (the implemented extension of the TaskSet class), the *min_wait* and *max_wait* which are respectively the minimum and maximum time, in milliseconds, that a simulated user will wait between executing each task. Beyond the standard HTTP error status codes (400-499) automatically detected by Locust, exception responses which are returned with a success HTTP status code (e.g. 200) can be opportunely caught and reported as task failure.

Operationally, when a test is started, each instance of the *Locust* classes (each concurrent user) start calling its *task_set* which pick one of it's task and call it. It will then wait a random number of milliseconds, between *min_wait* and *max_wait*, and then lunch a new task, and so on.

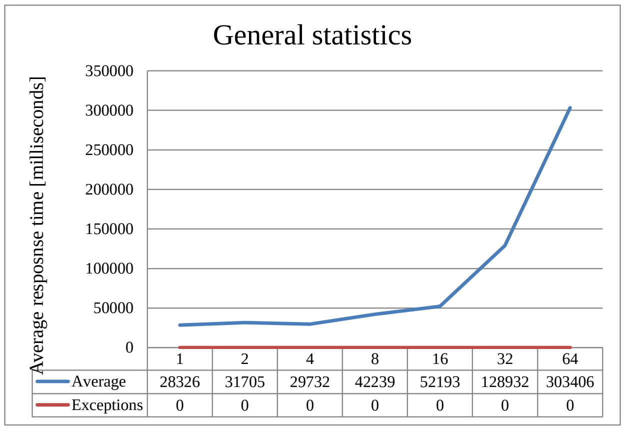
For the scope of the study, because the HELIDEM geo-portal is essentially an interface to a number of geo-processing features offered trough the WPS that orchestrate other services, the load testing is focused on the implemented WPS processes. Thus, the load test was conducted setting up a single user type – HELIDEM geo-portal user – performing four different tasks on four different processing areas. Tasks are contour extraction, profile calculation, basin analysis and derivative elaboration; areas as reported in Table 78 were sampled in order to include large, medium, small and very small basins with respect to the Alpine morphology. The combination of task and area for each request is selected randomly.

Table 7. Processing areas used in load testing; bbox is expressed as [minx,miny, maxx,maxy] in CH1903 coordinate system and areas in sqkm.

Area name	Area	bbox	Basin type
Maggia	1236	673126,106645, 706994,143139	Large
Verzasca	488	694828,114091, 712831,141198	Medium
Breggia	210	719542,76682, 733188,92109	Small
Cama	43	733520,120819, 740044,127433	Very small

Different load tests were conducted simulating respectively 1, 2, 4, 8, 16, 32 and 64 concurrent users; each test was run for approximately 2 hours. The low number of simulated concurrent users depends on the fact that this is a specific portal expected to be used by a limited number of specialists and therefore with a very limited traffic with respect to other services like those offered by social network where million of concurrent users are easily registered. Because of the expected duration of the processes (several tenths of a second), the *max_wait* was set to 2 minutes while the *min_wait* was set to 5 seconds; used *hatch-ratio* value was 1.

Figure 8 shows the average response time for the different scenarios of concurrent users without discrimination of request type. This plot clearly shows that the quality of the system is very high in fact no exception were raised over a total number of 3380 requests. Nevertheless, the performance of the system quickly degrades with increasing number of concurrent users, particularly evident with more then 16 users.

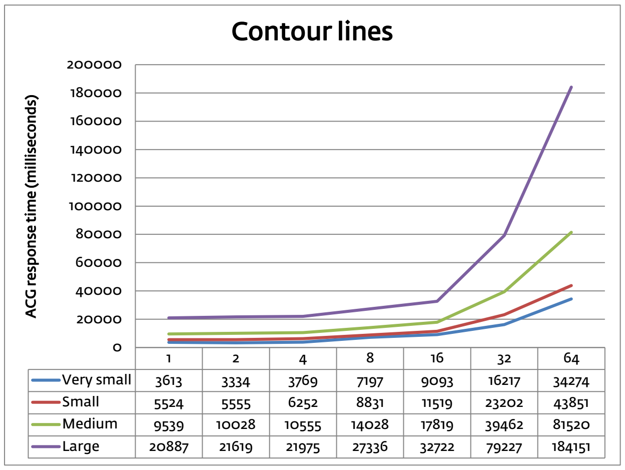


	1	2	4	8	16	32	64
Average	28326	31705	29732	42239	52193	128932	303406
Exceptions	0	0	0	0	0	0	0

Figure 8. Average response time during the whole test.

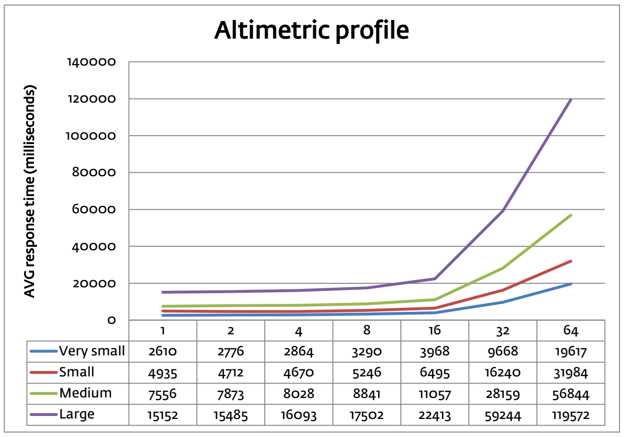
Figure 9, Figure 10, Figure 11 and Figure 12 illustrate the response time in milliseconds for each analyzed process (Locust's task) over the growing number of concurrent users.

The fastest process is the extraction of profile with response time that varies from 2.5 seconds (one users and very small basin) to almost 2 minutes (64 users and large basin) with an average of 18 seconds. The Contours extraction registered a response time between a minimum of 3.6 seconds and a maximum of 3 minutes with an average of 26.9 seconds. The time required by the user to get a response from the server in case of calculation of elevation derivative ranges from 8.4 seconds to 17.4 minutes with an average of 2.2 minutes. The longest process is the hydro-morphological analysis of watershed where a minimum of 24 seconds, a maximum of 23.5 minutes and an average of 3.15 minutes were recorded. WCS service response time (see Figure 13) vary smoothly with respect to WPS processes.



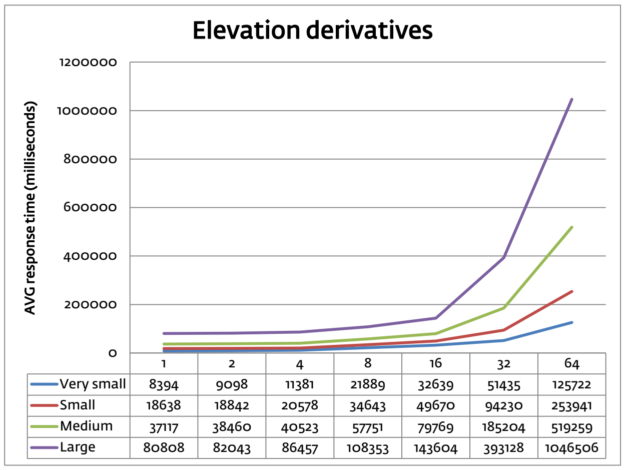
	1	2	4	8	16	32	64
Very small	3613	3334	3769	7197	9093	16217	34274
Small	5524	5555	6252	8831	11519	23202	43851
Medium	9539	10028	10555	14028	17819	39462	81520
Large	20887	21619	21975	27336	32722	79227	184151

Figure 9. Average response time during contour lines process.



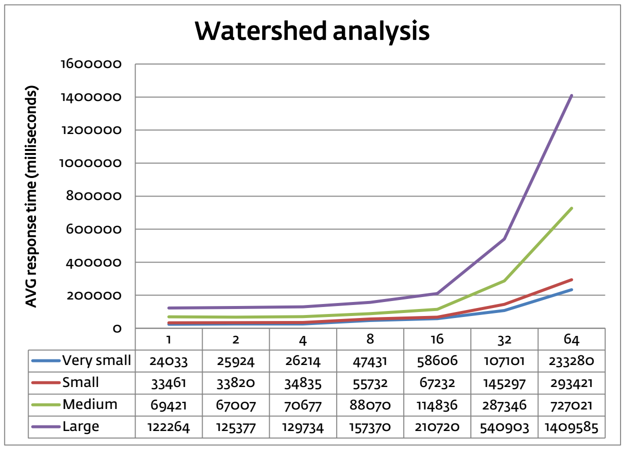
	1	2	4	8	16	32	64
Very small	2610	2776	2864	3290	3968	9668	19617
Small	4935	4712	4670	5246	6495	16240	31984
Medium	7556	7873	8028	8841	11057	28159	56844
Large	15152	15485	16093	17502	22413	59244	119572

Figure 10. Average response time during profile extraction process.



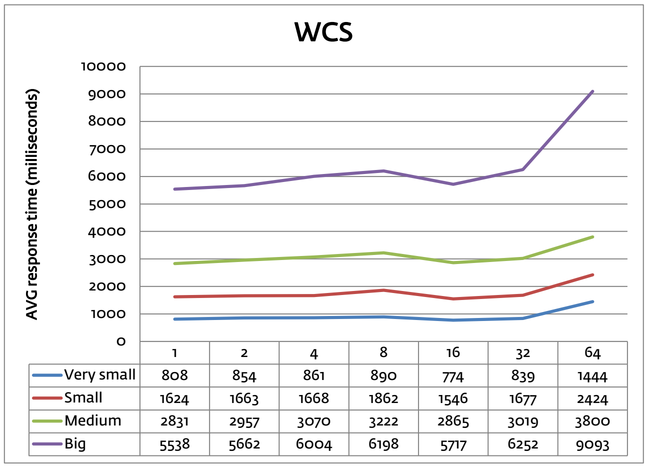
	1	2	4	8	16	32	64
Very small	8394	9098	11381	21889	32639	51435	125722
Small	18638	18842	20578	34643	49670	94230	253941
Medium	37117	38460	40523	57751	79769	185204	519259
Large	80808	82043	86457	108353	143604	393128	1046506

Figure 11. Average response time during elevation derivatives process.



	1	2	4	8	16	32	64
Very small	24033	25924	26214	47431	58606	107101	233280
Small	33461	33820	34835	55732	67232	145297	293421
Medium	69421	67007	70677	88070	114836	287346	727021
Large	122264	125377	129734	157370	210720	540903	1409585

Figure 12. Average response time during watershed analysis process.



	1	2	4	8	16	32	64
Very small	808	854	861	890	774	839	1444
Small	1624	1663	1668	1862	1546	1677	2424
Medium	2831	2957	3070	3222	2865	3019	3800
Big	5538	5662	6004	6198	5717	6252	9093

Figure 13. Average response time during data extraction from Web Coverage Service.

It worth to be noted that during the first phase of test settings we have experienced some issues related to disk space. In fact, results are stored for 24 hours and tend to exhaust disk space as the number of requests increase. Nevertheless, even if exception was raised, the system showed a good robustness to cope with errors during execution and to continue to operate despite anomalies: no crash and no downtime were recorded.

7. CONCLUSIONS

This work has presented the realization of a geospatial portal based on OWSs which uses Web Processing Service as the main component orchestrating Web Coverage Service for data gathering and Web Feature Service and Web Mapping Service for result dispatch and visualization. Load testing has been performed in order to understand the system behavior under real case scenarios and concurrent access.

Test results show that the system has a good robustness and good quality; in fact no system failure and no exception response were registered. Performances are relatively good if compared with desktop processing and considering data loading times. Nevertheless, system response time exponentially degrades with process complexity, increasing number of users and analyzed areas size. As a result, response time of WPS vary from about 2.5 seconds (to satisfy a single user requesting an altimetric profile over a very small basin area) to more then 20 minutes (to provide a watershed analysis over a large basin when 64 concurrent users are accessing the application).

Response time, as expected, is dependent on process algorithms complexity and increases with it. This is confirmed by tests, in fact higher latency can be observed from profile calculation to contour line extrapolation, to elevation derivatives creation, to watershed analysis. Looking at a single process the causes of service performance degradation are the presence of concurrent users and the size of the elaboration area.

The size of the elaboration area affect the response time because it set the area to be processed and / or because it defines the size of the data to be extracted and downloaded. To better understand the influence of remote access to data in a distributed system like this a WCS test was conducted. Results show that response time is sensible to the number of concurrent user, but it also shows that if compared with WPS response time, it is slowly degrading. This means that, when a higher number of concurrent users are operating its impact is reduced. In the case of DTM derivatives calculation the WCS relative cost in term of time, decrease from about 4% when a single user is accessing the system to about 0.7% with 64 concurrent users. Moreover, the relative impact of data gathering is depending on the complexity of the process: in simpler processes like the profile extraction it accounts in average for about the 25% of the whole process, in more complex processes like basin analysis it is almost negligible counting for less then 3%.

From the previously presented considerations on data gathering cost, it can be deduced that the most important cause of HELIDEM service performance degradation is the concurrent processing. In particular, the bottleneck of this system is the CPUs load. The CPU is also quickly exhausted: during the tests starting from 16 users the CPUs usage was around 80% and from 32 users a usage rate of 100% was observed.

A final consideration to be taken into account is the disk space. If results are made available to users for a period, trough standard services and/or compressed archive, then the disk may run out of space. This is particularly important when long data availability periods are used and processes produce large data: derivatives process in the case study outputs up to 10 raster maps and about 30 Mb of data for large basin. A combination between average size of process outputs, length of the period of output availability and maximum number of request expected in the same period should be used to define the required disk space.

To improve the performance of the system under intensive concurrent accesses, the increase of computational power and disk space seems to be the natural solution to override these limitations: a scalable cloud computing service, like the Amazon Elastic Compute Cloud (Amazon EC2), could resolve the issue.

Nevertheless more research on process optimization should be conducted. For example, it would be interesting to verify the impact of asynchronous programming over response time; in fact it is a well known approach to reduce the user waiting time by freeing resources when are not needed (generally during I/O operations like data download or file reading and writing) and making them available to other requests in accordance with the non-blocking paradigm. Another interesting option which could be investigated is the optimization of data storage, access and serving, like the Rasdaman array database [36].

In summary this research shows that the used software stack is robust and of good quality and that response time for processing digital terrain model services, also when complex operations are required, is reasonable when processing power is balanced with the number of concurrent users.

8. ACKNOWLEDGMENTS

Our thanks to the HELIDEM project team for the collaborative work, the Geodesy Division of swisstopo and to the UMUG office of the Canton Ticino for supporting this project.

9. REFERENCES

[1] Vieux, B. E. 2001. Distributed hydrologic modeling using GIS (pp. 1-17). Springer Netherlands.

[2] Barzaghi, R., Conte, R., Falletti, G., Maggi, A., Martino, M., Migliaccio, F., ... & Tselfes, N. 2006. Exploitation of GOCE data for a local estimate of gravity field and geoid in the Piemonte area (northern Italy). In Atti del 3rd International GOCE User Workshop (pp. 6-8).

[3] Wilson, J. P., & Gallant, J. C. 2000. Digital terrain analysis. Terrain analysis: Principles and applications, 1-27.

[4] Box, D., Ehnebuske, D., Kakivaya, G., Layman, A., Mendelsohn, N., Nielsen, H. F., ... & Winer, D. 2000. Simple object access protocol (SOAP) 1.1.

[5] de la Beaujardiere, J. 2006. OpenGIS® web map server implementation specification. Standard Specification, 06-042.

[6] Vretanos, P. A. 2010. OpenGIS Web Feature Service 2.0 Interface Standard. Open Geospatial Consortium Inc, Version, 2(0).

[7] Baumann, P. 2010. OGC WCS 2.0 interface standard—core. Open Geospatial Consortium Inc., Wayland, MA, USA, OpenGIS® Interface Standard OGC.

[8] Schut, P., & Whiteside, A. 2007. OpenGIS web processing service. OGC project document.

[9] The Apache Software Foundation. 2013. Apache HTTP Server Version 2.2 Documentation. Online resource: http://httpd.apache.org/docs/2.2.

[10] Youngblood, B. 2013. GeoServer Beginner's Guide. Packt Publishing Ltd..

[11] The Apache Software Foundation. 2013. Apache Tomcat 7 – Documentation Index. Online resource: http://tomcat.apache.org/tomcat-7.0-doc/index.html

[12] Turton, I. 2008. Geo tools. In Open source approaches in spatial data handling (pp. 153-169). Springer Berlin Heidelberg.

[13] Gosling J., Joy B., Steele G., Bracha G., Buckley A.. 2013. The Java Language Specification, Java SE 7 Edition. Addison-Wesley Professional, ISBN 9780133260229.

[14] Brovelli, M. A., Mitasova, H., Neteler, M., & Raghavan, V. 2012. Free and open source desktop and Web GIS solutions. Applied Geomatics, 4(2), 65-66.

[15] Fenoy, G., Bozon, N., & Raghavan, V. 2013. ZOO-Project: the open WPS platform. Applied Geomatics, 5(1), 19-24.

[16] International Organization for Standardization and International Electrotechnical Commission, ISO/IEC 9899:1999, Programming languages—C

[17] Van Rossum, G. 2007. Python Programming Language. In USENIX Annual Technical Conference.

[18] Neteler, M., Bowman, M. H., Landa, M., & Metz, M. 2012. GRASS GIS: A multi-purpose open source GIS. Environmental Modelling & Software, 31, 124-130.

[19] Hogan, B. P. 2011. HTML5 and CSS3: Develop with Tomorrow's Standards Today. Pragmatic Bookshelf.

[20] Flanagan, D. 2006. JavaScript: the definitive guide. " O'Reilly Media, Inc.".

[21] ZHANG, J. F., WANG, J. X., & JIA, X. R. 2010. Design and implementation of Data Maintenance System based on ExtJS [J]. Railway Computer Application, 1, 014.

[22] Hazzard, E. (2011). OpenLayers 2.10 Beginner's Guide. Packt Publishing Ltd.

[23] GeoExt Community. 2010. *GeoExt Documentation*, Online resource: http://www.geoext.org/docs.html.

[24] Adair M., Greenwood R., Richard D., Irons S. , Terral O. 2012. Proj4js -- Javascript reprojection library, https://github.com/proj4js/proj4js

[25] Cannata, M., Molinari, M. E., Luan, T. X., & Long, N. H. 2012. Web Processing Services for Shallow Landslide. International Journal of Geoinformatics, 8(1).

[26] Díaz, L., Granell, C., Gould, M., & Huerta, J. 2011. Managing user-generated information in geospatial cyberinfrastructures. Future Generation Computer Systems, 27(3), 304-314.

[27] Reitz K.. 2012. Requests: HTTP for Humans, http://docs.python-requests.org/en/latest

[28] Hunter J.D. 2007. *Matplotlib: A 2D Graphics Environment*, Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95

[29] Di Leo, M., & Di Stefano, M. 2013. An Open-Source Approach for Catchment's Physiographic Characterization. In AGU Fall Meeting Abstracts (Vol. 1, p. 06).

[30] Tarboton, D. G., & Ames, D. P. 2001. Advances in the mapping of flow networks from digital elevation data. In World water and environmental resources congress (pp. 20-24). USA: Am. Soc Civil Engrs.

[31] Lupp, M. 2007. Styled layer descriptor profile of the web map service implementation specification. Open Geospatial Consortium Inc. OGC, 1(0).

[32] Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. Civ. Eng, 101, 1258-1262.

[33] OpenOffice.org. 2010. Python-UNO bridge, http://www.openoffice.org/udk/python/python-bridge.html

[34] Romero, A. V. (2010). VirtualBox 3.1: Beginner's Guide. Packt Publishing Ltd. Chicago

[35] Jonatan Heyman, Carl Byström, Joakim Hamrén, User load testing tool for web services, ESN Social Software, 2011-11-05, Available at: http://locust.io and version 0.4 of Locust https://github.com/locustio/locust

[36] Owonibi, M., Baumann, 2012. P.: D-WCPS: A Framework for Service Based Distributed Processing of Coverages.. Proc. GEOProcessing 2012, The Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services, Valencia, Spain, January 30, 2012, pp. 215 – 221.