A New Geo-portal for MODIS/SEVIRI Image Products with Geolocation-Based Retrieval Functionality

Jorge Sevilla,[a] Yves Julien,[b] Guillem Sória,[b] José A. Sobrino[b], Antonio Plaza[a]

[a] Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Avda. de la Universidad s/n, Cáceres, Spain, 10003

[b] Global Change Unit, Imaging Processing Laboratory, University of Valencia, C/ Catedrático Agustín Escardino nº9, 46980 Paterna, Spain.

Abstract. A large number of remote sensing data sets have been collected in recent years by Earth observation instruments such as the moderate resolution imaging spectroradiometer (MODIS) aboard the Terra/Aqua satellite and the spinning enhanced visible and infrared imager (SEVIRI) aboard the geostationary platform Meteosat Second Generation (MSG). The advanced remote sensing products resulting from the analysis of these data are useful in a wide variety of applications, but require significant resources in terms of storage, retrieval and analysis. Despite the wide availability of these MODIS/SEVIRI products, the data coming from these instruments are spread among different locations and retrieved from different sources, and there is no common data repository from which the data or the associated products can be retrieved. In this paper, we take a first step towards the development of a geo-portal for storing and efficiently retrieving MODIS/SEVIRI remote sensing products. The products are obtained using an automatic system that processes the data as soon as they are provided by the collecting antennas, and then the final products are uploaded one day delay in the geo-portal. Our focus in this work is on describing the design and efficient implementation of the geo-portal, which allows for a user-friendly and effective access to a full repository of MODIS/SEVIRI advanced products (comprising tens of terabytes of data) using geo-location retrieval capabilities. The geo-portal, which is available online at `http://ceosspain.lpi.uv.es`, has been implemented as a web application composed of different layers. Its modular design provides quality of service and scalability (capacity for growth without any quality losing), allowing for the addition of components without the need to modify the entire system. On the client layer, an intuitive web browser interface provides users with remote access to the system. On the server layer, the system provides advanced data management and storage capabilities. On the storage layer, the system provides a secure massive storage service. An experimental evaluation of the geo-portal in terms of efficiency and product retrieval accuracy is also presented and discussed.

Keywords: MODIS, SEVIRI, geo-portal, digital repository, geolocation-based image retrieval, remote sensing products, share, web.

Address all correspondence to: Jorge Sevilla, University of Extremadura, Escuela Politécnica, Department of Technology of Computers and Communications, Hyperspectral Computing Laboratory, Avda. de la Universidad s/n,Cáceres, Spain, 10003; Tel: +34 927 257000 (Dial extension 51662 during voice message); E-mail: jorgesece@unex.es

1 Introduction

In the last years, a significant amount of data from satellite Earth Observation instruments has been made available to the public. Among these, the moderate resolution imaging spectroradiometer (MODIS) and the spinning enhanced visible and infra-red imager (SEVIRI) sensors improve on their predecessors. MODIS is a key instrument aboard the Terra/Aqua satellite[1–4] . Terra's orbit

around the Earth is timed so that it passes from North to South across the Equator around 10:30, while Aqua passes South to North over the Equator around 13.30. Global coverage of the Earth is provided by Terra/Aqua every one to two days. On the other hand, the SEVIRI instrument on board the geostationary platform Meteosat Second Generation (MSG)[5] provides (every 15 minutes) a low resolution scan of the Europe, Africa, the Middle East and the eastern tip of South America. The images have a sub-nadir resolution of 3 km that deteriorates towards the poles, reducing to an average of 5 by 3 km^2 over Europe.

In recent years, many research studies have focused on the calibration and exploitation of data coming from the aforementioned instruments[5–8] . For instance, in[9] an operational algorithm for retrieving the land surface temperature (LST) from SEVIRI data is provided. In[10] , a technique for estimating the sea surface temperature (SST) from SEVIRI data is given. A similar strategy is developed in[11,12] for MODIS data, while[13] assessed the presence of trends in MODIS time series. In[14] , a procedure for water vapor retrieval from SEVIRI observations is presented. Numerous classic studies, such as[15,16] , have also focused on the retrieval of advanced vegetation parameters from remote sensing data such as normalized difference vegetation index (NDVI) or fraction of vegetation cover (FVC), with the ultimate goal of providing advanced products covering relatively large areas[17] . Several other recent applications have exploited MODIS data in a diversity of areas, such as estimation of vegetation phenology[18] , mapping of snow cover[19] , characterization of soil moisture[20] , mapping of impervious cover[21] , agriculture monitoring[22] , analysis of land surface temperature (LST) and environmental factors[23] , deriving water fraction and flood maps[24] , or fire characterization[25] , among many others[26–28] . SEVIRI data has also been recently exploited for phenology estimation[29] , detection of tropical cyclones[30] , monitoring aerosols[31] , or estimating land surface radiation and energy budgets from ground measurements[32] . These studies are possible

thanks to the availability of a wide range of advanced remote sensing products from the data collected by those instruments.

There are several institutions currently distributing multispectral sensors products. On one hand, the NASA Earth Observing System Data and Information System (EOSDIS) provides end-to-end capabilities for managing NASA's Earth science data from various sources (satellites, aircraft, field measurements and various other programs) in which the data is distributed through the *LAADS Web*[33] system. This system provides quick and easy access to MODIS Level 1, Atmosphere and Land data products and Visible Infrared Imaging Radiometer Suit (VIIRS) sensor Level 1 and Land data products. On the other hand, the Land Processes Distributed Active Archive Center (LPDAAC) from USGS distributes mainly land products through several web systems, from which two of them are focused on MODIS products. First, the *USGS EE*[34] tool offers data from the Landsat missions and a variety of other data providers, this tool now provides access to MODIS land data products (from the NASA Terra and Aqua missions) and to ASTER level-1B data products over the U.S. territories (from the NASA ASTER mission). The second web system of USGS is *MRTWeb 2.0*[35] , which is a more complete data system that combines the interface of USGS Global Visualization Viewer[36] and the product database of MRT[37] . In our work, we propose an alternative repository that combines products from two very different resolution sensors: MODIS and SEVIRI, which provide images and products with different formats.

The Calibration of Earth Observation Satellites in Spain (CEOS-SPAIN) project has recently developed a full operational chain for advanced retrieval and processing of MODIS and SEVIRI data acquired in real-time at the Global Change Unit (UCG), Image Processing Laboratory (IPL)[38] , University of Valencia (see Fig. 1). These final products are validated (quality control) using the on-ground data acquired in a variety of test sites. The UCG/IPL participated actively in the

definition of products to be delivered to users from MODIS and SEVIRI data. The resulting remote sensing products are of great interest for the wide remote sensing community. These MODIS products cover Southwestern Europe and most of Northwestern Africa, while SEVIRI products cover half the surface of the Earth. However, the wide availability of these products is difficult due to the massive volume of information that they comprise. In the case of MODIS, the amount of data acquired and processed by UCG/IPL from July 2010 to October 2014 exceeds nine thousand images and two hundred thousand resulting products (more than eleven terabytes of products). In the case of SEVIRI, the amount of data acquired and processed by UCG/IPL from July 2007 to October 2014 comprises more than three hundred thousand images and six hundred thousand resulting products (more than seventeen terabytes of products). This massive amount of information can be extremely useful in climate change studies and in many other studies, specially if the products could be easily retrieved for a given area of interest. In fact, this capability has the potential to impact the activities of many remote sensing users such as farmers, engineers, water managers, as well as the scientific community.

In this paper, we describe a new geo-portal that has been specifically designed to provide efficient access to a large collection of products of MODIS/SEVIRI image acquisitions, which are generated in near real-time at the facilities of UCG/IPL. This work is the first to be published on the reception system at IPL. The geo-portal has been designed in the framework of the CEOS-SPAIN project, which ultimately pursues the development of an integrated system for optical data validation and information extraction, expected to serve as a reference to the remote sensing community through the heterogeneous data distribution and dissemination mechanisms planned for the project. The geo-portal has been designed using advanced computing infrastructure to speed-up the supported data distribution and processing tasks. The geo-portal works as a digital repository
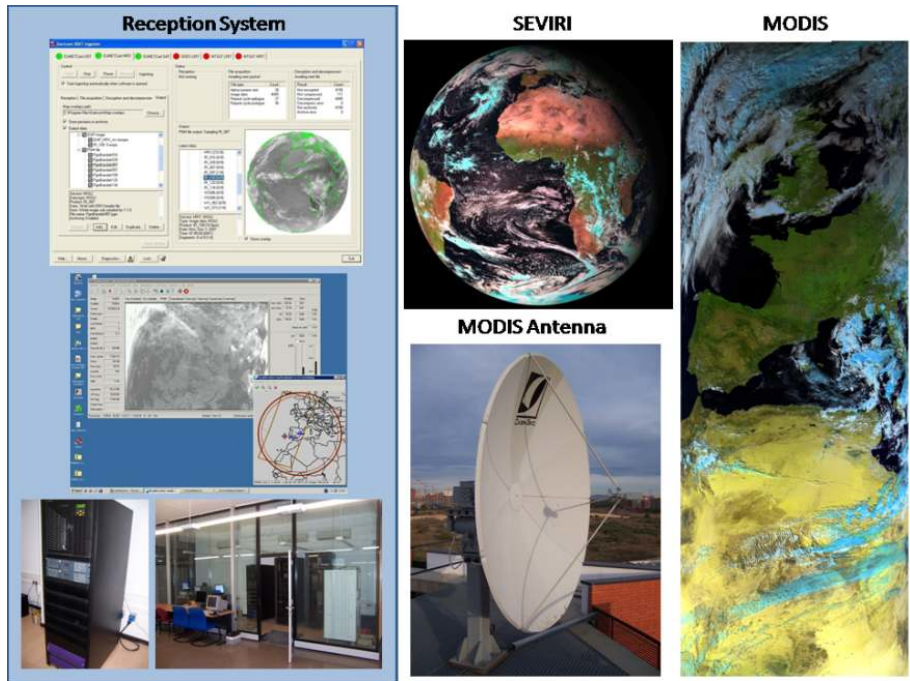
4

Fig 1 MODIS (Terra/Aqua) and SEVIRI (Meteosat Second Generation) real-time reception system at the Global Change Unit (UCG) within the Image Processing Laboratory (IPL) of the University of Valencia.

which has been designed in order to store and efficiently retrieve advanced MODIS/SEVIRI products. Of particular importance are the geolocation-based retrieval functionalities included in the geo-portal, which facilitates searching for remotely sensed scenes (in geolocated form) over a map using Google Maps* services. Although the geo-portal is focused on Southwestern Europe, it is designed to store images from any location. Another feature of our geo-portal is that it provides advanced products, which may be useful to end-users without the possibility to perform advanced processing of the collected remote sensing data. In fact, the idea of using the final products instead of the raw data is quite appealing from the viewpoint of the end-user, who can simply retrieve the processed products after applying different types of indices and perform a more intuitive exploitation of the data.

The proposed geo-portal has been implemented as a web application, available online at `http:`

---

*https://www.google.com/maps

5

`//ceosspain.lpi.uv.es`, and is composed of different layers. Its modular design provides quality of service and scalability, and allows adding and/or modifying components without the need to modify the entire system. These properties result from the fact that the geo-portal has been developed as a digital repository, using advanced software and hardware design techniques with the most advanced tools for such purposes currently available. On the client layer, an intuitive web browser interface provides users with remote access to the system. On the server layer, the system provides advanced data management capabilities. On the storage layer, the system provides a safe massive storage service handling several terabytes of data. Combined, these parts provide a completely open standardized repository of (easily searchable) MODIS/SEVIRI products that is expected to be useful in a variety of remote sensing data exploitation tasks, as originally intended by the CEOS-SPAIN initiative.

The remainder of this work is organized as follows. Section 2 describes the processing chain used to generate the advanced MODIS/SEVIRI products that are distributed through the geo-portal. This chain is applied as soon as the data are collected, and the system automatically stores the resulting products in the geo-portal. Section 3 describes the implementation of the geo-portal, which is composed of three main layers: 1) client layer, which defines the interaction between the user and the system through a web interface; 2) server layer, which efficiently manages the requests coming from end-users; and 3) storage layer, in charge of safe storage and retrieval of remote sensing products. Section 4 presents an experimental validation of the system, with particular attention to its geolocation retrieval accuracy and to the performance of the system in responding to the queries carried out by end-users. Section 5 concludes the paper with some remarks and hints at plausible future research lines.

2 Processing chain

We have implemented a processing chain for MODIS (Terra/Aqua) and SEVIRI (MSG) data acquired in real-time with the antennas located at the UCG/IPL of the University of Valencia. SEVIRI data encompasses the hemisphere centered on the $(0,0)$ latitude/longitude coordinates, while acquired MODIS data consist of Aqua and Terra overpasses within the reach of the antenna located at the IPL, which covers mostly Western Europe and Maghreb. These data allow for the generation of classic and advanced remote sensing products for inter-comparison with other sensor products. The proposed chain has been designed for processing image acquisitions every day, and is applied in near real-time (i.e., soon after the data are collected). The chain is implemented in the interactive data language (IDL)[†]. This section describes succinctly the process carried out over MODIS (subsection 2.2), and then SEVIRI (subsection 2.3) data. Both cases need a previous pre-processing step which is described in subsection 2.1.

2.1 Pre-processing

For both MODIS and SEVIRI sensors, images are acquired as digital counts, which are converted to radiances using the gain and offset values provided in the meta-data associated to the images. Then, the day-time visible and near infra-red channels of the data are atmospherically corrected using SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum[39]. The choice of standard input values for atmospheric correction with SMAC software results of a compromise between atmospheric correction accuracy and instantaneous processing of the received data. By using standard values, the atmospheric corrected values may be slightly for atmosphere type (continental), atmospheric pressure (1013 hPa), aerosol optical thick-

---

[†]http://www.exelisvis.com/ProductsServices/IDL.aspx

ness (0.05) and ozone concentration (0.33 atm.cm), the atmospherically corrected values may be slightly over or underestimated, although the resulting error is generally lower than the error resulting from the use of the top of atmosphere values. This method runs on a pixel by pixel basis, and needs additional information such as (standard values in parentheses): atmosphere type (continental), atmospheric pressure (1013 hPa), aerosol optical thickness (0.05) and ozone concentration (0.33 atm.cm). Additionally, a water vapor product (see below) is needed as an input of the SMAC software. Finally, for thermal infra-red channels, brightness temperatures are estimated by inverting Planck's law:

$$\mathbf{T}_i = \frac{\dfrac{\mathbf{c}_2}{\lambda_{eff}}}{ln\left(\dfrac{\mathbf{c}_1}{\lambda_{eff}{}^5 \cdot \mathbf{B}(\mathbf{T}_i)} + 1\right)}, \tag{1}$$

where $\mathbf{T}_i$ is the brightness temperature in each thermal band, $\mathbf{c}_1 = 1.1491047 \cdot 10^8 W/(m^2 sr \mu m^{-4})$, $\mathbf{c}_2 = 1.4387752 \cdot 10^4 K \cdot \mu m$, $\mathbf{B}(\mathbf{T}_i)$ is the at sensor radiance, expressed in $W/(m^2 sr \mu m)$, and λ_{eff} is the effective wavelength for each band i.

2.2 MODIS products

The complete set of products available in our system is described hereafter. Ancillary data include Latitude, Longitude, Elevation (Height), Sensor Azimuth Angle, Sensor Zenith Angle, Solar Azimuth Angle, Solar Zenith Angle. Several masks can also be downloaded: Land/Sea mask (LSM), day/night mask (Day), cloud mask (Clouds), fire mask (Fire), and snow mask (Snow). These masks have been calculated with MOD35 and MOD14 software [‡]. Advanced products, including Vegetation Condition Index (VCI)[40] and Bidirectional Reflectance Distribution Function (BRDF)[41] corrected NDVI and reflectances for bands 1 to 7, will be included in a close future. Besides, other

---

[‡]http://eostation.scanex.ru/software.html

products are available in our system:

- NDVI[15] is traditionally calculated as $NDVI = (NIR - RED)/(NIR + RED)$, where NIR stands for reflectivity in Near Infra-Red, and RED for reflectivity in red channels, which for MODIS correspond, respectively, to MODIS bands 2 and 1:

$$NDVI = \frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}.$$

(2)

  where ρ_1 and ρ_2 are the at-surface reflectivities obtained from sensor bands located in RED and NIR spectral regions.

- FVC is estimated from the NDVI following[16] for day-time acquisitions, as a normalization of NDVI between standard bare soil and dense vegetation values. In the case of MODIS, these values are 0.15 and 0.90, respectively[17] :

$$FVC = \frac{NDVI - 0.15}{0.90 - 0.15}.$$

(3)

- Emissivities are estimated for day-time acquisitions from FVC and MODIS band 1 information, following the methodology presented in[12] . These emissivities correspond to MODIS thermal bands 31 and 32, and are estimated differently depending on the vegetation proportion within a given pixel. In the case of night-time acquisitions, such method cannot be implemented due to the lack of solar radiation; therefore the emissivity estimates during the previous day are re-projected to a latitude/longitude grid, averaged and re-projected back to the night-time acquisition configuration for further calculations. Due to computational constraints, the re-projection algorithm is limited to the Iberian Peninsula and North West

Table 1 MODIS emissivity values

	ε	$\mathbf{\Delta}\varepsilon$
vegetation: $NDVI > 0.5$	0.99	0
Mixed: $0.2 \leq NDVI \leq 0.5$	$0.971 + 0.018FVC$	$0.006(1 - FVC)$
Bare soil: $NDVI < 0.2$	$0.9832 - 0.058\rho_1$	$0.0018 - 0.060\rho_1$

Africa. Therefore, night products are also limited to this area. As Table 1 shows, the emissivities are expressed as average emissivities ε (for bands 31 and 32) and spectral difference of emissivities $\mathbf{\Delta}\varepsilon$.

- Water vapor is then estimated using the method developed in[12] . This method is based on the attenuation of surface reflected solar radiation and clouds in near infra-red due to water vapor. To this end, water vapor absorbing bands centered at 0.905, 0.936 and $0.94\mu m$ (bands 17, 18 and 19) are used in addition to a water vapor transparent band centred at $0.865\mu m$ (band 2) as follows:

$$\mathbf{W} = 0.192 \cdot \mathbf{W}_{17} + 0.453 \cdot \mathbf{W}_{18} + 0.355 \cdot \mathbf{W}_{19} \tag{4}$$

where:

$\mathbf{W}_{17} = 28.449 \cdot \mathbf{G}_{17}{}^2 - 54.434 \cdot \mathbf{G}_{17} + 26.314,$

$\mathbf{W}_{18} = 27.884 \cdot \mathbf{G}_{18}{}^2 - 23.017 \cdot \mathbf{G}_{18} + 5.012,$

$\mathbf{W}_{19} = 19.914 \cdot \mathbf{G}_{19}{}^2 - 26.887 \cdot \mathbf{G}_{19} + 9.446,$

with:

$$\mathbf{G}_{17} = \mathbf{L}_{17}/\mathbf{L}_2,$$

$$\mathbf{G}_{18} = \mathbf{L}_{18}/\mathbf{L}_2,$$

$$\mathbf{G}_{19} = \mathbf{L}_{19}/\mathbf{L}_2,$$

where $\mathbf{W}$ is the atmospheric water vapor $(g \cdot cm^{-1})$, and $\mathbf{L}_2$, $\mathbf{L}_{17}$, $\mathbf{L}_{18}$ and $\mathbf{L}_{19}$ are the radiance values (RAD-TOA) of MODIS in 2, 17, 18 and 19 bands respectively. In the case of night-time acquisitions, such method cannot be implemented due to the lack of solar radiation, therefore the water vapor estimates during the previous day are re-projected to a latitude/longitude grid, averaged and re-projected back to the night-time acquisition configuration for further calculations. Even though different water vapor values are estimated for any given day, only one water vapor value (daylight average of valid values) is considered per day. Assuming a constant water vapor during 24 hour may lead to over and under-estimations of this parameter, resulting in small differences in final LST. A sensibility analysis shows that an error of $1g \cdot cm^2$ in water vapor leads to a 0.23 K error in LST.

- Land surface temperature (LST) is estimated for day- and night-time acquisitions using the method developed in[12] :

$$LST = \mathbf{T}_{31} + \mathbf{a}_1 + \mathbf{a}_2(\mathbf{T}_{31} - \mathbf{T}_{32}) + \mathbf{a}_3(\mathbf{T}_{31} - \mathbf{T}_{32}) + (\mathbf{a}_4 + \mathbf{a}_5 \cdot \mathbf{W})(1 - \varepsilon) + (\mathbf{a}_6 + \mathbf{a}_7 \cdot \mathbf{W})\delta\varepsilon, \quad (5)$$

where $\mathbf{T}_{31}$ and $\mathbf{T}_{32}$ are brightness temperature for MODIS bands 31 and 32 respectively, ε and $\delta\varepsilon$ are respectively the average emissivity and the spectral emissivity difference for these bands, W is the total amount of water vapor estimated above, and $\mathbf{a}_1 = 1.02, \mathbf{a}_2 = 1.79, \mathbf{a}_3 = 1.20, \mathbf{a}_4 = 34.83, \mathbf{a}_5 = -0.68, \mathbf{a}_6 = -73.27$ and $\mathbf{a}_7 = -5.19$.

Table 2 SEVIRI emissivity values

	ε_{108}	ε_{120}
vegetation: $NDVI > 0.5$	0.99	0.99
Mixed: $0.2 \leq NDVI \leq 0.5$	$0.968 + 0.021 FVC_{seviri}$	$0.976 + 0.015 FVC_{seviri}$
Bare soil: $NDVI < 0.2$	$0.977 - 0.048 Vis06$	$0.981 - 0.026 Vis06$

- Sea surface temperature (SST)[42] is estimated for day- and night-time acquisitions using:

$$SST = \mathbf{T}_{31} + (\mathbf{a}_1 + \mathbf{a}_2 \cdot \mathbf{W})(\mathbf{T}_{31} - \mathbf{T}_{32}) + \mathbf{a}_3 \cdot \mathbf{W} + \mathbf{a}_4, \tag{6}$$

where $\mathbf{T}_{31}$ and $\mathbf{T}_{32}$ are brightness temperature for MODIS bands 31 and 32 respectively, $\mathbf{W}$ is the total amount of water vapor estimated above, and $\mathbf{a}_1 = 1.90, \mathbf{a}_2 = 0.44, \mathbf{a}_3 = 0.05$ and $\mathbf{a}_4 = 0.34$.

2.3 SEVIRI products

The SEVIRI products available in our processing chain can be summarized as follows:

- NDVI is retrieved as in the case of MODIS data, by using red (visible band at 0.6 microns) and near-infra-red (visible band 0.8 at microns) information.

- Emissivities are estimated for day-time acquisitions from NDVI and red information at 0.6 microns ($Vis06$), following the threshold method. These emissivities correspond to SEVIRI infrared $\mathbf{T}_{108}$ and $\mathbf{T}_{120}$ spectral bands centred at 10.8, and 12.0 μm, and are estimated differently depending on the vegetation proportion (PV) within a given pixel. These emissivities are expressed as ε_{108} and ε_{120}. However, as Table 2 shows the coefficients for emissivity estimation for different NDVI ranges of value:

where $FVC_{seviri} = \dfrac{(NDVI - 0.2)^2}{0.009}$.

- Water vapor is then estimated using the method developed by[12] . This method is based on the difference in temperature between two different acquisitions ($\mathbf{T}^A$ and $\mathbf{T}^B$) for both infrared channels ($\mathbf{T}_{108}$ and $\mathbf{T}_{120}$) centred at 10.8, and 12.0 μm, and depends of the observation zenith angle (θ). This temperature difference should be greater than 10 K for the method to be accurate:

$$\mathbf{W} = \mathbf{a} \cdot arg^2 + \mathbf{b} \cdot \beta + \mathbf{c}, \tag{7}$$

where $\mathbf{a} = -15.1 \cdot sec\theta + 5.1$, $\mathbf{b} = 16.4 \cdot sec\theta - 2.8$, $\mathbf{c} = 0.336 \cdot sec\theta - 0.117$, and $\beta = \dfrac{1}{sec\theta} ln \left(\dfrac{\mathbf{T}_{108}^A - \mathbf{T}_{108}^B}{\mathbf{T}_{120}^A - \mathbf{T}_{120}^B} \right)$. However, this method provides only a few values per day, which can lead to errors in the case of rapidly evolving weather conditions. To address this issue, methods for instantaneous estimation of water vapor are currently being investigated[43] .

- LST is estimated for day- and night-time acquisitions using the method developed in[14] as follows:

$$LST = \mathbf{T}_{108}$$

$$+ \left[1.34 - \frac{0.11}{cos^2\theta}\right] \cdot (\mathbf{T}_{108} - \mathbf{T}_{120})$$

$$+ \left[0.29 + \frac{0.08}{cos^2\theta}\right] \cdot (\mathbf{T}_{108} - \mathbf{T}_{120})^2$$

$$+ \left[60.67 - \frac{10.01}{cos^2\theta}\right] \cdot (1 - \varepsilon)$$

$$\text{(8)}$$

$$+ \left[-6.71 + \frac{2.47}{cos^2\theta}\right] \cdot \mathbf{W} \cdot (1 - \varepsilon)$$

$$+ \left[-125.91 + \frac{15.09}{cos^2\theta}\right] \cdot \mathbf{\Delta}\varepsilon$$

$$+ \left[19.44 - \frac{4.27}{cos^2\theta}\right] \cdot \mathbf{W} \cdot \mathbf{\Delta}\varepsilon$$

$$+ \left[-0.44 + \frac{0.57}{cos^2\theta}\right],$$

where $\mathbf{T}_{108}$ and $\mathbf{T}_{120}$ are brightness temperature for the infra-red bands centred at 10.8 and 12.0 μm (Ir108 and Ir120) respectively, ε and $\mathbf{\Delta}\varepsilon$ are respectively the average emissivity and the spectral emissivity difference for these bands, W is the total amount of water vapor estimated above, and θ is the observation zenith angle.

- SST is estimated for day- and night-time acquisitions using the method developed by[10] :

$$SST = \mathbf{T}_{108} + (0.99 \cdot cos\theta + 0.21) \cdot (\mathbf{T}_{108} - \mathbf{T}_{120}) + \left(\frac{0.364}{cos\theta} + 0.15\right) \cdot (\mathbf{T}_{108} - \mathbf{T}_{120}) + \frac{0.327}{cos\theta} + 0.11,$$

$$\text{(9)}$$

where $\mathbf{T}_{108}$ and $\mathbf{T}_{120}$ are the brightness temperatures for MSG bands Ir108 and Ir120, re-

spectively.

3 Geo-portal

This section is devoted to the description of our proposed geo-portal system. In subsection 3.1, we describe the system architecture of the geo-portal and the different layers that compose it. In subsection 3.2, we describe the structure of the database that stores the associated meta-data of the MODIS/SEVIRI scenes and their products. Subsection 3.3 describes how the queries to the database are performed in the system. Finally, subsection 3.4 describes the procedure used by the geo-portal to retrieve and provide advanced remote sensing scenes and their products to end-users.

3.1 System architecture

As shown in Fig. 2, the architecture of the proposed system is composed by different layers, which can be defined by their roles. The system follows a modular design in which the communication between layers is performed using standard data exchange formats and transfer protocols, so that any layer can be easily modified and/or enhanced as long as its connectivity with the remaining layers of the system is maintained. More specifically, our design has been carried out using free software tools such as `Symfony2`[§], a full-stack web framework, while the adopted format for data exchange is JavaScript Object Notation (`JSON`[¶]), an open standard format that uses human readable text to transmit data objects. In the following, we describe the different layers that compose the geo-portal.
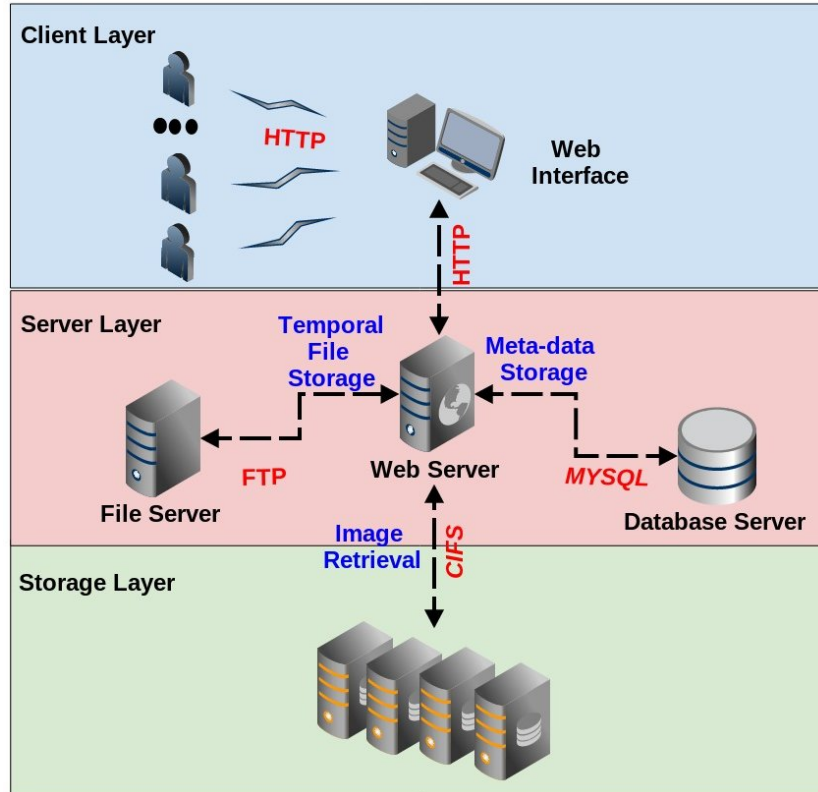

[§]http://symfony.com
[¶]http://www.json.org

15

Fig 2 Architecture of the proposed geo-portal.

3.1.1 Client layer

This layer defines the interactions between the users (through an internet web browser) and our geo-portal, and is responsible for providing users with remote and interactive access to the system. The web interface has been designed using HTML5[‖], which is the most widely used web programming language, and CSS3[**], which is a standard style language for improving the HTML5 web appearance. In addition, we also used jQuery[††], which is a JavaScript library for controlling the web behaviour. The interaction between the user and the web interface is captured by the event handlers of the jQuery libraries. The web interface transmits the requests to the server layer via HTTP[*], an application protocol for distributed, collaborative, hypermedia information systems and

---

[‖]http://www.w3.org/TR/html51
[**]http://www.w3.org/Style/CSS
[††]http://jquery.com
[*]http://www.w3.org/Protocols/

16

the foundation of data communication for the World Wide Web. Most of the views are actually generated in the server, using `Symfony2`, a robust web development framework.

3.1.2 Server layer

The system is implemented as a web application, in the sense that it provides remote access to a set of clients (i.e., the end-users of our system) through internet connexions, so that web pages are delivered to the clients while user requests are received and immediately processed. An advantage of this approach is that no additional software has to be installed on the client node, since only a web browser is required for interacting with the geo-portal. The services provided by the system are managed and executed on the server layer, which is composed of several elements with different roles. As Fig. 2 shows, the web server handles web interface requests (via `HTTP`) and manages the system resources, such as meta-data and file data storage. The meta-data are associated to the image scene and their products which are used for efficient data retrieval as it will be explained in subsequent sections. The server layer can be considered as the main engine of the system since it is in charge of managing and connecting the different components that integrate it. The server layer is also in charge of storing image and product meta-data, following a database scheme that is described in subsection 3.2. For the management of the database we have selected `Mysql`[‡‡], an open-source database manager which provides all the necessary requirements for our system, such as stored functions and procedures, fast queries, and low computational cost. This database manager is widely used for building high performance webs. The `Mysql` database manager allows for the efficient execution of stored procedures directly, which is needed in our context in order to deal with the high computational cost of executing a matching algorithm for data retrieval involving

---

[‡‡]http://www.mysql.com

several terabytes of data. On the other hand, we also used a `FTP`* file storage server in this layer for providing remote user access to non-remotely accessible storage layer files. This module is ultimately in charge of providing a temporal file storage of the files related to the user download requests. The files are automatically deleted once a given limit time (parametrized value) is expired. This strategy allows for efficient downloads of the products retrieved by the geo-portal after the end-user queries are completed.

3.1.3 Storage layer

The proposed geo-portal has been specifically designed bearing in mind the need to manage and share a large amount of remote sensing data products obtained after applying the MODIS/SEVIRI processing chain described in section 2. The hardware resources utilized in the storage layer comprise a set of `DELL NAS`* processing nodes. First, a `DELL PowerVault NX3200`* system is dedicated to storing the MODIS related products. It consists of a chassis with up to 12 hard `SAS`* drives, and an `Intel Xeon E5-2609`* CPU with 32 GB of RAM memory. Security and high availability is provided by a `PERC H710`* integrated RAID controller which combines multiple hard disks into a logical unit for storage purposes. The overall storage capacity comprises 30 terabytes, using RAID 5 and a hot spare unit. On the other hand, a second `DELL` node is dedicated to storing SEVIRI products. It consists of another `DELL PowerVault NX3200` connected to a `DELL PowerVault MD1200`* which provides 24 terabytes of additional storage capacity, for a total of 54 terabytes capacity for the SEVIRI products. Both the `PowerVault MD1200` and the

---

*http://www.w3.org/Protocols/rfc959/
*http://www.dell.com/us/business/p/network-file-storage
*http://www.dellstorage.com/WorkArea/DownloadAsset.aspx?id=2966
*http://www.dell.com/learn/us/en/04/campaigns/dell-hard-drives
*http://ark.intel.com/products/64588/intel-xeon-processor-e5
*http://www.dell.com/downloads/global/products/pvaul/en/perc-technical-guidebook.pdf
*http://www.dell.com/us/business/p/powervault-md1200/pd

`PowerVault NX3200` are connected through `PERC H800`* adapters. As mentioned before, the main purpose of this layer is to provide advanced storage capabilities, so all interactions are performed through other layers of the system that access the this layer directly. The storage layer and the server layer are located inside a private network which is remotely non-accessible. Both layers are connected by using the common internet file system (`CIFS`)*, a protocol that defines a standard for remote file access allowing up to millions of connections at the same time.

3.2 Database Structure

Fig. 3 illustrates the database scheme used to store remote sensing products in the proposed geo-portal. The database has been designed in order to store information not only from the products, but also from the original remote sensing data (although in our current version only the products are available through the geo-portal). The information stored in the database for each image entry comprises the number of samples, lines, bands, data type, byte order, wavelength information, interleave, thumbnails, geolocation and sensor used (limited to MODIS/SEVIRI in the present version but fully extensible to other sensor data in future developments). Most importantly, our system also contains information about the presence of cloud areas (called *cloudinesss* hereinafter) in the scene. Clouds were identified through IMAPP[44] Level2 Cloud Mask (MOD35) software. In order to minimize the database size, the geolocation data is confined to the four extreme image coordinates of each data set. In practice, each image is decomposed into thirty parts (called quarters) and we store geolocation and cloudiness information for each quarter. In addition, relevant information about the data products are also stored in the geo-portal, including type of product, resolution, sensor of acquisition, etc. All this information can be used for retrieval purposes using different

---

*http://www.dell.com/downloads/global/products/pvaul/en/perc-technical-guidebook.pdf
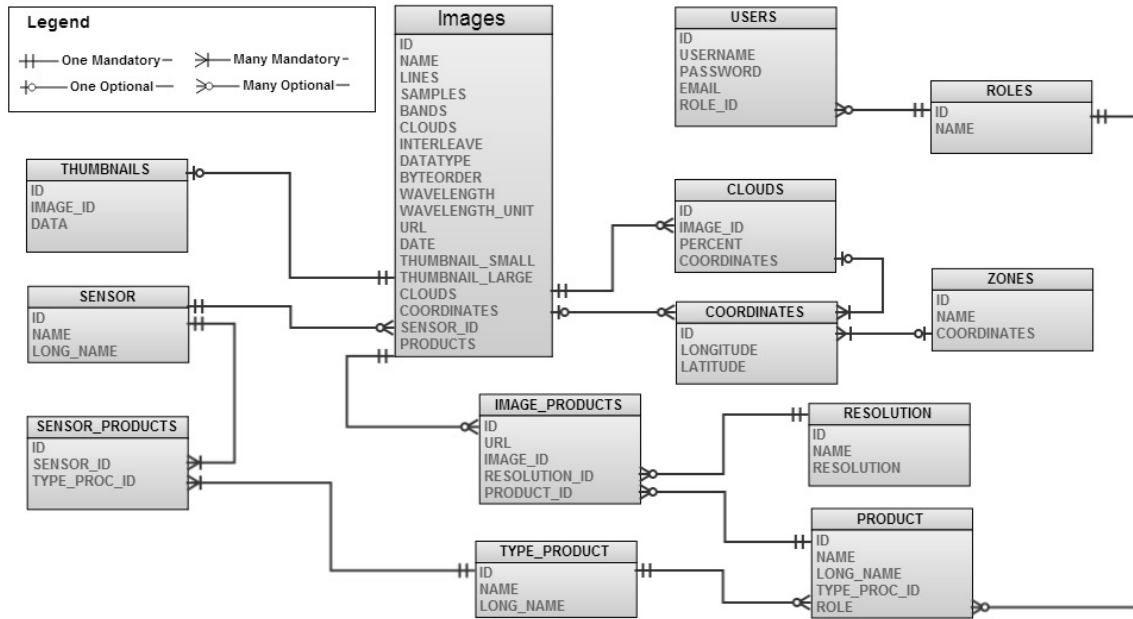*http://www.samba.org/samba/docs/man/Samba4-HOWTO/protocol.html

Fig 3 Structure of the database used to store data and products in the proposed geo-portal.

queries. At this point, we reiterate that the database structure has been designed to store not only the final products but also the raw data. Although in the current version of the geo-portal we only make available the final products to external users, the system is prepared to provide also other different data levels (including the raw data). In the following we describe widely the database tables design.

3.3 Queries

Although the aim of this work is to offer remote sensing products, the retrieval system is designed to retrieve remote sensing scenes which have associated products. Thus, the queries are performed using the image associated meta-data, which is the fastest procedure since all the products of an image have the same query parameters.

In this subsection we outline the queries that can be performed to retrieve remote sensing products in the proposed geo-portal system. The geo-portal includes an advanced retrieval system

which provides high query performance. The query filters include acquisition date, used sensor, geolocation or cloudiness. It should be noted that the last two filters are only available for MODIS images due to the geostationary character of the MSG platform, geolocation information for SE-VIRI data needs not being included within each image. As for cloud information for SEVIRI data, it is still not available, although we plan to include this information in the geo-portal in a near future. In the following, we first describe the searching methodology, which relies on a newly developed geolocation matching algorithm. Then, we describe the procedure used to perform the search from an end-user's point of view (with particular emphasis on the developed web browser interface and the searching options available in the system). Finally, we describe the procedure used by the system to allow for remote downloads of data from end-users.

3.3.1 Searching methodology

In order to deal with the time needed to perform image retrieval on a huge database such as the one available in our geo-portal, the searching functionality has been designed in a way that the queries are completely executed on the database manager engine. This is a very important aspect, since the possibility to perform the queries directly in `Mysql` allows for much increased performance as the queries are executed as simple `SQL` calls embedded in the database manager. In previous versions of the geo-portal we intended to use an external programming language such as C++ in order to perform the queries, which resulted in a much slower response time. As a result, the embedding of the queries into `Mysql` is considered to be one of the most attractive features of the proposed geo-portal system as this design choice guarantees a high quality of service. The interest region-based searches are implemented by means of a newly developed geolocation matching algorithm, which works as a stored procedure in the database to resolve geolocation-based queries. This algorithm

is particularly useful for the retrievals based on the MODIS sensor, since this instrument provides images with irregular geolocation that makes it difficult to link the data with a set of given geo-coordinates. This introduced significant difficulties since our idea was to visualize and retrieve the obtained products in interactive fashion, based on available map services such as Google Maps. In order to address this issue, we have developed a fast matching algorithm which compares two geographic regions using their four extreme coordinates. This methodology is summarized in Algorithm 1.

Algorithm 1 Geolocation matching algorithm

```
1:  procedure GELOCATIONMATCHING(IMAGES,selectedPoints,maxCloudiness)
2:      vectorResults= new Vector
3:      for all IMAGES do
4:          imagePoints = image.Points
5:          resultIn = false
6:          if (No selectedPoints) OR PLANEINTERSECTION(imagePoints, selectedPoints) OR ANYPOINTINPOLYGON(imagePoints, selected-
    Points) then
7:              resultIn=true
8:              if (maxCloudiness) then
9:                  quarters=image.quarters
10:                     for all quarters do
11:                         quarterPoints=quarter.Points
12:                         if (PLANEINTERSECTION(quarterPoints, selectedPoints) OR ANYPOINTINPOLYGON(quarterPoints, selectedPoints)) then
13:                             if quarter.cloudPercent > maxCloudiness then
14:                                 resultIn=false
15:                                 Break For loop
16:                             end if
17:                         end if
18:                     end for
19:              end if
20:          end if
21:          if resultIn then
22:              vectorResults.push(image)
23:          end if
24:      end for
25:      return vectorResults;
26: end procedure
```

As shown by Algorithm 1, the main idea of the proposed geolocation procedure is to establish the overlapping between a user-defined region and the corresponding image region. In other words, the user interactively defines an area of interest over a map service (in our case, by drawing a polygon in Google Maps) and the system should retrieve all the stored products of the images that have some area overlapping with the region defined by the user in interactive fashion. The matching pro-

Algorithm 2 Matching procedure used by the geolocation matching algorithm

1: **procedure** POINTINSIDE(polygonPoints,testx,testy)
2:     NUMPOINTS=4
3:     out=0
4:     **if** testx < MIN(polygonP.x) OR testx > MAX(polygonP.x) OR testy < MIN(polygonP.y) OR testy > MAX(polygon.y) **then**
5:         j=NUMPOINTS-1
6:         **for** i = 0 to NUMPOINTS **do**
7:             z = polygonP[j].x - polygonP[i].x* (testy - polygonP[i].y) / (polygonP[j].y - polygonP[i].y) + polygonP[i].x
8:             **if** (polygonP[i].y > testy =! polygonP.y[j] > testy) AND (testx < z) **then**
9:                 c =! c
10:             **end if**
11:             j=i
12:         **end for**
13:     **end if**
14:     **return** out
15: **end procedure**
16: **procedure** ANYPOINTINPOLYGON(imagePoints,selectedPoints)
17:     NUMPOINTS = 4
18:     **for** i = 0 to NUMPOINTS-1 **do**
19:         point = imagePoints[i]
20:         **if** POINTINSIDE(selectedPoints,point.x,point.y) **then**
21:             **return** true
22:         **end if**
23:     **end for**
24:     **for** i = 0 to NUMPOINTS-1 **do** point = selectPoints[i]
25:         **if** POINTINSIDE(imagePoints,point.x,point.y) **then**
26:             **return** true
27:         **end if**
28:     **end for**
29:     **return** false
30: **end procedure**
31: **procedure** PLANEINTERSECTION(imagePoints,selectedPoints)
32:     $NUMPOINTS = 4$
33:     $j = NUMPOINTS - 1$
34:     **for** i = 0 to NUMPOINTS-1 **do**
35:         **if** $(Line(imagePoints[j], imagePoints[i]) \cap Line(selectedPoints[0], selectedPoints[1]))$ OR
36: $(Line(imagePoints[j], imagePoints[i]) \cap Line(selectedPoints[1], selectedPoints[2]))$ OR
37: $(Line(imagePoints[j], imagePoints[i]) \cap Line(selectedPoints[2], selectedPoints[3]))$ OR
38: $(Line(imagePoints[j], imagePoints[i]) \cap Line(selectedPoints[3], selectedPoints[0]))$ **then**
39:             **return** true
40:         **else**
41:             j=i
42:         **end if**
43:     **end for**
44:     **return** false
45: **end procedure**

cedure is described in detail in Algorithm 2. As shown by this algorithm, all comparisons are based on the four extreme coordinates for the user-defined region and for the associated quarters. In order to manage the computational cost of the geolocation procedure, this technique has been designed in the form of an efficient stored procedure implementation, which exploits the high performance of `Mysql` stored procedure engine and provides low latency in the associated complex queries. At this point, it is important to reiterate that the geographic matching of the regions/quarters is performed using geometric concepts which are based on interpreting the coordinates as points and the regions as planes formed by those points, as described in detail in Algorithm 2. In order to cover all the matching possibilities, we first evaluate if there is any intersection of lines from both planes and, if this is not the case, we evaluate if any point from either plane is inside the other plane by using a point-in-polygon approach developed in[45] . As described in Algorithms 1 and 2, the geolocation matching technique executes different queries based on different filter inputs, which can be simply summarized as follows:

- *No filters:* all the scenes from the same sensor are retrieved.

- *Interest region filter:* only those scenes in which its geographic region matches with the interest region are retrieved.

- *Cloudiness threshold filter:* only those scenes in which all quarters contain cloudiness below the threshold are retrieved.

- *Cloudiness threshold and interest region filters:* only those scenes in which all quarters within the region of interest contain cloudiness below the threshold are retrieved.

3.3.2 Searching Procedure from End-User Point of View

In this subsection we outline how the searching procedure is conducted from an end-user's point of view. Fig. 4 shows the browser interface of the geo-portal, which allows the user to select several filters using a friendly interface. As shown by Fig. 4, the leftmost panel of the browser allows specifying the type of sensor for the retrieval of associated image products, as well as the maximum percentage of clouds allowed in the data to be retrieved. The leftmost panel also allows for the retrieval of image products by date (or range of dates) and by geographic coordinates. On the other hand, as Fig. 4 shows the main panel of the browser includes a map obtained from Google Maps service. The map panel is fully interactive, and allows the user to draw a polygon directly on the map so as to perform the retrieval only in the geographical area delimited by the polygon. In order to achieve this functionality, the geolocation matching technique described in Algorithm 1 and its associated matching procedures, described in Algorithm 2, are used. In the following, we summarize the different steps that a simple query involves:

- First, the user can define simple criteria for the query, such as sensor or dates of acquisition, in the leftmost panel of the browser as indicated in Fig. 4.

- In a second step, the user can refine the query by drawing a polygon in the map panel. The geolocation matching technique will be used to retrieve only the images located in a geographic area that has some overlapping with the polygon. The system provides other additional options for selecting the region of interest coordinates, such as simple input form or a set of pre-established zones, as indicated in the leftmost panel (see Fig. 4).

- As an optional step, the system allows defining a maximum cloudiness filter, which is used to restrict images with low visibility due to cloud contamination. After the user clicks the
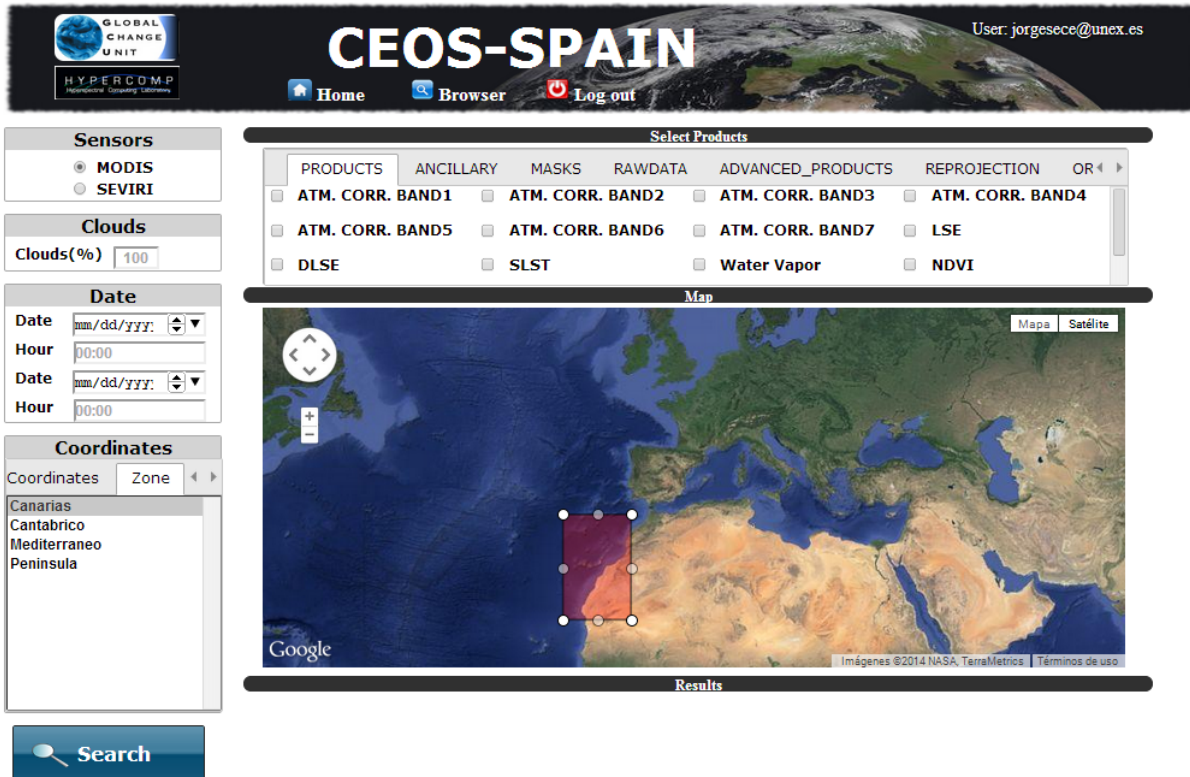
25

Fig 4 Browser interface of the developed geo-portal.

search button, all the images matching the aforementioned searching criteria are retrieved

and the user can now download the desired image products. As Fig. 5 shows, the system

retrieves, in a table, a list of the image acquisitions together with their associated products;

this table is fully interactive and allows to browse among the image meta-data and products.

In the following subsection we explain how the downloads (which can involve multiple users

accessing concurrently the database) have been implemented in our geo-portal.

3.4 Downloading

The developed geo-portal has been designed to facilitate the access to the products which are

required by users. Once the scenes of interest have been retrieved using the searching procedure

described in the previous subsection, the user can select and download the products associated
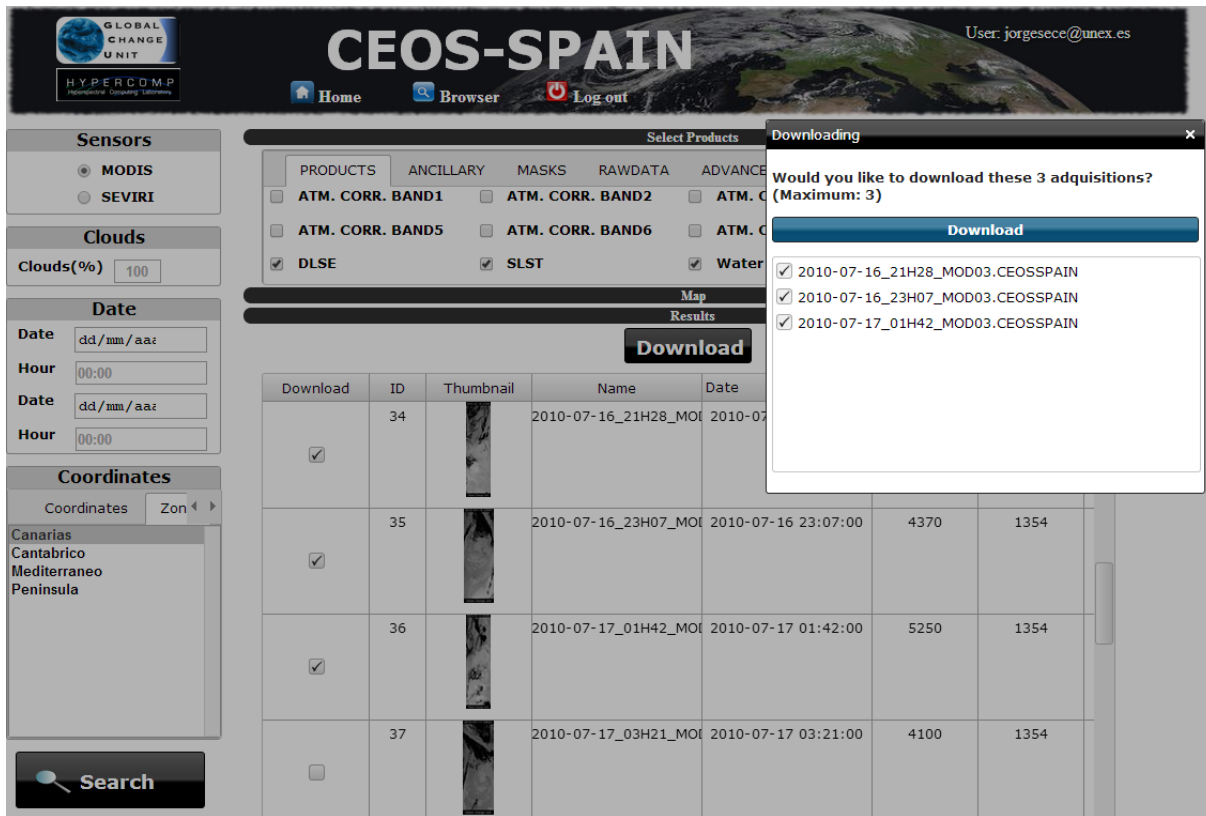
Fig 5 Search results retrieved by the browser interface.

to the image matching the search criteria. In this regard, the geo-portal offers the capability to browse among the products of the images resulting from the filtering procedure and to select them individually for download. In addition, as Fig. 5 shows, the main panel of the browser includes a list of the available products which offers the capability to download the selected products of all the images selected in the table. A file storage server (based on the `FTP` protocol) is used for downloading purposes, so that the server creates a folder with all the requested products that are available for the user to download. The selected products will be deleted from the server after a number of days (variable value). In this way, once the user executes the download option by clicking the corresponding button in the browser (see Fig. 5), the system creates the temporal folder, copies the selected products to the folder and shows a pop-up message to the user together with the `FTP` download link. As mentioned before, the system has been implemented so as to

support a large number of simultaneous connections and concurrent data downloads, being able to execute the queries with high performance and with significant quality of service. In the following section we provide an experimental evaluation of the system in terms of geolocation accuracy and computational efficiency in resolving the incoming queries.

4 Experimental results

In this section, we evaluate the proposed geo-portal from two different perspectives: the efficiency in the retrieval of data products and the accuracy of the proposed geolocation matching algorithm. The section is organized as follows. In subsection 4.1, we describe the MODIS/SEVIRI data sets and products used for our experimental validation. In subsection 4.2, we illustrate the accuracy of the proposed geolocation matching algorithm used to retrieve products from our geo-portal. Finally, subsection 4.3 performs an evaluation of the computational performance and response times measured in the retrieval process.

4.1 Data and products

The experiments have been carried out using the full database of MODIS/SEVIRI images which have been processed (using the methodology described in section 2) at UCG/IPL from July 2010 to January 2014 (in the case of MODIS) and from July 2007 to January 2012 (in the case of SEVIRI). This comprises a total of 7626 scenes and 165406 products of MODIS (occupying a total space of more than seven terabytes of products) and a total of 163813 scenes and 313345 products of SEVIRI (occupying a total space of more than ten terabytes of products). For the MODIS data, it should be noted that the MODIS instrument provides images from (plus/minus) 55-degree scanning pattern at the EOS orbit of 705 km, thus achieving a 2330-km swath. Thus, the features of these

images are different in every acquisition. This includes width, height or geolocation. In this way, two images from the same area can have different dimensions (width of 1354 plus/minus 10 pixels and height of 5000 plus/minus 200 pixels). This aspect complicated the geolocation process for these images, as indicated before. In order to illustrate this issue further, the following subsection provides an evaluation of the geolocation retrieval accuracy in the case of MODIS data.

4.2 Geolocation retrieval accuracy

The accuracy of the proposed geolocation-based retrieval system has been evaluated using a case study based on a query centred in the Canary Islands (35 degrees north latitude, minus 9 degrees south latitude, 20 degrees east longitude and minus 20 degrees west longitude), as indicated in the user-drawn polygon depicted in Fig. 4. For illustrative purposes, the query was performed for a specific day (July 12, 2010), in which five different MODIS images were returned. For simplicity, we did not apply the cloudiness threshold in this experiment to remain focused on the evaluation of the geolocation results. For illustrative purposes, Fig. 6 graphically illustrates the geolocation accuracy achieved by our system in this particular experiment. More specifically, in Fig. 6 the red polygons represent a region of interest defined by the user (see Fig. 4), and the blue polygons represent the available MODIS images collected at different times on July 12, 2010. The intersection areas are displayed in cyan color. As shown by Fig. 6, three different images exhibited some degree of overlapping with the user-defined area. Note that the images acquired by MODIS at 01h24m and 12h29m did not overlap at all with the user-defined area, hence these images were not retrieved by the geo-portal. In turn, the image acquired by MODIS at 03h02m, 10h48m and 21h52m exhibited some (partial) overlapping with the user-defined search area, thus being retrieved. We conducted several additional experiments using a larger number of involved
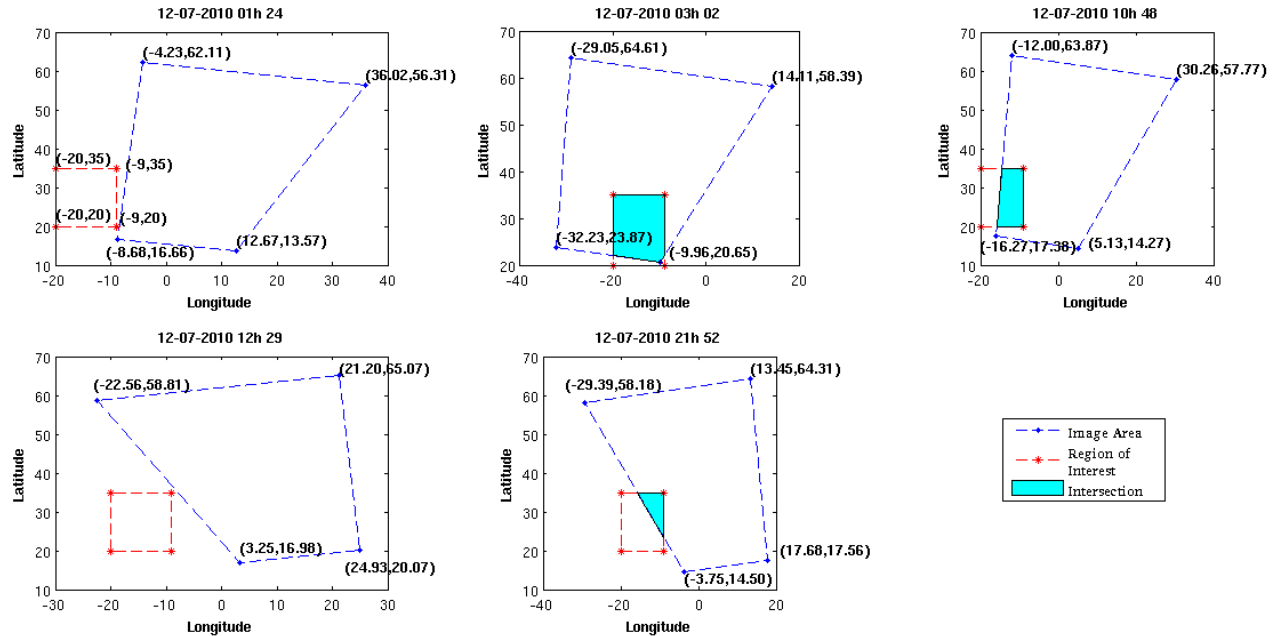
Fig 6 A case study illustrating the geolocation matching accuracy of the developed geo-portal. The red polygons represent a region of interest defined by the user (centered over the Canary Islands, as depicted in Fig. 4) and the and blue polygons represent the available MODIS images at different hours on July 12, 2010. The intersection areas are displayed in cyan color. In our experiment, only the cases in which there was overlapping between the user-defined area and the available images resulted in data retrievals, which confirmed the good performance of the geolocation matching algorithm.

images and user-defined searching areas with the same accuracy results, which confirmed the high precision of the geolocation matching algorithms developed for the geo-portal.

4.3 Efficiency in Retrieval

In this subsection we evaluate the retrieval performance of the proposed geo-portal. For this purpose, we considered three different types of queries and measured the time in seconds from the query execution to the query completion, i.e. the effective time that the geo-portal took in order to provide the data solicited by the user for download purposes. Specifically, the types of queries considered in our tests were the following ones:

- In the first type of query, we used a *no filter* scenario so that all the images from the same sensor could be retrieved. We performed this query for all the MODIS data and also for all

Table 3 Time (in seconds) that the geo-portal took to complete several types of queries applying different filters when exploring the whole database of MODIS/SEVIRI images. The images are filtered by date from July 2010 to January 2014 (in the case of MODIS) and from July 2007 to January 2012 (in the case of SEVIRI). In the queries based on a region of interest, we used an area centred over the Canary Islands depicted in Fig. 4. The cloudiness threshold considered in experiments was 80%, so that we filtered out the images containing more clouds than established in this threshold.

	Only Date filter (SEVIRI images)	Only Date filter (MODIS images)	Region of interest (MODIS images)	Cloudiness and region of interest (MODIS images)
Images retrieved	163813	7626	5153	2592
Time (seconds)	0.008	0.008	9.101	37.078

the SEVIRI data, so that we can evaluate the performance of the system in retrieving all the data stored in the system by sensor type.

- In the second type of query, we analyzed the performance of the *interest region filter*. In this case, we used the same case study illustrated in the previous subsection, in which a query centered in the Canary Islands (35 degrees north latitude, minus 9 degrees south latitude, 20 degrees east longitude and minus 20 degrees west longitude) was used after a user-defined region was established in the map panel of the geo-portal browser as depicted in Fig. 4.

- The third type of query analyzed the *cloudiness threshold filter* in combination with the *interest region filter*. This means that, for the considered user-defined region of interest centered in the Canary Islands and depicted in Fig. 4, we retained only those images that contain less than 80% cloudiness and discarded the rest. In all cases, we avoided establishing a filter for the retrieval of images based on the date, in order to have a complete exploration of the whole database in multi-temporal fashion.

As shown by Table 3, the retrieval times were quite fast for the three types of considered queries. The fastest cases correspond to the retrieval of all MODIS images and all SEVIRI images, which took less than one second to be completed. This is because the only filter applied

31

in the query was the one given by the sensor type, which allowed for a very fast retrieval of the images corresponding to each type of sensor. In any event, given the very high number of images retrieved (7626 for MODIS and 163813 for SEVIRI), we consider that the retrieval times were extremely fast. This indicates that the architecture of the proposed system can handle large volumes of data while providing very fast response times for the queries, implemented directly in the `Mysql` database manager. When the region of interest was defined, the query based on the considered region centered around the Canary Islands returned 5153 MODIS scenes, which is also a considerable amount of information. In this case, the geo-portal could provide the result of the query in approximately 9 seconds. Finally, the most complex query executed was completed in approximately 37 seconds. Such response time is due to the fact that the cloudiness threshold was applied in combined fashion with the region of interest filter, which means 30 combined comparisons for each image, since the cloudiness into the image is split into 30 quarters. Thus, this query is more complex in terms of execution than the previously considered ones. Still, the retrieval time in this case is considered to be quite appropriate bearing in mind that the query took as input the whole MODIS image database which comprises approximately eight terabytes of data. At this point we reiterate that the queries retrieve the images and also their products, no product-related filters are available in the searching process, since all the products of an image have the same query parameters.

5 Conclusions and future work

In this paper, an on-line geo-portal has been developed to manage a standardized data repository for multispectral images and their advanced products, which supports dissimilar sensors in order to provide a wide perspective of the studied areas. A main characteristic of our geo-portal is that

it provides access to processed final products. This eliminates the need to perform the processing of the data offline by integrating the data processing into the geo-portal. The data processing is accomplished in real-time, i.e. as soon as the MODIS/SEVIRI data are collected by the receiving antenna, our system obtains the resulting products and stores them in our database, which also handles the raw data. This makes the final product immediately available to the wide remote sensing community and allows for a detailed exploration of the high-level products obtained after the data processing.

The geo-portal has been designed in a modular way and implemented using advanced software and hardware infrastructures, allowing for the fast execution of queries embedded into our database engine as stored procedures. Most importantly, the geo-portal includes the possibility to perform geolocation-based product retrieval by defining regions of interest (and applying several types of filters) in a user-friendly browser that allows for a complete interaction with a map server such as Google Maps. Here, the map server is used to perform interactive based queries on user-defined regions of interest. These regions are very easy to define, and the user has the possibility to select an area and visualize the overlapping products available in the system in a very intuitive manner, which enhances the usability of the system.

In terms of storage and retrieval capacity, the geo-portal currently provides impressive features: more than nine thousand MODIS acquisitions and two hundred thousand resulting products (products occupy a total space of more than eleven terabytes of data) and more than three hundred thousand SEVIRI acquisitions and six hundred thousand resulting products (products occupy a total space of more than seventeen terabytes of data). These products can be retrieved in a very fast manner and downloaded for further analyses. It should be noted that the amount of data available for download keeps growing as new acquisitions are processed and added to the database. This

results in a massive amount of information that is now fully available to the remote sensing community through the developed geo-portal. As future work, we plan to include images and products collected from other sensors, in addition to additional MODIS/SEVIRI products. We are also planning on integrating additional content-based image retrieval functionalities into the geo-portal, in the vein of the developments presented recently in[46] .

References

1 X.-X. Xiong et al., "Overview of NASA Earth Observing Systems Terra and Aqua moderate resolution imaging spectroradiometer instrument calibration algorithms and on-orbit performance", *Journal of Applied Remote Sensing* 3(1),032501-032501-25, (2009).

2 G. Toller et al., "Status of earth observing system Terra and Aqua moderate-resolution imaging spectroradiometer level 1B algorithm", *Journal of Applied Remote Sensing* 2(1),023505-023505-19 (2008).

3 B.-C. Gao, R.-R. Li and X.-X. Xiong, "Possible enhancement to specifications of several Moderate Resolution Imaging Spectroradiometer visible and near-IR channels for improved earth observations", *Journal of Applied Remote Sensing* 1(1), 013504-013504-7 (2007).

4 L. Wang et al., "Impact assessment of Aqua MODIS band-to-band misregistration on snow index", *Journal of Applied Remote Sensing* 1(1), 013531-013531-11 (2007).

5 F. Pasternak, P. Hollier, and J. Jouan, "SEVIRI, the new imager for meteosat second generation," *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* 3, 1094-1099 (1993)

6 X. Xiong, N. Che, and W. Barnes, "Terra MODIS on-orbit spatial characterization and performance," *IEEE Transactions on Geoscience and Remote Sensing* 43, 355-365 (2005).

7 S. Platnick et al., "The MODIS cloud products: algorithms and examples from Terra," *IEEE Transactions on Geoscience and Remote Sensing* 41, 459-473 (2003).

8 H. Shen, C. Zeng, and L. Zhang, "Recovering reflectance of AQUA MODIS band 6 based on within-class local fitting," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 185-192 (2011).

9 J.-C. Jiménez-Muñoz et al., "Temperature and emissivity separation from MSG/SEVIRI data," *IEEE Transactions on Geoscience and Remote Sensing*, 1-15 (2014, in press).

10 M. Romaguera, J. A. Sobrino, and F. S. Olesen, "Estimation of sea surface temperature from SEVIRI data: algorithm testing and comparison with AVHRR products," *International Journal of Remote Sensing* 27(22), 5081-5086 (2006).

11 M. Atitar and J. A. Sobrino, "A split-window algorithm for estimating LST from Meteosat 9 data: Test and comparison with in situ data and MODIS LSTs," *IEEE Geoscience and Remote Sensing Letters* 6, 122-126 (2009).

12 J. A. Sobrino, J. El Kharraz, and Z.-L. Li, "Surface temperature and water vapour retrieval from MODIS data," *International Journal of Remote Sensing* 24(24), 5161-5182 (2003).

13 J. A. Sobrino and Y. Julien, "Trend analysis of global MODIS-Terra vegetation indices and

land surface temperature between 2000 and 2011," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, 2139-2145 (2013).

14 J. A. Sobrino and M. Romaguera, "Water vapour retrieval from Meteosat 8/SEVIRI observations," *International Journal of Remote Sensing* 29(3), 741-754 (2008).

15 C.J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment* 8(2), 127-150 (1979).

16 G. Gutman and A. Ignatov, "The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models," *International Journal of Remote Sensing* 19(8), 1533-1543 (1998).

17 F.Camacho de Coca et al., "Prototyping of cover product over Africa based on existing CYCLOPES and JRC products for VGT4Africa," *Proceedings of the 2nd Workshop on Recent Advances in Quantitative Remote Sensing (RAQRS)* 1, 722-727 (2006).

18 B. Tan et al., "An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 361-371 (2011).

19 L. Lei, Z. Zeng, and B. Zhang, "Method for detecting snow lines from MODIS data and assessment of changes in the Nianqingtanglha mountains of the Tibet plateau," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5, 769-776 (2012).

20 T.-Y. Chang et al., "Estimation of root zone soil moisture using apparent thermal inertia with MODIS imagery over a tropical catchment in northern Thailand," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5, 752-761 (2012).

21 J. Knight and M. Voth, "Mapping impervious cover using multi-temporal MODIS NDVI

data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 303-309 (2011).

22 N. Torbick et al., "Monitoring rice agriculture in the Sacramento Valley, USA with multi-temporal PALSAR and MODIS imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 451-457 (2011).

23 F. Tian et al., "Studies on the relationships between land surface temperature and environmental factors in an inland river catchment based on geographically weighted regression and MODIS data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5, 687-698 (2012).

24 D. Sun, Y. Yu, and M. D. Goldberg, "Deriving water fraction and flood maps from MODIS images using a decision tree approach," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4*, 814-825 (2011).

25 A. S. Ardakani et al., "Spatial and temporal analysis of fires detected by MODIS data in northern Iran from 2001 to 2008," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 216-225 (2011).

26 B. Geiger et al., "Land surface albedo derived on a daily basis from Meteosat Second Generation observations," *IEEE Transactions on Geoscience and Remote Sensing* 46, 3841-3856 (2008).

27 M. O. Rasmussen et al., "Directional effects on land surface temperature estimation from Meteosat Second Generation for Savanna landscapes," *IEEE Transactions on Geoscience and Remote Sensing* 49, 4458-4468 (2011).

28 B. Hirn, C. Di Bartola, and F. Ferrucci, "Combined use of SEVIRI and MODIS for detecting,

measuring, and monitoring active lava flows at erupting volcanoes," *IEEE Transactions on Geoscience and Remote Sensing* 47, 2923-2930 (2009).

29 J. A. Sobrino, Y. Julien, and G. Soria, "Phenology estimation from Meteosat Second Generation data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, 1653-1659 (2013).

30 N. Jaiswal and C. M. Kishtawal, "Objective detection of center of tropical cyclone in remotely sensed infrared images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, 1031-1035 (2013).

31 R. M. A. Timmermans et al., "The added value of a proposed satellite imager for ground level particulate matter analyses and forecasts," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2, 271-283 (2009).

32 S. Liang et al., "Review on estimation of land surface radiation and energy budgets from ground measurement, remote sensing and model simulations," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 3, 225-240 (2010).

33 LAADS, "Level 1 and Atmosphere Archive and Distribution System," *Available online: http://ladsweb.nascom.nasa.gov/* (accessed on 07th June 2014) .

34 EarthExplorer, "USGS EarthExplorer," *Available online: http://earthexplorer.usgs.gov* (accessed on 07th June 2014) .

35 M. 2.0, "USGS MRT Wev 2.0," *Available online: https://mrtweb.cr.usgs.gov/* (accessed on07th June 2014).

36 Glovis, "USGS Global Visualization Viewer," *Available online: http://glovis.usgs.gov/* (accessed on 07th June 2014).

37 MRT, "MODIS Reprojection Tool (MRT)," *Available online: https://lpdaac.usgs.gov/tools/modis reprojection tool* (accessed on 07th June 2014).

38 GCU/IPL, "Global Change Unit, Image Processing Laboratory at University of Valencia," *http://www.uv.es/ucg*.

39 H. Rahman and G. Dedieu, "SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum," *International Journal of Remote Sensing* 15(1), 123-143 (1994)

40 F.-N. Kogan, "NOAA plays leadership role in developing satellite technology for drought watch", *Earth Observation Magazine*, 18-21, (Sep. 1994).

41 W. Lucht et al., "A comparison of satellite-derived spectral albedos to ground-based broad-band albedo measurements modeled to satellite spatial scale for a semidesert landscape", *Remote Sensing of Environment*,74,85-98 (2000).

42 J. El Kharraz, "Determinación de la temperatura de la superficie terrestre a partir de datos MODIS," *PhD Thesis, Univertity of Valencia, Spain* (2009)

43 Y. Julien et al., "Near Real-Time estimation of Water Vapor column from MSG SEVIRI thermal infrared bands: implications for Land Surface Temperature retrieval", *IEEE Transactions on Geoscience and Remote Sensing*, submitted.

44 IMAPP, "International MODIS/AIRS Processing Package," *http://eostation.scanex.ru/software.html* (accessed on 27th May 2014) .

45 M. Shimrat, "Algorithm 112: Position of point relative to polygon," *Communications of the ACM* 5(8), 434- (1962).

46 J. Sevilla and A. Plaza, "A new digital repository for hyperspectral imagery with unmixing-based retrieval functionality," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1-15 (2014, in press).

Jorge Sevilla received his BS and MS degrees in computer sciences from the University of Extremadura in 2006 and 2008, respectively, where he is currently pursuing the Ph.D. degree. He has been a Research Intern with the National Institute of Nuclear Physics, Catania, Italy. He has also been a Visiting Researcher with the West University of Timisoara, Timisoara, Romania. His research interest include hyperspectral image analysis, efficient implementations, development of digital repositories, and content-based image retrieval. He has served as reviewer for several journals and conferences. Jorge Sevilla is a member of the Hyperspectral Computing Laboratory (Hypercomp).

Yves Julien is also a researcher at the Global Change Unit (http://www.uv.es/ucg) at the University of Valencia, Spain. He obtained in 2008 his Ph.D. in Earth Physics and Thermodynamics from the University of Valencia (Spain) and in Electronics, Electrotechnics, and Automatics (specialized in Remote Sensing) from the University of Strasbourg (France). Author of more than 22 international papers (http://www.uv.es/juy/publications.htm). His research interests include temperature and vegetation indices interactions, as well as time series analysis for land cover dynamic monitoring.

Guillem Sória received the Ph.D degree in Earth Physics and Thermodynamics from the University of Valencia (Spain) in 2006. Author of more than 20 international papers . He is currently a Research Scientist with the Global Change Unit (http://www.uv.es/ucg) at the University of Valencia. His research interests include thermal remote sensing and the design and implementation of experimental field campaigns related to thermal parameters retrieval.

José A. Sobrino is Professor of Physics and Remote Sensing and head of the Global Change Unit (http://www.uv.es/ucg) at the University of Valencia, Spain. Author of more than 100 papers. Coordinator of the European projects WATERMED and EAGLE. Member since November of 2003 of ESAC (Earth Science Advisory Committee) of ESA (European Space Agency). Chairperson of First, Second and Third International Symposium RAQRS (Recent Advances in Quantitative Remote Sensing) http://www.uv.es/raqrs. His research interest include atmospheric correction in visible and infrared domains, the retrieval of emissivity and surface temperature from satellite images, and the development of remote sensing methods for land cover dynamic monitoring.

Antonio Plaza is an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp). His main research interests comprise remotely sensed hyperspectral image analysis, signal processing, and efficient implementations of large-scale scientific problems on high performance computing architectures, including commodity Beowulf clusters, heterogeneous networks of computers and grids, and specialized computer architectures such as field-programmable gate arrays (FPGAs) or graphical processing units (GPUs). He was the Coordinator of the Hyperspectral Imaging Network, a European project with total funding of 2.8 MEuro. He authored more than 400 publications, including 125 JCR journal papers (77 in IEEE journals), 20 book chapters, and over 240 peer-reviewed conference proceeding papers (94 in IEEE conferences). He has guest edited seven special issues on JCR journals (three in IEEE journals). He has been a Chair for the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (2011). He is a recipient of the recognition of Best Reviewers of the IEEE Geoscience and Remote Sensing Letters (in 2009) and

a recipient of the recognition of Best Reviewers of the IEEE Transactions on Geoscience and Remote Sensing (in 2010), a journal for which he served as Associate Editor in 2007-2012. He is also an Associate Editor for IEEE Access, and was a member of the Editorial Board of the IEEE Geoscience and Remote Sensing Newsletter (2011-2012) and the IEEE Geoscience and Remote Sensing Magazine (2013). He was also a member of the steering committee of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2012). He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) in 2011-2012, and is currently serving as President of the Spanish Chapter of IEEE GRSS (since November 2012). He is also currently serving as the Editor-in-Chief of the IEEE Transactions on Geoscience and Remote Sensing journal (since January 2013). Additional information: http://www.umbc.edu/rssipl/people/aplaza.

List of Figures

6   A case study illustrating the geolocation matching accuracy of the developed geo-portal. The red polygons represent a region of interest defined by the user (centered over the Canary Islands, as depicted in Fig. 4) and the and blue polygons represent the available MODIS images at different hours on July 12, 2010. The intersection areas are displayed in cyan color. In our experiment, only the cases in which there was overlapping between the user-defined area and the available images resulted in data retrievals, which confirmed the good performance of the geolocation matching algorithm.

List of Tables