Front. Earth Sci. 2013, 7(3): 295–309
DOI 10.1007/s11707-013-0365-z

RESEARCH ARTICLE

A geospatial web portal for sharing and analyzing greenhouse gas data derived from satellite remote sensing images

Hao LIN[1], Bailang YU (✉)[1], Zuoqi CHEN[1], Yingjie HU[2], Yan HUANG[1], Jianping WU[1], Bin WU[1], Rong GE[1]

1 Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China
2 Department of Geography, University of California Santa Barbara, Santa Barbara, CA 93106, USA

Abstract Greenhouse gas data collected by different institutions throughout the world have significant scientific values for global climate change studies. Due to the diversity of data formats and different specifications of data access interfaces, most of those data should be first downloaded onto a local machine before they can be used. To overcome this limitation, we present a geospatial web portal for sharing and analyzing greenhouse gas data derived from remote sensing images. As a proof-of-concept, a prototype has also been designed and implemented. The workflow of the web portal contains four processes: data access, data analysis, results visualization, and results output. A large volume of greenhouse gas data have been collected, described, and indexed in the portal, and a variety of data analysis services, such as calculating the temporal variation of regionally averaged column CO_2 values and analyzing the latitudinal variations of globally averaged column CO_2 values, are integrated into this portal. With the integrated geospatial data and services, researchers can collect and analyze greenhouse gas data online, and can preview and download the analysis results directly from the web portal. The geospatial web portal has been implemented as a web application, and we also used a study case to illustrate this framework.

Keywords greenhouse gas data, geospatial web portal, online spatial analysis

1 Introduction

Understanding the trend and the impact factors of global climate change is one of the most important tasks faced by climatologists, geographers, and policy makers in the past decades. It has been proven that global climate change is dominated by anthropogenic activities which have been increasing the proportion of greenhouse gas in the atmosphere (Karl and Trenberth, 2003). A variety of greenhouse gas, including water vapor (H_2O), carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and ozone (O_3) exist in the atmosphere, absorbing and reflecting radiation within the thermal infrared range (Karl and Trenberth, 2003). This process is similar to the mechanism of a greenhouse which can be heated up by incoming solar radiation (Le Treut et al., 2007). Monitoring and analyzing the concentration of greenhouse gases in the atmosphere is important for understanding its impacts on global climate change.

So far, large amounts of greenhouse gas data have already been collected by numerous institutions using different instruments. Particularly, many data are from ground observations and remotely sensed images. Data from ground observations are obtained by the ground stations distributed all around the world. They monitor the concentration of greenhouse gases in the atmosphere at specific locations routinely, and tabular data (such as TXT, DBF, or Excel files) are the major results released by these stations. Remote sensing data are derived from the original images obtained by various sensors equipped on earth-observation satellites, such as Atmospheric Infrared Sounder (AIRS) on the Aqua (Aumann et al., 2003; Strow et al., 2003), Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) on the ENVISAT-1 (Bovensmann et al., 1999), Infrared Atmospheric Sounding Interferometer (IASI) on the METOP (Zhang et al., 2008), and Thermal and Near-infrared Sensor for Carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) on the GOSAT (Yokota et al., 2009). Those remote sensing data often cover a global scale, and therefore can be used for both regional and global climate change research. These data mentioned are

Received October 4, 2012; accepted February 20, 2013

E-mail: blyu@geo.ecnu.edu.cn

often in different formats since they are collected by different sensors. For example, the AIRS/Aqua Level 3 daily CO_2 data are in the Hierarchical Data Format (HDF) format, while the SCIAMACHY XCO_2 Level 2 data of WFM-DOAS version 2.1 retrievals (WFMD v2.1) are in the WFM-DOAS ASCII file (WAS) format.

The emergence of the Internet significantly facilitated the sharing and exchanging large amount of geospatial data distributed in different locations. Many agencies have established online databases or Spatial Data Infrastructures (SDIs) to support meteorological and climatic data (including greenhouse gas data) sharing (Longley and Maguire, 2005; Tait, 2005; van der Wel, 2005; Woolf et al., 2005; Titov et al., 2009). Thanks to many data providers, the greenhouse gas data derived from remote sensing images now can be directly downloaded from the Web. Although these web systems have provided different services, their main functions are basic data searching and downloading instead of online data processing (Woolf et al., 2003, 2005). The studies on greenhouse gases, such as the assessment of inversion measurements from different sensors (Barkley et al., 2006), the comparison of satellite inversion observations with in situ measurements (Blond et al., 2007), and spatio-temporal analysis of greenhouse gas distribution based on satellite data (Alkhaled et al., 2008; Bai et al., 2010; Wang et al., 2011; Zhang et al., 2011), require synergistic utilizations and specialized analysis of a large volume of data from different platforms. In those studies, scientists experience complex data processing procedures repeatedly. They have to search and download original data from different databases, convert data to specific formats compatible using a variety of processing software to fuse, assess, and analyze data, and output results. This complex process is time-consuming and labor intensive, and largely decreases the research efficiency.

A geospatial web portal (geoportal) is a web-based system that presents an entry point to geospatial data and services on the web, or, more simply, a website where such geospatial content can be discovered (Longley and Maguire, 2005; Tait, 2005). A variety of information acquisition and processing methods have been proposed for sharing and interoperating multi-source, heterogeneous, and distributed scientific data. For example, metadata are utilized to describe the contents of spatial data to the public (Goodchild and Zhou, 2003); web services are used to access and handle information; and diverse interoperation technologies are adopted to convert data between different formats (Woolf et al., 2005; Nativi and Domenico, 2009). These methods can be applied to the management and processing of greenhouse gas data.

This paper presents a geospatial web portal for sharing and analyzing greenhouse gas data derived from remote sensing images. In the following sections, we first introduce the features of satellite inversion greenhouse gas data and existing online systems for data sharing and processing. Then, the framework of the portal is described in section 4. Next, we present the implementation and function description of the portal in section 5. In section 6, a case study is used to illustrate how this portal may facilitate the acquisition and utilization of greenhouse gas data. Finally, conclusions are drawn and future work is discussed.

2 Related works

In this section, we describe the main features of greenhouse gas data and make references to existing web-based greenhouse gas databases as well as related websites. We then summarize the functional requirements of the geospatial web portal by combining these two aspects.

2.1 Features of greenhouse gas data

Greenhouse gas data records the information about greenhouse gas as proportion in the atmosphere, and focus on CO_2 and CH_4. There are a variety of greenhouse gases, but their contributions to the greenhouse effect are different. The degree of importance are sorted as: H_2O (36%–72%), CO_2 (9%–26%), CH_4 (4%–9%), O_3 (3%–7%) (Karl and Trenberth, 2003). Water vapor accounts for the largest percentage of the greenhouse effect, but the average residence time of a water molecule in the atmosphere is only nine days (AGU, 1995), compared to years or centuries of the staying time of other greenhouse gases such as CH_4 and CO_2, and water vapor is less affected by human activities. Aside from water vapor, other major greenhouse gases are well-mixed, and take years to leave the atmosphere (IPCC, 2001). They have more scientific value for greenhouse effect research compared with water vapor, although they are present at extremely low concentrations in the atmosphere. Mixing ratios are used as one of the three principal measures of atmospheric to describe atmosphere composition (Jacob, 1999). Mixing ratios of trace gases are commonly given in units of parts per million volume (ppmv or simply ppm), parts per billion volume (ppbv or ppb), or parts per trillion volume (pptv or ppt). For example, the present-day CO_2 concentration is 365 ppmv (365×10^{-6} mol/mol). Greenhouse gas data are in units of ppm or ppmv (CO_2), ppb or ppbv (CH_4) and ppt or pptv (O_3), corresponding to one geographic reference point (e.g., defined by latitude and longitude coordinates) for each value.

There are two types of greenhouse gas data with various formats as mentioned in section 1: ground observation data and remote sensing data. Compared to ground observations, remote sensing has access to large-scale information in a short time and makes it possible to collect data on dangerous or inaccessible areas. When designing a greenhouse gas metadatabase, it is better to focus on remote sensing data for future extension of the portal. However, the acquisition of remote sensing data can be largely affected by cloud cover and many other factors. The accuracy of greenhouse gas data are also influenced by different radiative transfer models. Thus, ground observation data are also important to validate the remote sensing data. But at the current stage and this paper, we focus on those greenhouse gas data derived from satellite remote sensing images.

Based on the discussion above, we summarize four main features of greenhouse gas data derived from satellite remote sensing images:

(i) Spatiality: Each greenhouse gas value is recorded with a unique geographic reference. Thus, they can be used in spatial analysis.

(ii) Multi-source: Data products from each sensor are released in their own official websites. While some can be accessed online, others can only be downloaded by particular tools or require an order beforehand.

(iii) Large volume: For instance, the AIRS data, which includes a total of 240 Level 2 CO_2 Standard Product Files and Level 2 CO_2 Support Product Files, are produced each day. Similar situation happens to the AIRS Level 2 physical retrieval products. As days and years pass by, these data have grown to a huge data collection and their volumes are increasing rapidly every day.

(iv) Format Heterogeneity: The greenhouse gas data from different sources collected by different sensors are often saved in different formats such as HDF, Network Common Data Form (NetCDF), WAS, ASCII Grid Format (GRID), etc.

2.2 Survey of existing web portals for greenhouse gas data

Currently, many databases have been established by government agencies and institutions to support the sharing of meteorological data, which also includes greenhouse gas data. We have investigated and surveyed the following authoritative and well known data service system around the world.

The Atmospheric Scientific Database and the Resources and Environment Remote Sensing Database of the Chinese Academy of Sciences have collected a lot of atmospheric and remote sensing data for online data retrieval. The marine and atmospheric data service system of the National Center for Atmospheric Research (NCAR) once utilized a Bulletin Board System (BBS) platform and now adopts a webpage of data list for low-cost data sharing. Global Atmosphere Watch (GAW) of World Meteorological Organization and the Intergovernmental Panel on Climate Change (IPCC) Data Distribution Centre contain parts of greenhouse gas observed data for online retrieval and download. But both of these have little satellite remote sensing data. NASA's Global Change Master Directory pays more attention to the search and download of remote sensing data which contains little greenhouse gas data. The National Environmental Satellite, Data, and Information Service (NESDIS) offer several information services about monitoring systems of earth and environment assessment, focusing on the retrieval and download of O_3 and chemical content rather than greenhouse gas data. Global atmospheric satellite remote sensing data of GIF, ASCII or GRIB2 formats can be accessed from this website. The National Institute for Environmental Studies (NIES) of Japan releases global atmospheric composition data as charts and these data can be downloaded as Excel files, however, this data primarily concerns the atmosphere of Japan. The United Nations Framework Convention on Climate Change (UNFCCC) and the European Environment Agency (EEA) supply large quantities of figures regarding greenhouse gas emission. Some greenhouse gas data on the list of these two web portals can be downloaded as Excel files.

Almost all of these websites have excellent retrieval and download functions for scientific data, but few of them offer further online analysis functions of these data. In addition, these databases are designed for common climate data sharing rather than greenhouse gas data and global warming research. Existing websites can provide useful references and examples for us to design and improve our own geospatial web portal. We particularly discuss details about data access, data analysis, results visualization, and results output of two representative online databases with completed online analysis functions: the Atmospheric Science Data Integration Analysis Platform and GIOVANNI.

Atmospheric Science Data Integration Analysis Platform[1)] is provided by Institute of Atmospheric Physics and Computer Network Information Center of the Chinese Academy of Sciences. It provides a simple processing and analysis platform for atmospheric science practitioners.

(i) Data access: Data of this platform are stored in mirror database for domestic research personnel in China and can be accessed directly. Currently, users can browse and download atmospheric data from data providers including NCAR and IPCC. Users can get latest version of data from a mirror database directly without worrying about the data quality. However, only scientific data of NetCDF fomat can be analyzed.

(ii) Data analysis: Processing functions are offered to registered users and contains integration analysis and online run capability. NetCDF data can be selected for online analysis. After selecting data for computing, users can choose the algorithm for computation or plot and set parameters, such as Canonical Correlation Analysis (CCA), Empirical Orthogonal Function (EOF), Principal Components Analysis (PCA) or Principal Oscillation Pattern (POP). Besides choosing algorithms provided by platform, users can also write R scripts themselves to run

1) http://adoap.csdb.cn/

these data. However, users cannot implement regional or time series analysis based on particular spatiotemporal properties such as a specific region or a time period.

(iii) Results visualization: After data analysis, results are shown in a plot or map graphics, which are generated as Encapsulated PostScript (EPS) format for users to preview.

(iv) Results outputs: Analysis results can be downloaded as images only. In addition, users can save or share work flows for test and verification. After completing online analysis of scientific data, users can give some feedback online or participate in the training about how to use this platform.

GIOVANNI[1] is developed by the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). It is a web-based application that provides a simple and an easy way to visualize, analyze, and access vast amounts of Earth science remote sensing data. The AIRS data, in which we are interested, are supported by GIOVANNI.

(i) Data access: GIOVANNI can access vast amounts of Earth science remote sensing data released by NASA, particularly from satellites. However, data released by satellites of other agencies such as SCIAMACHY or GOSAT, in which we are interested for this study, cannot be accessed or analyzed online in GIOVANNI.

(ii) Data analysis: Currently, GIOVANNI consists of several portals tailored to meet the needs of different earth science research communities. They can be classified into five categories: atmospheric, meteorological, oceanic, and hydrologic as well as application and education. In the online analysis page, users can choose one type of data product for analysis, for instance, AIRS Level 3 monthly CO_2 data. After the user has selected space-time conditions such as spatial region or time period as well as type of visualization such as chart or map, data are analyzed in the background and processing messages are displayed in a popup page.

(iii) Results visualization: Analysis results are displayed as chart or shaded map with color bar.

(iv) Result output: Chart results can be output and downloaded as PNG pictures while map results can be downloaded as KMZ files for further display in Google Earth on users' own computers.

In summary, existing web portals have three main limitations. First, the greenhouse gas data required for a particular application are often scattered on the websites of different data providers. In most cases, it is difficult for researchers to access and analyze data by simply choosing one web portal. For instance, GIOVANNI have almost perfect functions for data access and online analysis, but except the data from NASA, it cannot do online analysis for the data from others agencies (e.g., SCIAMACHY data). Second, users have to specifically choose data of the same format (e.g., NetCDF) or one kind of satellite sensors (e.g., AIRS) for further online analysis. These systems still cannot directly process data in heterogeneous formats from different sources. Finally, when requested to output analysis results as charts, existing web portals only provide pictures (e.g., ESP or PNG) for downloading. However, tabular files (such as Excel files) are as valuable as pictures for chart results, and therefore should also be provided for the users.

Our web portal overcomes the above limitations by providing one-stop services for data access and online analysis. Particularly, we focus on greenhouse gas data derived from satellite remote sensing images, but this web portal can also be applied to handle other types of data from different sources.

3 Data for the portal

The greenhouse gas data that we shared in our web portal are from ten data set products (Table 1), with a 400 GB total data size. Except for the product named AIRS L2 CO_2 Grid, all of the others are collected and downloaded directly from official websites, including AIRS[2] L2 CO_2 Swath, AIRS L3 CO_2 Product, GOSAT[3] L2 CO_2, IASI[4] L2 EPS, SCIAMACHY[5] L2 XCO_2, SCIAMACHY L3 XCO_2, FY-3A[6] L2 O_3, OMTO3e[7] L3 O_3, and OMI_-DOAS L3 O_3. AIRS L2 CO_2 Grid is summed and averaged from AIRS L2 CO_2 Swath with three kinds of temporal resolutions, respectively, covering a calendar month, eight days, and one day. The spatial parameters are averaged into $1° \times 1°$ grid cells, from $-180.0°$ to $+180.0°$ longitude and from $-90.0°$ to $+90.0°$ latitude. Table 1 also shows the details of the data, such as platform and instrument, time span, temporal resolution, spatial resolution, content, format, and quantity.

The formats of the greenhouse gas data we focus on in this research are HDF (released by AIRS and FY-3A), HDF5 (released by GOSAT and IASI, OMTO3e and OMI_DOAS), WAS, as well as GRID (WAS and GRID are released by SCIAMACHY). HDF is a free library and platform independent data format for storage and exchange of scientific data, designed to organize large amounts of numerical data. HDF has two major versions, HDF4 and

1) http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html
2) http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings
3) http://www.gosat.nies.go.jp/index_e.html
4) http://www.eumetsat.int/Home/index.htm
5) http://www.iup.uni-bremen.de/sciamachy/NIR_NADIR_WFM_DOAS/downloads/
6) http://satellite.cma.gov.cn/PortalSite/default.aspx
7) http://disc.sci.gsfc.nasa.gov/giovanni/additional/data-holdings/PIP/ozone.shtml

Table 1 List of greenhouse gas data available in the spatial web portal

Name	Platform/Instrument	Time span	Temporal resolution	Spatial resolution	Content	Format	Quantity
AIRS L2 CO_2 Swath	Aqua/AIRS	2002.9–2010.12	About 12 h	1° × 1°	CO_2	HDF	737338
AIRS L3 CO_2 Product	Aqua/AIRS	2002.9–2012.2	Daily, weekly, monthly	2° × 2.5°	CO_2	HDF	3473
AIRS L2 CO_2 Grid	Aqua/AIRS	2002.9–2010.12	Daily, weekly, monthly	1° × 1°	CO_2	HDF	3482
GOSAT L2 CO_2	GOSAT /TANSO-FTS	2009.4–2011.2	3-day	10.5 km × 10.5 km	CO_2	HDF5	9054
IASI L2 EPS	METOP/IASI	2009.1–2010.12	/	25.0 km × 25.0 km	H_2O, CO_2, N_2O, CH_4, O_3	HDF5	2263
SCIAMACHY L2 XCO_2	ENVISAT-1 /SCIAMACHY	2003.1–2009.12	Monthly	30.0 km × 60.0 km	CO_2, CH_4	WAS	84
SCIAMACHY L3 XCO_2	ENVISAT-1 /SCIAMACHY	2003.1–2009.12	Monthly	0.5° × 0.5°	CO_2, CH_4	GRID	84
FY-3A L2 O_3	FY3A/TOU	2009.7–2011.12	Daily	0.5° × 0.5°	O_3	HDF	903
OMTO3e L3 O_3	Aura/OMI	2004.10–2011.12	Daily	0.25° × 0.25°	O_3	HE5	1354
OMI_DOAS L3 O_3	Aura/OMI	2004.10–2011.12	Daily	0.25° × 0.25°	O_3	HE5	2819

HDF5, including filename extensions of hdf, hdf4, h5, hdf5, he4, he5. Both WAS and GRID are ASCII formats.

4 Framework of the portal

This paper proposes a composite framework to effectively support online data access, analysis, visualization, and output of greenhouse gas data. In this section, we first discuss the requirements for the portal design, and then present a framework of the portal.

The essential requirements of a spatial web portal can be summarized as three points: storage and management, retrieval and access, as well as customization and analysis. First of all, the portal should support storage and management of multi-source and heterogeneous greenhouse gas data for users to access without requiring users to consider the sources and formats of greenhouse gas data. Such data are the basis for the further spatial analysis functions requested by the users. Second, the portal should allow users to retrieve those data directly online. Such data can be retrieved by their spatial (e.g., selecting a regional coverage), temporal (e.g., choosing a time range) and thematic indexes (e.g., using topic-related keywords). Finally, the portal should support the capacities of greenhouse gas data online customization and analysis. Users should be allowed to download the simple analysis results using a web browser instead of having to install some client application beforehand.

Based on the requirements above, we designed a multi-layer framework which consists of two servers (Fig. 1): a web server and a data server. The data server is used to store all data, which contains real data and metadata. The web server provides an interactive user interface, and can be visited by web browsers. It has access to the data server, and can output online analysis results to the client side. The web server contains four components, which are data access, data analysis, result visualization, and result output.

5 Implementation

In this section, we describe the key technologies utilized to implement these four functional modules. Figure 2 shows the detailed components of the four modules.

5.1 Data access

So far, ten data sets (shown in Table 1) are provided on this web portal, and each of them contains hundreds and thousands of pieces of data. Such data are the sources for online spatial analysis functions, and they can also be accessed and downloaded directly from the portal. We designed the data access module as two sub modules, data deployment and data retrieval. In the following section, we will discuss these in sequence.

5.1.1 Deployment of real data and metadatabase

Both real data and metadata are deployed on the data server by offline processes, and each data set has been stored in a separated local disk drive, all greenhouse gas data can be accessed by the metadatabase.

The metadatabase contains two parts: metadata of the data set and metadata of the data within the data set. Metadata of data set is used to describe the information of ten data sets, pointing to all data in one data set. A metadata standard is essential for constructing a structured metada-

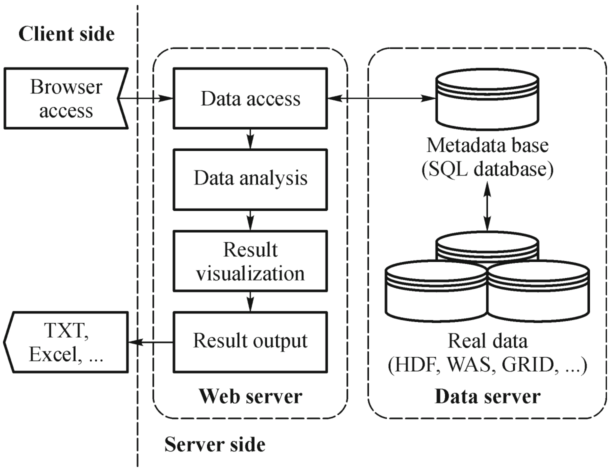


Fig. 1 The architecture of the framework.

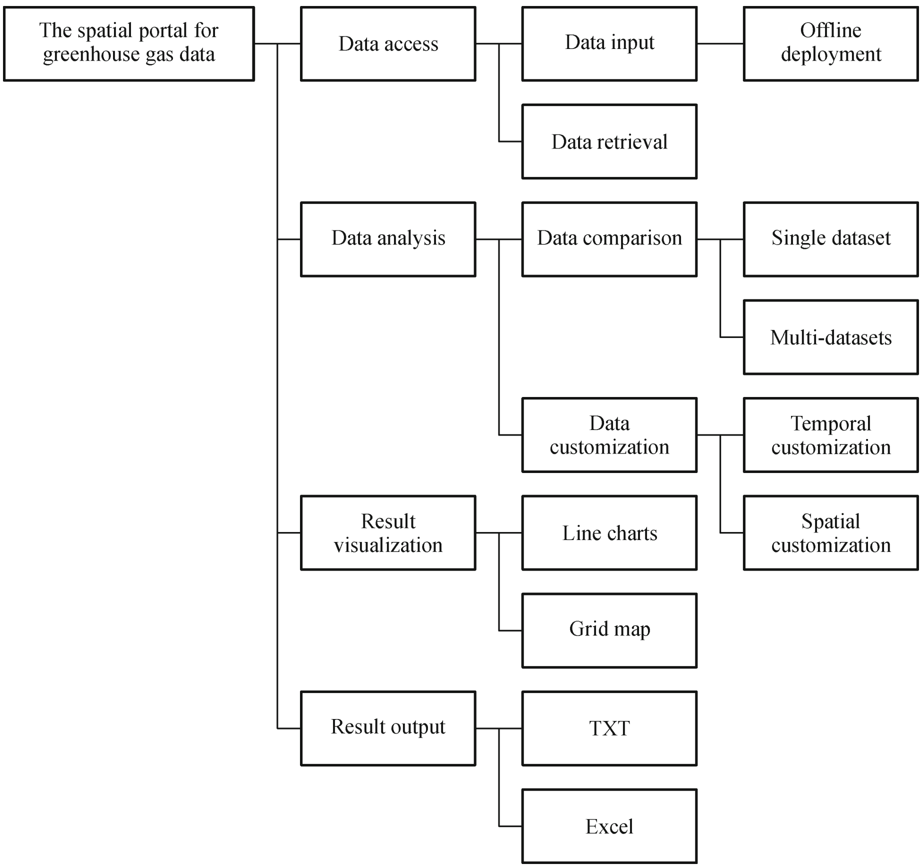


Fig. 2 Functional modules of the portal.

tabase. Two standards are chosen for references. The first one is Content Standard for Digital Geospatial Metadata (CSDGM) which is maintained by the Federal Geographic Data Committee (FGDC). The second one is ISO 19115 which defines how to describe geographical information and associated services, including contents, spatial-temporal purchases, data quality, access, and rights to use. By considering these metadata standards, we define the content structures of metadata in order to access the greenhouse gas data stored on the server. SQL Server 2005 is used for constructing metadatabase, and the recorded metadata contains the data's identification, contact and authorities and publication details, spatial-temporal information, quality and limitations, reference information, as well as maintenance description. Metadata of the data within the data set is used to describe a single piece of data in one data set, pointing to each piece of real data. According to the features of greenhouse gas data we discussed in section 2.1, metadata of data in the data set can be simplified into five parts: data name, data set name, time range, spatial extent, and preview image.

Metadata operating is realized on the server side. While data set metadata can be edited directly by humans, metadata of data in data set need to be operated by batch processing programs because of their large data volume as shown in Table 1. We batch import these description information of single piece of data into SQL Server by Environment for Visualizing Images (ENVI)/Interactive Data Language (IDL). The ENVI is an application for remote sensing multispectral and hyper spectral image analysis written in IDL. IDL is a programming language used for two-dimensional and multidimensional data visualization, analysis, and application development. Over 90 types of images and vector-data formats can be processed by ENVI, and it also supports the conversions between different image formats, such as HDF, HDF5, WAS, and GRID which are common data types used by greenhouse gas data.

5.1.2 Data retrieval

By using metadata searching, greenhouse gas data in the web portal can be directly downloaded or used as the input for online analysis. Users can check the data set list to see all available data sets, and can search a particular data set by choosing selection criteria.

The data set list shows the information of all available data sets including simplified information (Fig. 3(a)) and detailed information (Fig. 3(b)). Simplified information is listed in seven columns which are similar to Table 1. When the user clicks the "details" link, detailed metadata information will be displayed in new pages. Data amount of each data set is associated with the metadatabase, which can automatically show the changes after new data are imported or old data are deleted.

Data retrieval allows users to obtain data sets by setting retrieval options, temporal indexes (Fig. 4(a)), as well as coordinate range (Fig. 4(b)). Google Maps JavaScript API V3[1)] is chosen to implement spatial search in this portal. Greenhouse gas data are stored in the database with the upper-left latitude, upper-left longitude, lower-right latitude, lower-right longitude, center latitude, and center longitude. After users drag a box on Google Maps, the geometric parameters of the box (such as latitude and longitude of the upper-left point) are calculated and output to the corresponding textbox. If the center latitude and longitude of particular data are located within the boundary of the box, then those data satisfy the spatial requirements, and the search results will be displayed (Fig. 4(c)). The user can choose to download a particular piece of data or can simply check all the checkboxes for a bulk download.

5.2 Data analysis

The analysis methods provided by the portal enable users to analyze their data online. Such analysis methods also allow users to input their own parameters and analyze data in a customized way. There are two data process modules: data comparison and data customization. These modules are implemented using IDL, ArcObjects, Java, and JSP. In the following paragraphs, we first describe the technical foundations used to implement the data analysis functions, and then discuss the analysis methods and their theoretical basis.

5.2.1 Technical foundation

Two approaches are designed to implement online data analysis functions. The first one is using IDL to analyze heterogeneous greenhouse gas data directly on the server side. The second approach is unifying different formats before analysis is performed.

On the one hand, IDL is chosen for server-side analysis because it supports all formats of greenhouse gas data in the web portal. We can access data directly by IDL without considering their formats. Export Bridge is employed to integrate IDL objects into the module so that IDL objects can be called directly by Java or Component Object Model (COM) applications. The web portal is developed by JSP (Java Server Pages), and Java wrapper objects are used to interact with the IDL objects. In a word, ENVI/IDL is used for automatic processing on the server, while JSP is employed to implement the interactions between users and the portal interface as well as the connections among system modules.

On the other hand, parts of data are unified in advance for polygon region analysis, which can be used for enhancing inadequate spatial analysis ability of IDL, improving the availability of online analysis functions. ArcGIS is used to integrate heterogeneous data. With the feature of spatiality, greenhouse gas data can be converted into map data. In our web portal, we use two methods for unifying data. One way is converting all data into GeoTiff format data which can be accessed by ArcObjects for spatial analysis. The other way is to transform data into ArcSDE, which can be deployed in ArcGIS Server as map layer for online access.

5.2.2 Analysis methods

Four analysis methods have been provided by the web portal: single data set comparison, multiple data set

1) https://developers.google.com/maps/documentation/javascript/

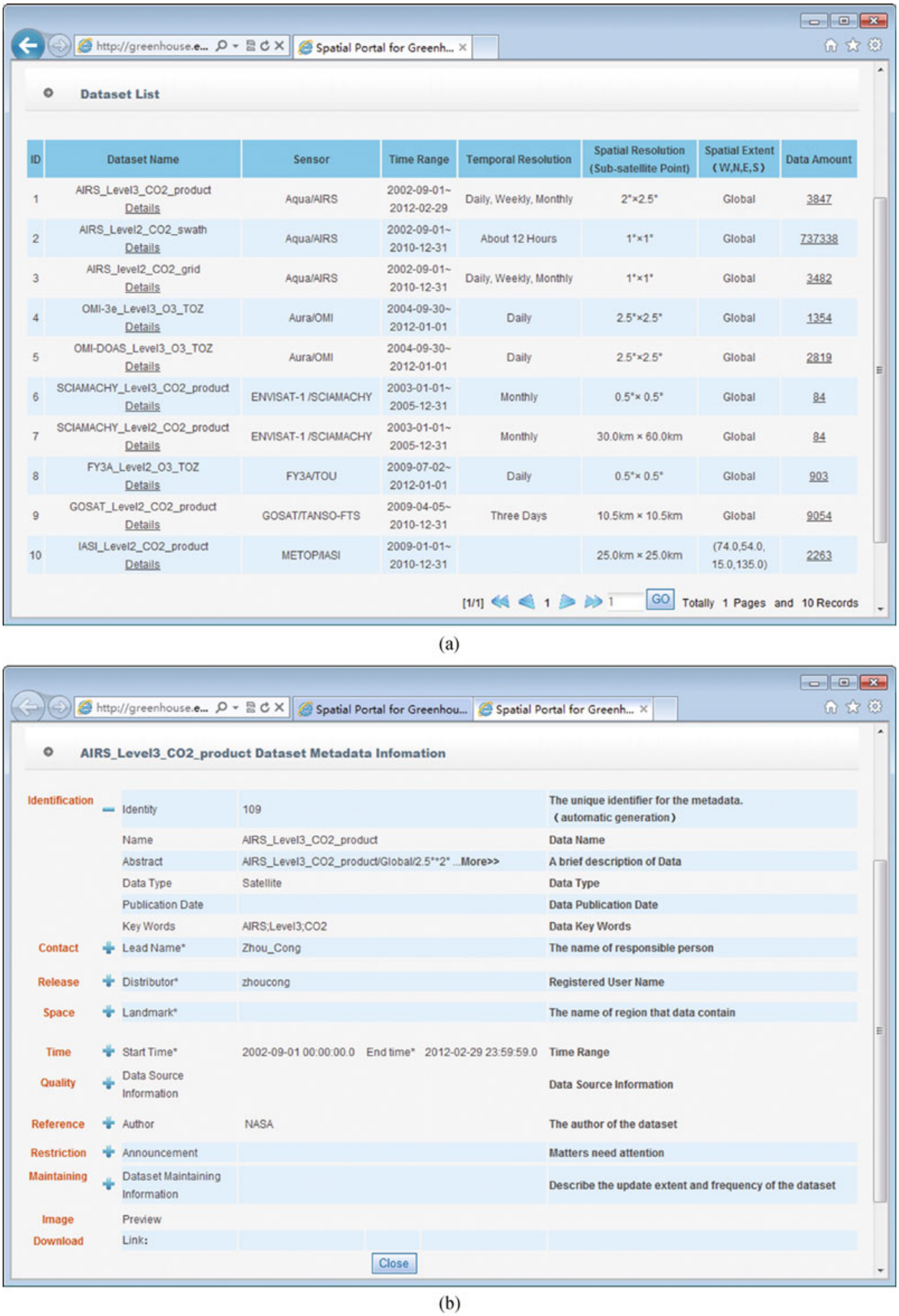
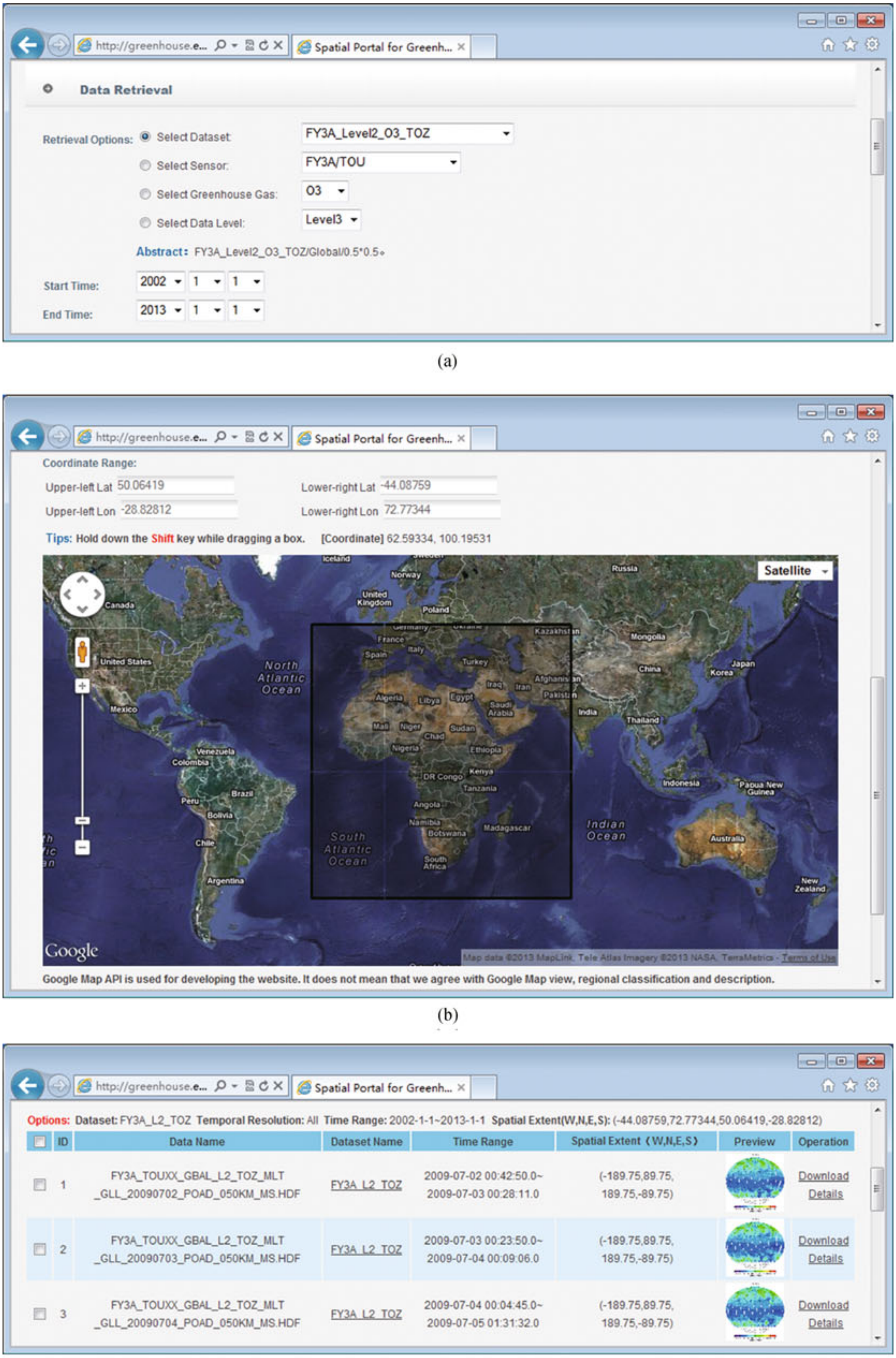


Fig. 3 Data set list of the portal. (a) Simplified information, (b) metadata in details.

comparisons, temporal customization, and spatial customization. All of the analysis methods have been implemented for a variety of greenhouse gases. In this section, we only discuss the methods for the CO_2 data, since the methods for other greenhouse gas data sets are similar.

(i) Temporal variation of averaged column CO_2 values in one point: This analysis constitutes the modules of one data set comparison and temporal customization. the CO_2 values ($p_{tx_0y_0}$) in one point of longitude (x_0) and latitude



(a)

(b)

(c)

Fig. 4 Data retrieval. (a) options selection (b) spatial retrieval (c) result list.

(y_0) can be compared with the other values at the same point in a time series (t), whose expression is defined as

$$f(t) = p_{tx_0y_0}, \tag{1}$$

where p stands for the value of CO_2 (unit is PPMV). Thus, daily, weekly, or monthly variations of columnar CO_2 values can be calculated at any point on earth.

(ii) Temporal variation of region averaged columnar CO_2 values: This analysis constitutes the modules of several data set comparisons and temporal customization. It is applied to calculating the averaged CO_2 values (p_{txy}) from a certain geographic extent and comparing them in a time series. The geographic extent can be given as a box or a polygon drawing in a map by users. We assume the coordinate region of a box is composed of maximum longitude, minimum longitude, maximum latitude, and minimum latitude. The whole expression is given as

$$f(t) = \frac{\sum_{x=x_{\min}}^{x=x_{\max}} \sum_{y=y_{\min}}^{y=y_{\max}} p_{txy}}{N_x \cdot N_y}, \tag{2}$$

where N_x and N_y stand, respectively, for the number of x and y according to the spatial resolution.

(iii) Latitudinal variations of globally averaged column CO_2 values: The modules of several data sets comparison and spatial customization consist of this analysis method. The earth is divided as ten degrees per band and 18 latitude bands (n) in total and the averaged value of each band is calculated in a certain time range. The whole expression is given as

$$f(n) = \frac{\sum_{x=-180}^{x=180} \sum_{y=-90+10n}^{y=-90+10(n+1)} \sum_{t=t_{\min}}^{t=t_{\max}} p_{txy}}{N_x \cdot N_y \cdot N_t}, \quad (0 \leqslant n \leqslant 17), \tag{3}$$

where N_t stands for the number of t according to the temporal resolution. Through analysis of latitudinal variations, scientist can get the spatial distribution of several data set products released from various satellite sensors in each band. Latitudinal variation corresponds to temperature variation and intensity variation of solar radiation, which not only associates closely with the concentration of greenhouse gas data but also influence the performance of satellite sensor. Scientists can use these data for analysis of both equipment difference and greenhouse gas distribution.

(iv) Seasonal variation of total columnar CO_2 values: This method contributes to several data sets comparison and spatial customization. A year is divided into four seasons and the averaged value of each band is calculated per season. The seasons (m) are divided as winter (m = 0) for the last December to February, summer (m = 1) March to May, spring (m = 2) June to August and autumn (m = 3) September to November. The whole expression is given as

$$f(n,m) = \frac{\sum_{x=-180}^{x=180} \sum_{y=-90+10n}^{y=-90+10(n+1)} \sum_{t=t(m)_{\min}}^{t=t(m)_{\max}} p_{txy}}{N_x \cdot N_y \cdot N_t}, \quad (0 \leqslant n \leqslant 17, 0 \leqslant m \leqslant 3), \tag{4}$$

5.3 Result visualization

Result visualization allows users to preview analysis results before downloading. It can help users to evaluate the analysis result and adjust parameters. After online data analysis, the results can be shown in two ways: line charts (Fig. 5(a)) and a grid map with color bar (Fig. 5(b)).

Flot[1)] is employed to realize line charts. Greenhouse gas data are analyzed by ENVI/IDL and output as two one-dimensional arrays: one for time series and the other for the value series of CO_2 in PPMV corresponding to time series. Both arrays are transferred to Flot and then shown as line charts. The line charts can be integrated into Google Maps or shown directly in a web page. The grid map is implemented by ArcGIS Server and ArcSDE. Map are pre-published by ArcGIS Server and data are preprocessed and transformed into ArcSDE.

5.4 Result output

Users can download the analysis results as a TXT file or an Excel document if such results satisfy the users' requirements. The results can be applied to further research of greenhouse gas and global climate change.

The TXT file is automatically created by the web portal, and the charts in Excel file have been created as Excel templates with null values beforehand. Each online analysis method corresponds to one specific Excel template. These templates are named with unique names and all stored in the server side. When users choose to download analysis results as an Excel file, the specific blank template is copied to a target folder named by user ID. Then, the analysis results output from previous steps are automatically filled into the blank template.

6 Case study

The web portal has been implemented as a prototype named Spatial Portal for Greenhouse Gas Data[2)]. Currently, the web portal provides online analysis functions for three types of data: HDF data of the AIRS Level 3 CO_2

1) http://www.flotcharts.org
2) http://greenhouse.ecnu.edu.cn/

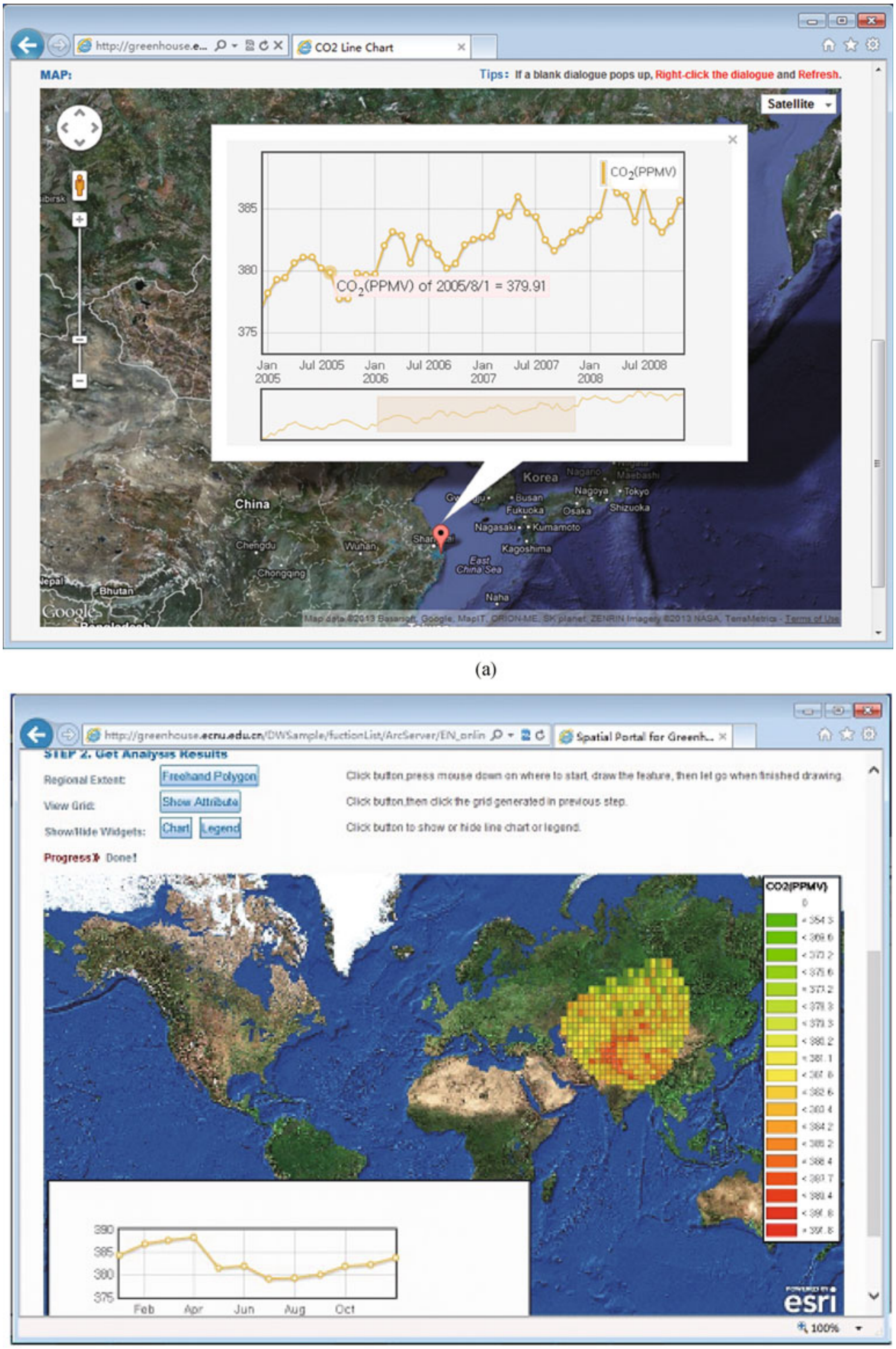
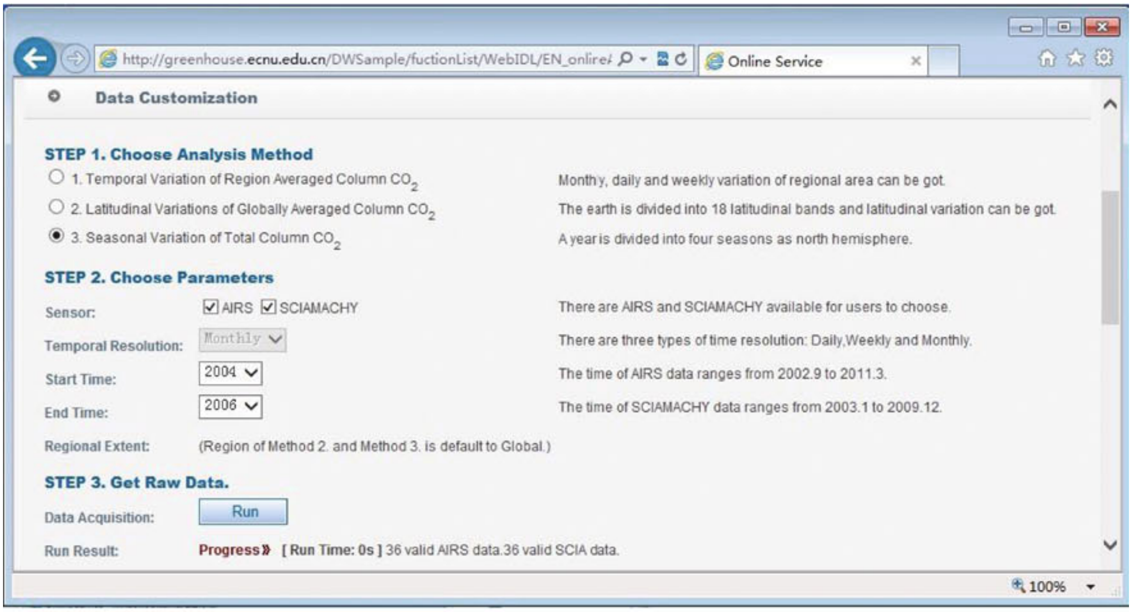


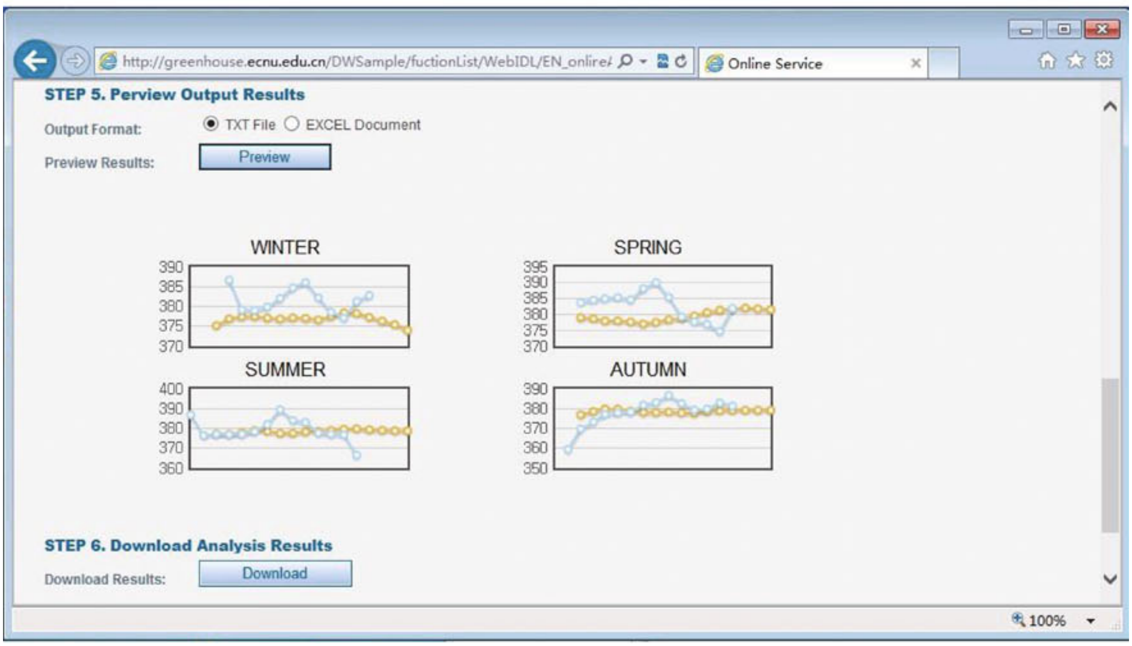
Fig. 5 Results visualization. (a) line chart (b) grid map.

product, GRID data of the SCIAMACHY Level 3 CO_2 product, and HDF5 data of the GOSAT Level 2 CO_2 product. All analysis methods for AIRS can be finished within one minute, and it takes two to five minutes for analyzing SCIAMACHY depending on the size of the data. Countdown clocks have been integrated with each analysis method and users can check how much time has been used for each online analysis. In this section, we use one study case to illustrate how data from this portal are accessed and analyzed, as well as how results are visually



(a)

(b)



(c)

Fig. 6 Seasonal variation of total columnar CO_2 values derived from AIRS and SCIAMACHY covering Dec 2004–Nov 2006. (a) Data access (according method of choice), (b) online analysis (according method of choice), (c) data visualization represented and output.

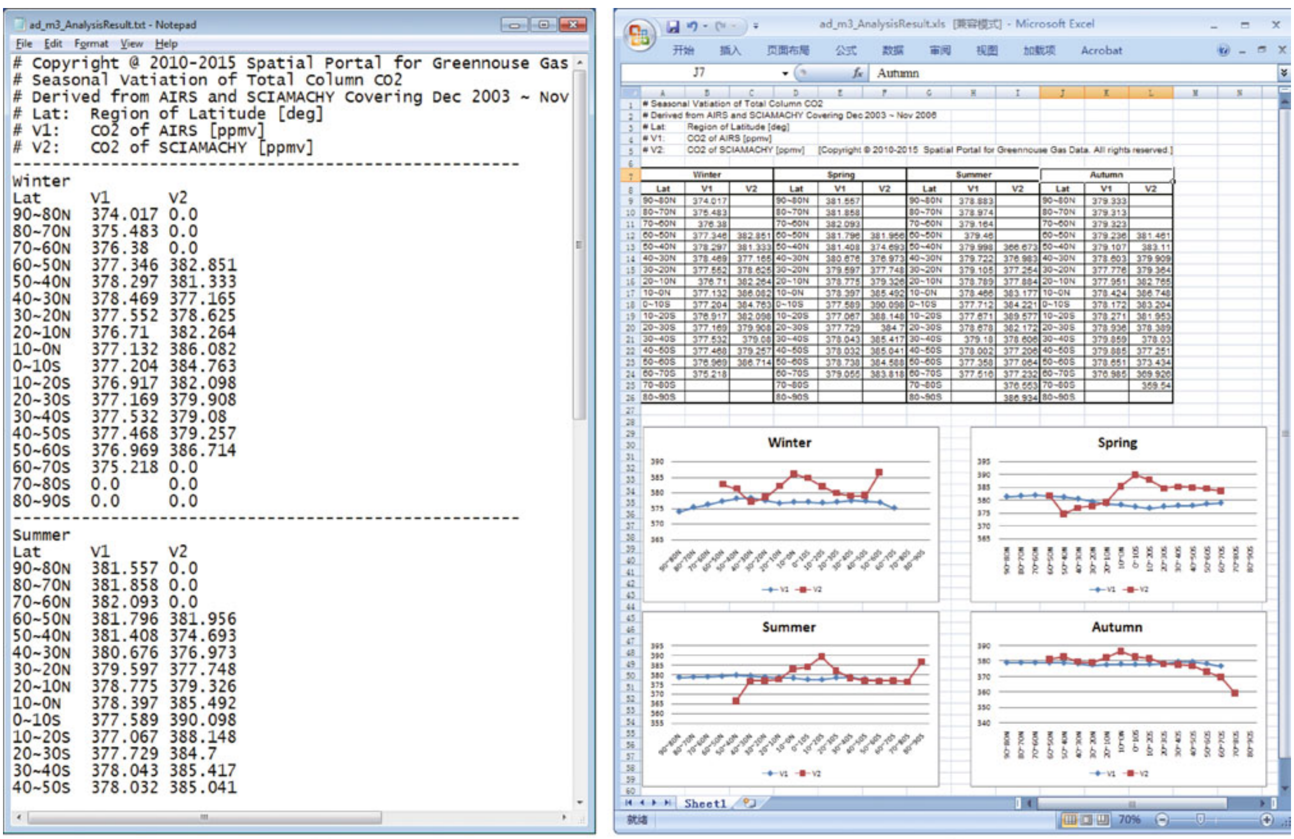
The analysis method chosen for this case is seasonal variation of total columnar CO_2 values. Five parameters are used for this case, including sensor, temporal resolution, start time, end time, and regional extent. AIRS and SCIAMACHY are chosen as the sensors, and temporal resolution is set as one month, and the regional extent is set as the whole globe by default (Fig. 6(a)). Users can also set the extent by drawing a polygon boundary on the map according the needs of other analysis methods. After users click the button to analyze these selected data (Fig. 6(b)), the analysis results are shown in four line charts which represent winter, spring, summer, and autumn (Fig. 6(c)), and the results can be output in two kinds of formats (TXT and Excel) (Fig. 7).

This study case demonstrates that the web portal can present the seasonal variation of total columnar CO_2 values derived from AIRS and SCIAMACHY covering Dec 2004 to Nov 2006. AIRS and SCIAMACHY are calculated at the same time and their analysis results are displayed in one line chart template. From these results, we can see distinct negative relevance between these two sensors and SCIAMACHY has better sensibility of latitudinal variations than AIRS.

7 Conclusions and future work

In this research, a geospatial web portal for greenhouse gas data has been designed and implemented as a prototype. This portal has been designed as a one-stop service for managing, retrieving, analyzing, and representing heterogeneous greenhouse gas data from multiple sources. We described the functions of this web portal, and then discussed the key technologies for implementation. Finally, we used one study case to evaluate its online analysis functions. Compared with existing web portal systems, our system not only realizes common data retrieval and online analysis function, but also meets the specific analysis requirements of greenhouse gas data, enhancing the multi-source data analysis and results output functions beside simple pictures.

However, the web portal described in this paper is still in a preliminary stage and more work needs to be done. More types of greenhouse gas data including CH_4 and O_3 and more data sources can be integrated into the portal to enhance its capability to analyze heterogeneous geospatial data. We will continue perfecting the functions of online processing and further improve the performance of the portal.



(a) (b)

Fig. 7 Data output and download. (a) In Txt file, (b) In Excel.

Acknowledgements This work is supported by the National Basic Research Program of China (Grant No. 2010CB951603) and the National Natural Science Foundation of China (Grant No. 41001270). The authors thank five anonymous reviewers for their constructive comments and suggestions.

References

AGU (1995). Water Vapor in the Climate System Special Report. Washington, D C: American Geophysical Union Publications Office

Alkhaled A A, Michalak A M, Kawa S R, Olsen S C, Wang J W (2008). A global evaluation of the regional spatial variability of column integrated CO_2 distributions. J Geophys Res, D, Atmospheres, 113 (D20): D20303

Aumann H H, Chahine M T, Gautier C, Goldberg M D, Kalnay E, McMillin L M, Revercomb H, Rosenkranz P W, Smith W L, Staelin D H, Strow L L, Susskind J (2003). AIRS/AMSU/HSB on the aqua mission: design, science objectives, data products, and processing systems. IEEE Trans Geosci Rem Sens, 41(2): 253–264

Bai W G, Zhang X Y, Zhang P (2010). Temporal and spatial distribution of tropospheric CO_2 over China based on satellite observations. Chin Sci Bull, 55(31): 3612–3618

Barkley M P, Monks P S, Engelen R J (2006). Comparison of SCIAMACHY and AIRS CO_2 measurements over North America during the summer and autumn of 2003. Geophys Res Lett, 33(20): L20805

Blond N, Boersma K F, Eskes H J, van der A R J, van Roozendael M, de Smedt I, Bergametti G, Vautard R (2007). Intercomparison of SCIAMACHY nitrogen dioxide observations, in situ measurements and air quality modeling results over Western Europe. J Geophys Res, D, Atmospheres, 112(D10): D10311

Bovensmann H, Burrows J P, Buchwitz M, Frerick J, Noel S, Rozanov V V, Chance K V, Goede A P H (1999). SCIAMACHY: mission objectives and measurement modes. J Atmos Sci, 56(2): 127–150

Goodchild M F, Zhou J Y (2003). Finding geographic information: collection-level metadata. GeoInformatica, 7(2): 95–112

IPCC (2001). 6.3 Well-mixed Greenhouse Gases. Working Group I: The Scientific Basis IPCC Third Assessment Report–Climate Change 2001, Retrieved 2012, from http://www.grida.no/publications/other/ipcc_tar/?src =/climate/ipcc_tar/wg1/218.htm

Jacob D (1999). Introduction to Atmospheric Chemistry. New Jersey: Princeton University Press

Karl T R, Trenberth K E (2003). Modern global climate change. Science, 302(5651): 1719–1723

Le Treut H, Somerville R, Cubasch U, Ding Y, Mauritzen C, Mokssit A, Peterson T, Prather M (2007). Historical Overview of Climate Change Science. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K B, Tignor M, Miller H L, eds. Climate Change 2007: The Physical Science Basis. Cambridge: Cambridge University Press

Longley P A, Maguire D J (2005). Geoportals. Comput Environ Urban Syst, 29(1): 1

Nativi S, Domenico B (2009). Enabling interoperability for Digital Earth: Earth Science coverage access services. International Journal of Digital Earth 2(Suppl 1): 79–104

Strow L L, Hannon S E, de Souza-Machado S, Motteler H E, Tobin D (2003). An overview of the AIRS radiative transfer model. IEEE Trans Geosci Rem Sens, 41(2): 303–313

Tait M G (2005). Implementing geoportals: applications of distributed GIS. Comput Environ Urban Syst, 29(1): 33–47

Titov A, Gordov E, Okladnikov I, Shulgina T (2009). Web-system for processing and visualization of meteorological data for Siberian environment research. International Journal of Digital Earth, 2(Suppl 1): 105–119

van der Wel F J M (2005). Spatial data infrastructure for meteorological and climatic data. Meteorol Appl, 12(1): 7–8

Wang K, Jiang H, Zhang X Y, Zhou G M (2011). Analysis of spatial and temporal variations of carbon dioxide over China using SCIAMACHY satellite observations during 2003–2005. Int J Remote Sens, 32 (3): 815–832

Woolf A, Cramer R, Gutierrez M, van Dam K K, Kondapalli S, Latham S, Lawrence B, Lowry R, O'Neill K (2005). Standards-based data interoperability in the climate sciences. Meteorol Appl, 12(1): 9–22

Woolf A, Haines K, Liu C L (2003). A web service model for climate data access on the grid. Int J High Perform Comput Appl, 17(3): 281–295

Yokota T, Yoshida Y, Eguchi N, Ota Y, Tanaka T, Watanabe H, Maksyutov S (2009). Global Concentrations of CO_2 and CH_4 Retrieved from GOSAT: first preliminary results. SOLA, 5: 160–163

Zhang L, Dong C H, Zhang W J, Zhang P (2008). METOP-on-board super-high spectrum resolution infrared atmospheric sounding interferometer (IASI). Meteorological Science and Technology, 36 (5): 639–642 (in Chinese)

Zhang X Y, Jiang H, Wang Y Q, Han Y, Buchwitz M, Schneising O, Burrows J P (2011). Spatial variations of atmospheric methane concentrations in China. Int J Remote Sens, 32(3): 833–847

AUTHOR BIOGRAPHIES

Hao Lin received the B.S. degree in Geographical Information System from East China Normal University, Shanghai, China, in 2010. She is currently a master student in Cartography and Geographical Information System at Department of Geography, East China Normal University. Her current research interests focus on Development of Web GIS and Remote Sensing.

Bailang Yu received the B.S. degree and Ph.D. degree from East China Normal University, Shanghai, China in 2002 and 2009, respectively. He is currently an Associate Professor of Key Laboratory of Geographic Information Science (Ministry of Education) and Department of Geography, East China Normal University. His research interests include object-oriented analysis for remotely sensed images, LiDAR, Urban Remote Sensing, and GIS (Geographical Information System). His recent research projects have been funded by the National Natural Science Foundation of China, the Specialized Research Fund for the Doctoral Program of Higher Education, and the Fundamental Research Funds for the Central Universities of China.

Zuoqi Chen received the B. S. degree in Geographical Information System from Fujian Normal University, China, in 2011. He is currently a master student in Cartography and Geographical Information System at Department of Geography, East China Normal University. His current research interests focus on GIS software development.

Yingjie Hu received the M.S. degree in Geographical Information System from East China Normal University, Shanghai, China, in 2009. He is currently a Ph.D. student at the Department of Geography of the University of California, Santa Barbara, USA. His current research interests focus on Geospatial semantics, Personal GIS, Semantic Web, Semantic Desktop, and mobile computing.

Yan Huang received the B.S. degree in Geographical Information System from Chengdu University of Technology, Chengdu, China, in 2010. She is currently a Ph.D. student in Cartography and Geographical Information System at Department of Geography, East China Normal University. Her current research interests focus on object-oriented analysis for remotely sensed images and LiDAR.

Jianping Wu received the B. S. degree from Nanjing University, Nanjing, China in 1983, the M. S. degree from Peking University, Beijing, China, in 1986, and the Ph.D. degree from East China Normal University, Shanghai, China, in 1996. He is currently a Professor of Key Laboratory of Geographic Information Science (Ministry of Education) and Department of Geography, East China Normal University. His research interests include Geographical Information System Development and Remote Sensing Application.

Bin Wu received the B. S. degree in Geographical Information System from East China Normal University, Shanghai, China, in 2011. He is currently a master student in Cartography and Geographical Information System at Department of Geography, East China Normal University. His current research interests focus on Geographical Information System Development and Remote Sensing.

Rong Ge is currently an undergraduate student in Geographic Science at Department of Geography, East China Normal University, Shanghai, China.