Development of a Web-based visualization platform for climate research using Google Earth

Xiaojuan Sun a,d,*, Suhung Shen b,c, Gregory G. Leptoukh c, Panxing Wang a, Liping Di b, Mingyue Lu a

* China Institute of Manufacturing Development, Nanjing University of Information Science & Technology, 219 Ningliu Road, Nanjing, Jiangsu 210044, China
b Center for Spatial Information Science and Systems (CSISS), George Mason University, Fairfax, VA 22030, USA
c Center for Spatial Information Science and Systems (CSISS), George Mason University, Fairfax, VA 22030, USA
d Key Laboratory of Meteorological Disaster of Ministry of Education, Nanjing University of Information Science & Technology, 219 Ningliu Road, Nanjing, Jiangsu 210044, China

A R T I C L E   I N F O
Article history:
Received 18 March 2011
Received in revised form 28 August 2011
Accepted 13 September 2011

Keywords:
Data visualization
KML
Google Earth
Climate study
Data sharing

A B S T R A C T
Recently, it has become easier to access climate data from satellites, ground measurements, and models from various data centers. However, searching, accessing, and processing heterogeneous data from different sources are very time-consuming tasks. There is lack of a comprehensive visual platform to acquire distributed and heterogeneous scientific data and to render processed images from a single accessing point for climate studies. This paper documents the design and implementation of a Web-based visual, interoperable, and scalable platform that is able to access climatological fields from models, satellites, and ground stations from a number of data sources using Google Earth (GE) as a common graphical interface. The development is based on the TCP/IP protocol and various data sharing open sources, such as OPeNDAP, GDS, Web Processing Service (WPS), and Web Mapping Service (WMS). The visualization capability of integrating various measurements into GE extends dramatically the awareness and visibility of scientific results. Using embedded geographic information in the GE, the designed system improves our understanding of the relationships of different elements in a four-dimensional domain. The system enables easy and convenient synergistic research on a virtual platform for professionals and the general public, greatly advancing global data sharing and scientific research collaboration.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Climate anomalies, especially climate extremes, impact significantly humans, the environment, and the economy (Kunkel et al., 1999; Changnon, 2000, 2005; Mendelsohn and Smith, 2002). On the other hand, certain human activities, such as deforestation, increased land use, and pollution, have become significant factors affecting the climate. Asian monsoons play an important role in the global atmospheric circulation system and the global climate (Wang et al., 1991; Shen and Lau, 1995; Wang and Ding, 2005). To understand how human activities affect Asian monsoons and, in turn, the global climate, an international Monsoon Asia Integrated Regional Study (MAIRS) (http://www.mairs-essp.org/) was initiated in 2003. The Atmospheric Data and Services Center of the Nanjing University of Information Science & Technology (NADSC), China, in collaboration with the NASA Goddard Earth Science Data and Information Service Center (GES DISC), is a MAIRS data center. Its atmospheric, ocean, and ecosystem data support Chinese and international research programs in studying land cover/land use changes, monsoons, and climate variations associated with human activities.

The current data service at NADSC is mainly via HTTP that allows users to download data onto their local server. The data archived at NADSC are limited. The massive and multiple format satellite remote sensing data are archived at other data centers. With current internet accessing speed, it is a time-consuming task to download all desired data. Moreover, it is expensive to store and distribute all datasets. Therefore, interoperability with remote data servers, for example, NASA GES DISC, and open data sharing with other data centers, is an urgent problem to be resolved.

The goal of designing a system is to provide a single platform that displays information clearly and effectively by means of graphics through accessing data and images from local and remote services. Many important geophysical variables, such as temperature, pressure, precipitation, vegetation index, and aerosols, are needed for climate studies. To better understand the climate and the climate change, it is beneficial to convert the numerical data, measured from meteorological stations and satellites or derived from numerical models, into images, and to make them accessible through a uniform platform. In general, it requires huge efforts for users to download, process, and analyze...
the datasets in order to obtain the final graphics for scientific research. Also, acquiring and processing of satellite remote sensing data require special skills and sufficient computer hardware and software supporting system. Thus, there is a need of a comprehensive platform that provides images for climate study with ease of use to the end user.

Open-source Project for a Network Data Access Protocol (OPeNDAP) provides a software framework that simplifies all aspects of scientific data networking, allowing uniform and simplified access to remotely sited data. GrADS Data Server (GDS) is another data server that allows interactive and easy access, processing, and visualization of the Earth science data through GrADS (Grid Analysis and Display System) (Doty, 1995).

The Open Geospatial Consortium (OGC) has been developing open standards and protocols for a consortium of GIS vendors, government agencies, academic institutions, and private entities. It has proposed a number of standards for geospatial data exchange, with the intention of promoting interoperability with the use of Web services as OGC Web Services (OWS). Several OWS specifications are available and practically implemented in many organizations to allow public community accessing huge datasets and processing graphics from geospatial data (Lee and Reichardt, 2005). For example, selected satellite and model data are available through OWS services at NASA GES DISC. In addition, HTTP/FTP protocol and HTTP protocol-based Giovanni-WMS service (Acker and Leptoukh, 2007; Berrick et al., 2009) are available for providing remote data and images.

Google Earth (GE), a virtual globe, provides the capability of integrating satellite imagery, aerial photography, and digital map data into a three-dimensional interactive virtual template of the world. There is renewed hope that every sort of information on the Earth and its environment, from the levels of toxic chemicals to the incidence of disease, will become available through OWS services at NASA GES DISC. In addition, HTTP/FTP protocol and HTTP protocol-based Giovanni-WMS service (Acker and Leptoukh, 2007; Berrick et al., 2009) are available for providing remote data and images.

2. System architecture

The goal of the system is to provide a platform for managing and visualizing vast volumes of distributed, heterogeneous, observational, and model data and images for climate study. In order to satisfy the requirements of diverse data sources and multiple users, the design of the system must use a loosely coupled architecture. As shown in Fig. 1, the system has a service-oriented architecture, containing three layers: data layer, business layer, and presentation layer. In the diagram, the NADSC local services are those in the cyan colored solid boxes, while the remote services are those in the purple colored dotted boxes.

2.1. Data layer

The data layer consists of information about data, data format, and data sources that are stored in a relational database. The local or remote archived data can be point data (e.g., station measurement) and grid data (e.g., satellite Level 3 data or model output). Most station data are in ASCII format, while NASA Satellite and model data are stored in a number of formats, such as HDF, HDF-EOS, and NetCDF. It is extremely useful to serve vast volumes of diverse data through a single server. However, the development of a flexible framework to bring multiple data sources with different formats into a heterogeneous data management system is a challenge. Fortunately, many satellite and model data are accessible online through interoperable services. Therefore, only station data including geographic information have been downloaded to the local server while the other data at different interoperable remote services are not downloaded. For the latter, metadata of a product, such as data source, data format, temporal coverage, and service type, are recorded in the database. The data layer provides the data source information to the business layer.

2.2. Business layer

The business layer is the core part of the system, including three components: data services, data access and analysis, and data integration. The task of this layer is to access data from local and remote services, process, and integrate them into a uniform platform.

2.2.1. Data services, access, and analysis

Many remote data centers provide data via a Uniform Resource Locator (URL), such as HTTP, FTP, SOAP (Simple Object Access Protocol), or HTTP-based application protocols, such as OPeNDAP, GDS, WCS, and WMS. In this prototype, the system accesses remote data services located at the NASA GES DISC. The satellite and model data archived at the NASA GES DISC are accessible through services such as OPeNDAP, GDS, and WMS. Users are able to access the data through URL for processing without downloading data.

A relational database-PHP application has been developed in-house at NADSC. It serves ground station data. The station data were populated into a MySQL database and managed using PHP applications nested with SQL commands. The current system has employed the following three methods to access data and to provide images to the end users through a single interface.

(a) Generation of images dynamically on local server

Images of station data are generated dynamically on the local data server. Software has been developed with PHP5.0 as a...
Web Processing Service (WPS) for accessing, processing, and rendering station data. Fig. 2 illustrates the workflow of the WPS. When a user clicks on a weather station on the GE, it submits a request through URL. The Web browser sends the user's request to the PHP application. On receiving the request, the PHP application extracts monthly meteorological data from the database, processes the data, and produces images on-the-fly with the aid of the PHP interpreter and Apache server. The generated images are sent back to the Web browser immediately for displaying. For any selected station, two sets of images are generated: monthly climatology and monthly anomaly. In generating monthly climatology images, PHP applications query monthly climatology of pre-selected parameters from the database, such as precipitation, temperature, and wind, process and render data, and then send images back to the client's browser. For any parameter, the monthly anomaly values are calculated on-the-fly using the monthly value minus the climatology value of the same month. Linear least squares regression of the monthly anomaly is performed. On monthly anomaly plots, in addition to curves of anomaly time series, the linear trend line of each parameter and regression coefficients, such as slope and intercept, are displayed. Compared to satellite and model data, the data volume of station data is relatively small. The performances of on-the-fly data access, process, and generate images are great. Images are returned to the client's browser within only a few seconds. With the current design, the system is very easily expandable by adding more stations and new parameters.

(b) Pregeneration of images at local server
Theoretically, images of data at remote servers can be generated dynamically by accessing data through online

Fig. 1. Service-oriented architecture of the platform. Local services (solid boxes) are in NADSC; remote services (dotted boxes) are in NASA GES DISC. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Fig. 2. Workflow of the PHP application for generating images on the local server.
services. However, the data volumes of satellite and model data are huge. With current internet accessing speed at NADSC, on-the-fly data access to remote data servers will be very slow, which is not practical to an online system. Thus, images were pregenerated using OpenGrADS, GDAL, and WMS applications. For this, OpenGrADS, GDAL, and WMS packages were installed at NADSC. Application software has been developed to access remote data, process, and generate images for remote services. If the remote data are served through GDS, OpenGrADS application software is called; if the remote data are served through FTP, GDAL application software is called; if the remote data are served through OPeNDAP, WMS application software is called. For example, OpenGrADS applications have been used to pregenerate a series of global monthly climatologies by accessing the NASA reanalysis data through GDS at GES DISC. The generated images are stored on the local Web server and their URLs are saved in the database, which will be accessed from a KML file. The KML file has been created dynamically by a PHP application. This way, the seasonal and spatial variations of the climate states, e.g., temperature, pressure, geospatial height, winds, and humidity, are displayed in GE. The images are created with automated jobs that monitor and update images timely whenever the remote data are changed. In the future, when the internet access speed is improved enough, the generation of images can be done on-the-fly and imported to GE directly instead of pregenerating and storing on the local Web server.

(c) Access images from remote WMS services

If a remote data service provides images through their Web server or WMS, the system accesses images directly. For example, NASA GES DISC Giovanni (the NASA Goddard Interactive Online Visualization And Analysis Infrastructure) is a Web-based application that provides a simple and intuitive way to visualize, analyze, and access vast amounts of Earth science remote sensing data (Berrick et al., 2009; Acker and Leptoukh, 2007). The Giovanni fetches data from remote data services through multiple methods, such as, FTP, HTTP, GDS, and OPeNDAP, and generates scientific plots on-the-fly. The Giovanni provides images via the WMS protocol as well. The URLs of generating images via WMS could be imported directly to a KML file.

2.2.2. Data integration

The URLs of accessing data or images from local or remote services have been populated into the database. The integration of different resources is finished in a KML file. Geospatial elements such as Document, Folder, Placemark, Icon, MultiGeometry, 3D Models, GroundOverlay, ScreenOverlay, Network, Textual descriptions, TimeSpan, and TimeStamp are defined by KML Specific tags.2 Folder, Document, Placemark, GroundOverlay, NetworkLink tags, etc., are used in the system as the intermediate media, to manage the climate data. The KML file that integrates URLs are created dynamically by a PHP program, called KML generator.

The current system has employed two methods in KML to visualize data through GE: (a) define geospatial data for managing station information; (b) define URL locations for presenting images. Station data, including station name, station geolocation, and monthly meteorological parameters, are accessed from the database as shown in Fig. 2. The station geolocation and relevant description about the station are extracted from the database and used as the input by the KML generator to produce place-mark and station description. In defining URLs to present images, it could be (1) the image address on the local or remote Web server; (2) HTTP/FTP-based WMS and PHP applications with standard inputs. As described in the previous section, the different URLs based on different protocols are organized in the database that are used to invoke a variety of images generated from station data, satellite, and model data. The generated KML statements are accessed through GE to display station information and images from multiple sources that achieves the purpose of integration of different information including images.

2.3. Presentation layer

The presentation layer is the part simply to access the generated KML file. The KML file can be opened in GE or other virtual Earth tools, such as Google Map. All collected data and images are listed on the left panel, and a globe on the right panel. Detailed data information is described in the next section. When a user clicks on a parameter for a selected month, an image will be displayed on the globe. Multiparameters could be displayed at the same time by overlaying one on top of others.

The designed platform renders the data and images from local and remote sources for end users at a single entry point that enables the global and regional climate research using data from multiple sources more efficiently. It has enabled large volume Earth science data sharing ideas with diverse networks, heterogeneous data storage, and can be accessed by multiple users simultaneously.

3. Data

The current system has integrated three types of climate data. This section describes the data sources, preprocessing methods, and how they are integrated into the system.

3.1. Global monthly climatology data

The global monthly climatology presents the seasonal and spatial variations of the climate states. The current system has integrated a number of parameters, including temperature, pressure, geospatial height, winds, humidity, precipitation, and aerosols.

The global monthly climatology images of atmospheric parameters such as pressure, temperature, wind, and humidity are generated from NASA atmospheric reanalysis datasets, Modern Era Retrospective-Analysis for Research and Applications (MERRA) in the period of January 1979–December 2008 by accessing the data at the NASA GES DISC through the GDS service using OpenGrADS software. The images were pregenerated and saved on a local Web server. MERRA is a project of the NASA Global Modeling and Assimilation Office at NASA Goddard Space Flight Center (GSFC). The reanalysis has been done using the major new version of the Goddard Earth Observing System Data Assimilation System Version 5 (GEOS-5), covering the time period of the modern era of satellite remotely sensed data, from January 1979 to the present. The atmospheric assimilation focuses specially on the hydrological cycle on a broad range of weather and climate time scales (Bosilovich, 2008). The monthly climatology values have been calculated by averaging the same months in 30 years. During this process, no data were downloaded to local machine from remote data servers.

The monthly global precipitation climatology images are accessed directly from the NASA GES DISC MAIRS project through the Giovanni-WMS service. The monthly precipitation data in Giovanni is the satellite-gage merged product (Adler et al., 2003) from the NASA Global Precipitation Climatology Project (GPCP) (http://precip.gsfc.nasa.gov/). The monthly climatology values...
have been computed using all available months of GPCP version 2.1 in the period of January 1979–December 2008. The data in Giovanni have a spatial resolution of 1.0° × 1.0°, regredded from the original 2.5° × 2.5° dataset using GrADS with a box averaging method.

Similar to the monthly global precipitation climatology, the images of global aerosol optical depth climatology are accessed directly from the NASA Goddard GES DISC MAIRS project through the Giovanni-WMS service. The aerosol optical depth climatology in the Giovanni system was generated from the MODIS-Terra monthly aerosol optical depth at 550 nm for the period of January 2001–December 2008. The spatial resolution of the data is 1.0° × 1.0°.

3.2. Measurements from meteorological stations

Time series of monthly meteorological data for the latest 30 years (1980–2009) from 194 WMO basic ground stations over China have been downloaded from National Meteorological Information Center, China Meteorology Administration (NMIC/CMA) (http://cdc.cma.gov.cn/). Parameters are extracted from the product SURF_CLI_CHN_MUL_MON_CES, including monthly mean of daily averaged/minimum/maximum surface air temperature, monthly precipitation, days of precipitation, wind speed, daily sunshine hours, etc. Stations that were migrated to a different location during the study period have been dropped and the final dataset contains 183 stations. The monthly climatology dataset (SURF_CLI_CHN_MUL_MMON_19712000_CES) has been downloaded from NMIC/CMA as well. For each station, the climatology data were computed from the highly quality controlled monthly data from January 1, 1979 to September 30, 2001. Migration of stations has been considered when computing the climatology data.

The station data are archived at the local server and images are generated dynamically when a station is clicked on the GE interface.

3.3. Regional satellite observations

Another climate dataset in the current system is the regional satellite remotely sensed land surface products, including Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI) from MODIS-Terra. For studying regional land surface processes and climates, the higher resolution MODIS 1-km 8-day LST (MOD11A2.005) and monthly NDVI (MOD13A3.005) data have been downloaded from the USGS Land Processes Distributed Active Archive Center (LP DAAC) (https://lpdaac.usgs.gov/) and processed by NASA GES DISC MAIRS project. Due to large data size, both original products are stored in 10° × 10° tiles. The tiled data have been stitched together and reprojected to the equidistant cylindrical projection for the entire Asian Monsoon Asia region (60°E–150°E, 0°–60°N). The processed data are accessible through FTP. Using local WMS application software, the LST and NDVI images were pregenerated and saved on the local Web service. In the current system, images are created for each summer (June, July, and August) over eastern China (102°E–122.5°E, 21°N–41.5°N) from 2000 to 2009.

4. Synergistic climate analysis capability using GE

The designed system could be used easily to conduct climate studies. This section describes the usage of the system and demonstrates sample applications of the system to study climate and climate anomaly with integrated observations and model data from multiple data servers.

4.1. Usage of the system

The developed system has been installed at NADSC and is currently operational (http://nadsc.nuist.edu.cn/mairsprogram/data/mairs.kml). The system is highly portable. Steps to use the system are as follows: (a) download and install the GE software, if not installed already; (b) download the MAIRS KML file; (c) open the KML file with GE. Once the KML is opened, the GE interface is displayed as shown in Fig. 3. All data are in a folder named “Visualization for Climate Study” under “Temporary Places”. On Clicking “Visualization for Climate Study”, a short description about the project is shown. The menu on the left side lists three data groups: Global Monthly Climatology, Ground Station Observations, and Regional Satellite Observations (summer). The Global Monthly Climatology group consists of eight measurements such as temperature, pressure, and precipitation. Each measurement contains 12 months climatology of one or more geophysical parameters. Under Ground Station Observation group, 183 stations have been considered when computing the climatology data.
stations are listed. Clicking on a station from the list or on a station marked on the globe, a page with the station information such as station name, station ID, location, and altitude, and graphs of ground measurements are displayed, including (1) time series of climatology for surface daily mean/maximum/minimum temperature, precipitation, and wind speed; (2) time series of anomaly for temperature, precipitation, and wind; (3) time series of diurnal temperature. The last group, Regional Satellite Observations (summer), contains two 1-km resolution land surface products, land surface temperature, and vegetation index.

The GE interface enables one to display one or more images from different groups, overlaying one image on top of others. Combining GE-embedded geographical information, the system allows a user to study climatology of an interested region, the spatial and temporal variation of a parameter, and to comprehend relationships among different elements. Synergizing different data from satellites, models, and ground stations increases the ability to access, manipulate, and analyze the relationship between global climate and regional climate change. Two sample scientific scenarios are discussed in the following sections.

4.2. Seasonal variations of the Asian monsoon precipitation

The KML provides the time features such as TimeSpan and TimeStamp, which are very important for visualization of massive historical archives of geospatial data. In this system, TimeStamp is used to show the monthly climate change of a number of meteorological elements. As shown in Fig. 4, the spatial patterns of seasonal variations of the climatology precipitation over the Asian monsoon region can be easily visualized on a global scale. The precipitation increases gradually from the winter (January) to the spring (April), and reaches a maximum in the summer (July) in most monsoon regions. The changes of the monthly precipitation are presented intuitively and dynamically for any interested region as animation. Such visually presented information is helpful for understanding the seasonal variation of climatology.

The seasonal variation of the precipitation can be viewed from the station measurements as well. Station locations are shown by clicking on the “Ground Station Observation”. Zooming in and clicking on a station, for example, Nanjing, a city in eastern China, a page with time series of selected ground measurements is displayed. Fig. 5 shows the time series of climatology at Nanjing, China. The monthly cumulated precipitation varies from about 20 mm in the winter (December–January) to over 180 mm in the early summer (June–July). The latter is associated with precipitation in the Meiyu (Biayu) season over the Yantze River region. Fig. 5 illustrates seasonal changes of temperature and wind speed. The monthly temperature (daily mean, daily maximum/minimum) is the lowest in January and is the highest in July and August, indicating a ~25°C difference between the winter and the summer. Interestingly, the wind speed of this station has no significant seasonal changes.

Associated Asian monsoon circulations can be studied using MERRA climatology data such as wind, pressure, geopotential heights, and humidity. Fig. 6 shows wind vectors overlain on the humidity at 850 hPa for January and July. In January, over East

![Figure 4](image-url)  
**Fig. 4.** Global monthly climatology precipitation during seasons: (a) January; (b) April; (c) July; (d) October.
Asia, the major wind is northwesterly, bringing cold and dry air to east China. Over India, the wind blows from land to ocean. An anticyclone exists over Indian land. The wind directions are reversed in July. Over East Asia, winds at 850 hPa are from the south, bringing warm and moist air from the tropical Pacific Ocean. Similarly, over India, the air is from the southwest, bringing moist air from the Indian Ocean to the land.

In this scenario, the GPCP precipitation images are acquired from Giovanni directly, the station time series plots are generated dynamically on the local server, and the MERRA images are preprocessed and saved on the local server. Accessing data and images from different resources through a single portal, a user can explore data and study relationships between different parameters easily without downloading and processing all the original data. Potentially, many topics can be studied using multiple parameters at different levels and times, e.g., how the intensity of an inversion layer changes under the strong wind shear? When does the convergence of different wind cause precipitation? What are the different monsoon rainfall patterns in different parts of the world and why is that? On this basis, scientists can research and analyze in depth to reveal the relationship between them.

### 4.3. Regional climate change associated with urbanization

Ground station locations are rendered in the GE interface by a PHP application combining the station location data in KML. When a user clicks on a station, a page pops up with the station information and time series plots of selected ground measurements. Taking Beijing, China, station as an example (Fig. 7), the station information such as name, ID, location, and altitude are displayed on the top of the page. The time series plots of monthly climatology and anomalies of temperature, precipitation, wind, etc. from 1980.1 to 2009.12 are displayed next to the station information. The climatology plots show the seasonal variation of climate of Beijing. The maximum precipitation occurs in the warmest July and the minimum in the coldest January (Fig. 7a). The wind is strongest during spring to early summer (March, April, May).

The monthly anomalies at a station can be used to study local climate variations. The plots in Fig. 7 show that surface temperatures over Beijing have increased significantly during the past 30 years. The daily minimum temperature has increased the most, and the daily maximum temperature has increased the least.
Based on linear regression estimation, the daily minimum, mean, and maximum temperatures over Beijing have increased approximately by 0.62, 0.5, and 0.36 °C (10 years)^{-1}, respectively, during 1980–2009 (Fig. 7b), indicating that the minimum temperature has increased more than the maximum temperature. The difference of the warming trends between the daily maximum and minimum temperatures is indicated clearly in changes of the diurnal temperature (Fig. 7c). The diurnal temperature has been reduced by about 0.86 °C during the past 30 years. Both monthly precipitation and wind speed indicate negative trends. As shown in Fig. 7d, the monthly cumulated precipitation has been reduced by about 3.79 mm (10 years)^{-1} and the wind anomaly is dominated by negative values during the past 30 years and it has been reduced by about 0.02 m/s (10 years)^{-1}.

Using data from 305 weather stations, Liu et al. (2004) found that the daily maximum and minimum air temperature over China increased by 0.1278 and 0.3238 °C (10 years)^{-1} between 1955 and 2000. The strong warming trend of the surface air temperature over Beijing since 1980 is more than twice that over China as indicated in Liu’s paper. Beijing has experienced fast development, and the city area has been expanded greatly, especially since the reforms started in late 1978. Such local climate changes are likely associated with urbanization, which is consistent with the results in recent studies about climate effects of urbanization in Beijing (Chu and Ren, 2005).

5. Discussion and conclusion

Due to large data volumes, climate data, especially the satellite remote sensing data, are archived and distributed in a number of data centers. By now, data products at some data centers are accessible online through advanced new data sharing technology such as OPeNDAP, GDS, and WMS. However, datasets, even in the same data center, are not served with the same techniques. It requires a certain amount of time to learn each data access method. Using GE as the interface, we have developed a visual and comprehensive platform prototype that has demonstrated the effective integration of various climate data and images from distributed and heterogeneous data sources. In summary, this prototype has integrated data and images using four methods: (a) preprocessing images by accessing remote data, such as global pressure, temperature climatology, available through GDS and OPeNDAP services if the data volumes are large; (b) accessing images directly such as GPCP precipitation and aerosols from remote WMS services; (c) access local data, process, and create images on-the-fly if the data volumes are small, such as station data through a HTTP-PHP application; (d) accessing data from remote data services and generating images with a local WMS service. The URL addresses of images or HTTP-based applications are stored in a database and a software generates KML files dynamically. The generated KML files are accessed by GE from the user’s Web browser. Theoretically, all images can be generated on-the-fly. However, because of the large data volume, with limited internet access speed at the local service, the data access is slow. Moreover, the calculation of the climatology of a single parameter takes several minutes; the performance will be very poor if done on-the-fly. Thus, in the current prototype, most images generated by accessing remote data services were precreated by automated jobs. Images are updated timely when the remote data are modified.

Traditionally, raw satellite remote sensing data are searched and downloaded manually from the remote data sources and then are processed on the local machine. The developed system, using data online interoperable services such as WMS, OPeNDAP, and GDS, enables the large volumes of satellite and model data on the remote services to be accessed more effectively through Web services. The data collection and management task are reduced dramatically.
Moreover, the robust dynamic on-demand processing assists users to obtain visual results of selected elements via only a few clicks. A user does not need to download and process the data. The resulting maps are encoded in well-known formats that allow any GIS software or OWS server to visualize or do further analysis.

The designed system is high scalable. For example, the data measured from 183 weather stations over China have been implemented. Based on the same principle and approach, data from other weather stations could be integrated into the platform easily. Similarly, through WMS, GDS, OPeNDAP, and PHP applications in the business layer, more satellite observations and modeled data from other data sources could be integrated into the system to promote the scientific research of climate and other aspects of weather, health, and environment. With the use of KML, GE is providing a method for visualizing, integrating, and comparing diverse, heterogeneous data from different data sources, revealing new information and knowledge that would have been hidden otherwise.

The lesson learned while developing the systems is that close collaborations to various remote data services are extremely important. It is necessary to have an automated system to monitor remote services regularly for obtaining the operation status and data version updates information at the remote services in order to have a nonbroken operation system.

The future use of Virtual Globes for scientific research presents both hope and challenges. Horizontal images and vertical images (Chen et al., 2009) can be displayed along with vertical geospatial data from observations. A promising direction for scientific research is to use Virtual Globes as Web browser plug-ins, which make it possible to allow GE being integrated into a web page for facilitating the online analysis and synergy of huge volumes of geospatial data. The other potential improvement would be to better display higher resolution images when zooming into a small region as in Google Earth Engine.

Acknowledgments

This work is supported by NASA ROSES 2008 program (NNH082DA001N-LCLUC), National Natural Science Foundation of China (40901244), Social Commonwealth Research Program of the Ministry of Science and Technology, China (GYHY200806009) and Graduate Innovation project of Jiangsu Province, China (CX10B_288Z). The authors are grateful to Yueqi Bai, Aijun Chen, and Aiguo Han at GMU CSSIS for their technique support and valuable comments during the system development.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2011.09.010.

References