Enhancing Agricultural Geospatial Data Dissemination and Applications Using Geospatial Web Services

Weiguo Han, *Member, IEEE,* Zhengwei Yang, *Member, IEEE,* Liping Di, *Senior Member, IEEE,* Bei Zhang, and Chunming Peng

***Abstract*—There are many important publicly available agricultural geospatial data products for the agriculture-related research, applications, and educational outreach programs. The traditional data distribution method cannot fully meet users' on-demand geospatial data needs. This paper presents interoperable, standard-compliant Web services developed for geospatial data access, query, retrieval, statistics, mapping, and comparison. Those standard geospatial web services can be integrated in scientific workflows to accomplish specific tasks or consumed over the Web to create value-added new geospatial application by users. In addition, this paper demonstrates, via real world use cases, applications of those services and potential impacts on facilitating geospatial Cropland Data Layer (CDL) retrieval, analysis, visualization, dissemination and integration in agricultural industry, government, research, and educational communities. This paper also shows that the geospatial Web service approach helps improve the reusability, interoperability, dissemination, and utilization of agricultural geospatial data. It allows for integrating multiple online applications and different geospatial data sources, and enables automated retrieving and delivery of agricultural geospatial information for decision-making support.**

***Index Terms*—Cropland Data Layer (CDL), CropScape, geospatial web service, geospatial data sharing and interoperability, service chain.**

I. Introduction

AGRICULTURAL geospatial data products (such as irrigation data, soil survey data, and cropland data) are widely used in many research areas, ranging from agricultural sustainability, food security, and biodiversity, to natural resources monitoring, disaster assessment, carbon accounting, bioenergy, etc. The original geospatial data can be obtained from Web sites, such as Geospatial Data Gateway (GDG) of the United States Department of Agriculture (USDA), in a traditional manner of "searching—selection—ordering—downloading—processing." This time-consuming data disseminating process does not provide a direct channel for users to retrieve and customize geospatial data and information at different spatial scales for agricultural assessment and analysis [1].

The International Organization of Standardization (ISO), the Open Geospatial Consortium (OGC), and other international organizations have published a series of geospatial information standard specifications that have been widely adopted by the geographic information science community to provide open geospatial data and processing services. However, the agricultural community has not applied these harmonized specifications adequately in agricultural geospatial data distribution, and has not exploited their benefits in the sharing and interoperability of geospatial data and services. The Research and Development Division of the National Agricultural Statistics Service (NASS) and the Center for Spatial Information Science and Systems (CSISS) of George Mason University jointly developed an online geospatial application, named CropScape (http://nassgeodata.gmu.edu/CropScape/), which allows users to navigate, customize, visualize, and analyze on-demand Cropland Data Layer (CDL) data interactively and intuitively [2]. This Web application successfully utilized the geospatial standards in agricultural geospatial data distribution and analysis in considering the huge benefits of those standards.

CropScape has been extensively utilized by policy and decision makers, scientists, researchers, educators, and farm producers in research and operational applications since its official release in January, 2011. However, the ever increasing demand from users of the agricultural and other communities for providing on-demand CDL data and enhanced analytical capacities in an open and interoperable environment have to be addressed. The enhancements need to support machine-to-machine and man–machine interactions over the Web for processing, analyzing, and disseminating CDL data. More importantly, there is a growing demand to leverage these CDL data, information, and services automatically in third party geospatial applications, geospatial models, or geoprocessing workflows. Therefore, the implemented geospatial Web services need to be further improved to disseminate and analyze the CDL data in an open and standard manner. The functionalities of CropScape have been described in details [2]. This paper will focus on the CropScape's geospatial Web services and workflows and their applications.

The remainder of this paper is organized as follows. Section II introduces the related works. The geospatial Web services for the CDL data is described in Section III. Section IV gives examples of consuming these services in various cases. Section V discusses these services applications and future work, and summarizes the conclusions.

Manuscript received November 18, 2013; revised February 05, 2014; accepted March 19, 2014. Date of publication April 30, 2014; date of current version January 06, 2015. The work of L. Di was supported by the United States Department of Agriculture, National Agricultural Statistics Service (Grant 58-3AEU-0-0067).

W. Han, L. Di, and B. Zhang are with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030 USA (e-mail: whan@gmu.edu; ldi@gmu.edu; bzhang4@gmu.edu).

Z. Yang is with the United States Department of Agriculture, National Agricultural Statistics Service, Research and Development Division, Spatial Analysis Research Section, Fairfax, VA 22030 USA (e-mail: zhengwei_yang@nass.usda.gov).

C. Peng was with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030 USA. She is now with ESRI, Redlands, CA 92373 USA (e-mail: pg.janie@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2315593

II. Related Works

A. Geospatial Web Services

Service Oriented Architecture (SOA) has been widely adopted as a modern and agile approach to build composite applications in the distributed environment for its advantages in flexible system development, software components reuse, just-in-time integration, and cross-platform interoperability [3]. SOA offers an innovative approach to integrate the discoverable and interoperable Web services across organizations as the functionality of Web applications [4].

The World Wide Web Consortium (W3C) defines a series of standards of Web services, such as Web Services Description Language (WSDL) and Simple Object Access Protocol (SOAP). These standards have been widely adopted in the business world. OGC has developed numerous specifications for easy access and processing geospatial data to meet and exceed user expectations [5]. Among those specifications, geospatial data services generate the customized geospatial data according to user-specified parameters, and geospatial processing services offer data processing and analysis functions. For example, the Web Coverage Service (WCS) provides standard interfaces to geographical coverage; the Web Map Service (WMS) handles geospatial data rendering; the Web Feature Service (WFS) manipulates feature-level geospatial data creation, updating, and exchange; and the Web Processing Service (WPS) offers standard interfaces for discovery, publishing, and binding to geospatial process.

Geospatial Web services enable users to utilize geospatial data and computing resources available online and to automate geospatial data integration, processing, and analysis [6]. Infrastructure for Spatial Information in Europe (INSPIRE) Data Specifications Special Interest Group released the draft technical guidelines of data specification on agricultural and aquaculture facilities. The ISO and OGC standards were included in the guidelines to achieve interoperability [7]. AgroXML is a standardized language for data exchange between farm management information systems, and has been adopted in many agricultural projects in Germany and the European Union.[1] Agriculture and Agri-Food Canada (AAFC) offers agricultural related maps, geospatial data, and applications, which cover land, soil, water, climate, biodiversity, as well as agricultural census information, on its website.[2] WMS was adopted to serve several geospatial data products (e.g., the 1967 and 2000 Plant Hardiness Zones data), and provided users with graphic pictures of the requested map layer(s). A Web-based decision support tool integrating multiple types of key services was built for monitoring agricultural and ecosystem services in Tanzania [8]. Zhao *et al.* tried to build a prototype geospatial Web portal framework using OGC geospatial standards to make the agricultural statistics data discoverable, retrievable, and analyzable in the geographic context [9]. A geoprocessing service approach was proposed to integrate distributed geographical information services [10]. Soil Web was built to distribute the digital soil survey products for the states of California, Arizona, and Nevada in multiple types of interface, including WMS and WFS services, and Web service Application Programming Interface (API) for soil survey information query [11]. Although many agricultural geospatial data portals have been built, most of them are not designed from the data and GIS analyst's perspective [12].

B. CDL Data

Most agricultural geospatial data products, such as the CDL data, are public information, and are provided freely for social, academic, and scientific purposes. The CDL data, derived from mid-resolution satellite data, Common Land Unit data, administrative data, and annual agricultural survey data, provides the specific crop and other land cover classifications encompassing the Contiguous United States (CONUS) in a geo-referenced raster dataset [13]. This valuable geospatial cropland product has been produced by NASS since 1997, and covered all 48 conterminous states and the District of Columbia from the crop year of 2008. The CDL data has been used as an important data source in agricultural related research fields, such as pesticide control [14], land cover monitoring [15], biomass monitoring [16], crop rotation [17], bioenergy crop inventory [18], and carbon accounting [19].

Previously, the CDL data could be downloaded only at the state level in a compressed form from the NASS website or GDG from the USDA Natural Resources Conservation Service (NRCS). To obtain the CDL data for a study area (e.g., a watershed or an agricultural region), geospatial data processing has to be performed on the original files using the traditional desktop GIS software. The entire process requires GIS software and expertise, and is time consuming. A new data dissemination application CropScape was proposed to streamline the process of accessing, disseminating and using the CDL data.

C. CropScape

Funded by NASS, CropScape was developed to provide CDL users with useful functions of online visualization, geospatial navigation and querying, reformatting and transformation, delineation of area of interest, on-the-fly data analysis, on-demand data processing and dissemination, thematic map creation, etc, in an SOA environment. CropScape helps CDL users avoid the burden of installing geospatial visualization and analytical software or tools, and also allows them to access, visualize, retrieve, and analyze the CDL data at any geographic level through an intuitive user interface. According to our statistics, thousands of users from federal, state, and local governments, companies, academic institutions, and nonprofit organizations have visited this Web application. CropScape revolutionizes the traditional CDL data distribution channels, and greatly improves CDL data publishing, accessing, dissemination, and applications. CropScape was listed as one of the "*highlights of Agency Open Government IT Accomplishments that improve citizen engagement*" in the FY 2011 Report to Congress on the Implementation of The E-Government Act of 2002 [20].

In CropScape, WCS, WFS, and WMS were configured to support geospatial data retrieval and rendering in a standard and interoperable manner. The operations of these geospatial data services were called on either the server side or client side of

[1]AgroXML: http://www.agroxml.de.

[2]Agriculture and Agri-Food Canada: http://www.agr.gc.ca/index_e.php.

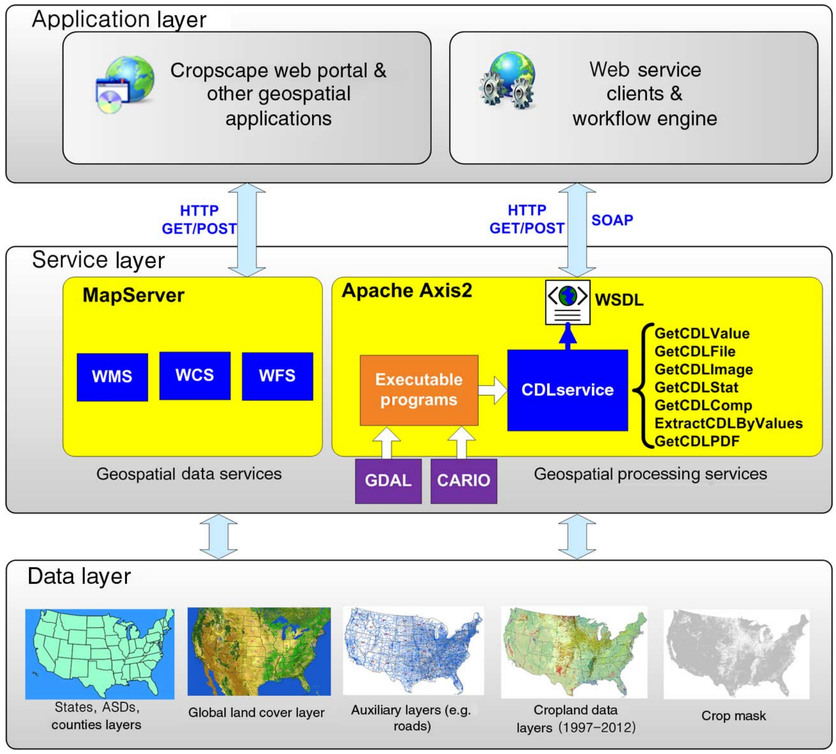

Fig. 1. Overall CropScape architecture (improved from [2]).

CropScape [2]. Users required that more map layers were served by these standard services, so that they could easily leverage them in other geospatial applications or desktop GIS software like ArcGIS. For example, the function for fetching CDL data within any specified area, particularly the area defined by the uploaded boundary file, was highly demanded to be exposed as standard Web services to support location-specific agricultural research and applications. Moreover, other useful analytical functions such as statistical calculation, change analysis, and thematic map creation were increasingly demanded by users for using these functions in other applications or even dynamic models [21]. The details of how to meet these requirements are given in Section III.

III. Geospatial Web Services of CropScape

A. System Architecture

The general architecture of CropScape is shown in Fig. 1. A detailed description of the architecture can be found in reference [2]. Here, only the components related to the geospatial Web services are described.

Following standard data encodings and service interfaces, open source MapServer software was used to serve geospatial raster and vector data in the services of WCS, WFS, and WMS in consideration of its popularity, high-performance, and robustness. W3C Web service was adopted in the implementation of Web geospatial processing services, owing to the fact that it has been widely used as an industry standard. Apache Axis2 was chosen as the runtime environment for these processing services for its efficient, modular, configurable, reliable and secured framework [22]. Geospatial Data Abstraction Library (GDAL) and two-dimensional (2-D) graphic Library Cairo were utilized to develop executable programs; these programs were then wrapped as a standard Web service named CDLService using Axis2 API to avoid redundant programming.

TABLE I
Map Layers List

Type	Name	Description	Service
Raster	CONUS CDLs	The annual CDL files for the CONUS from 1997 to 2013	WMS, WCS[b]
	State CDLs	The available annual CDL files for each CONUS state	
	Crop mask	The crop mask file derived from the 1997–2013 CONUS CDL files	
Vector	Boundaries	The boundary files of agricultural regions, states, ASDs, and counties for the whole CONUS and states, ASDs, and counties for each CONUS state[a]	WMS, WFS[b]
	Roads	The national freeway system and regional major highways	
	Waters	The rivers and lakes at the national and regional level	

[a]Created from the latest 2013 TIGER/Line Shapefiles (http://www.census.gov/geo/maps-data/data/tiger-line.html).
[b]Currently, only the projection of USA Contiguous Albers Equal Area Conic (USGS version) and World Geodetic System (WGS) 84 (EPSG:4326) are supported by these data services.

TABLE II
CDLSERVICE OPERATIONS LIST

Operation	Description	Inputs	Outputs
GetCDLValue	Query which crop was planted at the specified geographic location	Year, x, and y coordinates[a]	Crop value, category name, and color value
GetCDLFile	Generate the CDL data for the specified year and area	▪ Year and district code(s) ▪ Year and bounding box ▪ Year and multiple points coordinates ▪ Year and link of a feature vector file[b]	Link of the customized CDL file
GetCDLImage	Produce the images for the specified CDL file(s)	Link(s) of the CDL file(s), and output format (PNG or KML)	Link(s) of the output image file(s) or KML file
ExtractCDLByValues	Extract crops of interest from the CDL file and export them in a new raster file	Crop values and link of the CDL file	Link of the output raster file
GetCDLStat	Calculate crop acreage according to pixel counts and spatial resolutions	▪ Year, district code(s), and output format[c] ▪ Year, bounding box, and output format ▪ Year, multiple points coordinates, and output format ▪ Year, link of a feature vector file, and output format	Link of the output statistical result in the specified format
GetCDLComp	Compare the CDL files by pixel values to detect land cover change	Same as above	Link of the text file and the raster file of the output change result
GetCDLPDF	Create the CDL thematic map in PDF format	▪ Year, district code(s), paper size, and title ▪ Year, bounding box, paper size, and title ▪ Year, multiple points coordinates paper size, and title ▪ Year, link of a feature vectorfile, paper size, and title	Link of the PDF file of the thematic map

[a]All coordinates in the inputs must be in the projection of USA Contiguous Albers Equal Area Conic (USGS version).
[b]The feature vector file in the inputs can be a GML file or a compressed ESRI Shapefile (the. shp,. shx,. dbf, and. prj files must be contained and compressed without directory path in a. zip file).
[c]The output text formats include CSV, JSON, and TXT.

B. Geospatial Data Services

In the initial version of CropScape, two raster datasets (CONUS CDL and global land cover) were served using WMS and WCS, and all boundary layers (state, Agricultural Statistical District (ASD), and county) were served using WMS and WFS. These data services were deployed on a single server for both CropScape application and outside applications. It was found that the CropScape's performance and reliability were not acceptable. Moreover, the CDL data for each CONUS state and other contextual map layers (such as roads and rivers) were required to be included in these OGC services. To guarantee system performance, two high powered servers were set up for CropScape operational use and the MapServer was configured on both servers to handle raster data and vector data, respectively. In addition, two additional servers were used for outside applications use with the same configurations as those for CropScape use.

The detailed information of map layers provided in CropScape is given in Table I. WMS and WCS services were configured to serve not only the CDL files of the whole CONUS, but also the CDL files of each available individual CONUS state. Users can choose the services for the state of interest to them. Moreover, the crop mask layer was also served by WMS and WCS services. Users can apply the crop mask to focus their data exploration and analysis on the agricultural areas only. Similarly, the boundary layers on a CONUS and state scale can be accessed from WFS and WMS services. The CDL data and other map layers at a specified state can be explored in Web browser at the CropScape URL with state abbreviation (e.g., http://nassgeodata.gmu.edu/CropScape/IA). The road and water layers served by WMS and WFS were added as auxiliary layers to locate the specific area of interest (AOI). The detailed service description and request examples of these layers can be found in the Developer Guide of CropScape [23]. Users can utilize these services in OGC compliant software to retrieve the data layers for analysis and supplement visualization by overlaying them with other data layers.

C. Geospatial Processing Services

WCS *GetCoverage* request and WMS *GetMap* request return only grid data or images within a specified spatial extent defined by a bounding rectangle. The study area in most cases is normally an administrative or ecological region, such as a watershed area with an irregular shaped boundary. To facilitate CDL data applications and to better serve CropScape customers,

TABLE III
MAJOR CROP CATEGORIES IN THE CDL DATA

Value	Name	Color (R,G,B)	Value	Name	Color (R,G,B)
1	Corn	255,211,0	24	Winter wheat	165,112,0
2	Cotton	255,38,38	25	Other small grains	214,158,188
3	Rice	0,168,229	26	Dbl crop WinWht/Soybeans	113,113,0
4	Sorghum	255,158,12	27	Rye	173,0,124
5	Soybeans	38,113,0	28	Oats	160,89,137
6	Sunflower	255,255,0	29	Millet	113,0,73
10	Peanuts	112,165,0	30	Speltz	214,158,188
11	Tobacco	0,175,76	31	Canola	209,255,0
12	Sweet corn	221,165,12	32	Flaxseed	127,153,255
13	Pop or orn corn	221,165,12	34	Rape seed	209,255,0
14	Mint	127,211,255	35	Mustard	0,175,76
21	Barley	226,0,124	36	Alfalfa	255,165,226
22	Durum wheat	137,99,84	37	Other Hay/non alfalfa	165,243,140
23	Spring wheat	216,181,107	38	Camelina	0,175,76

a geospatial processing service, CDLService, was developed and extended to disseminate and analyze customized cropland and other land cover information in a more efficient and more flexible manner, and make machine to machine interactions possible.

CDLService provides seven operations on query, retrieval, statistics, mapping, and comparing of the CDL data, including *GetCDLValue*, *GetCDLFile*, *GetCDLImage*, *ExtractCDLByValues*, *GetCDLStat*, *GetCDLComp*, and *GetCDLPDF*. The WSDL of CDLService can be found at http://ws.csiss.gmu.edu:8080/axis2/services/CDLService?wsdl. The descriptions, input parameters, and output parameters of existing operations are listed in Table II.

These operations were wrapped from the corresponding executable programs that were built from GDAL and Cairo APIs and utilities. For example, an executable program was built to capture the specified crop types or crop changes (e.g., from corn to soybean) and export them in a raster file using the operation of *ExtractCDLByValue* (The major crops information in the CDL data is listed in Table III). In this program, GDAL API functions for opening the raster file, getting the dataset information, fetching the selected band and reading the raster data are first called to retrieve data information from the CDL file. Next, GDAL CreateCopy() function is called to create a new writeable raster file with the collected information, such as resolutions, color table, projection, etc. Then, the specified crop values are filled in the corresponding grid cells of the output rater file using GDAL raster I/O functions. Finally, the link of the new raster file will be returned in a response message by the operation of *ExtractCDLByValues*.

The performance of the CDLService once was one of the biggest issues for real world operation. To improve the CDLService performance, file caching was used in implementing CDLService. The CDL files, statistics results (in JSON, CSV, or TXT format), preview images, and KML files for the specified area with defined code in a given year are generated and saved

```
<?xml version='1.0' encoding='utf-8'?>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <ns1:GetCDLStat xmlns:ns1="http://cropscape.csiss.gmu.edu/CDLService/">
      <year>2011</year>
      <fips>19003,19029</fips>
      <format>CSV</format>
    </ns1:GetCDLStat>
  </soapenv:Body>
</soapenv:Envelope>
```

(a)

```
<?xml version='1.0' encoding='utf-8'?>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/">
  <soapenv:Body>
    <ns1:GetCDLStatResponse xmlns:ns1="http://cropscape.csiss.gmu.edu/CDLService/">
      <returnURL>
        http://nassgeodata.gmu.edu/nass_data_cache/CDL_2011_clip_352118264.csv
      </returnURL>
    </ns1:GetCDLStatResponse>
  </soapenv:Body>
</soapenv:Envelope>
```

(b)

Fig. 2. SOAP request and response example.

with unique filenames identified by year and code on the server when the first user requests data information from CropScape portal or CDLService service client. For the same data requested later by user(s), no repetitive processing will be performed and these file links will be returned promptly.

Users can consume CDLService by sending standard HTTP GET, HTTP POST, or SOAP requests. They can enter a simple HTTP GET request like http://nassgeodata.gmu.edu:8080/axis2/services/CDLService/GetCDLStat?year=2012&fips=19&format=csv in a Web browser, to get the quick answers of the questions, such as "What were the top three land cover categories in Iowa in 2012?" An example of SOAP request and response is shown in Fig. 2. It requests for crop statistics of two neighboring counties of Iowa, Adams County (FIPS code: 19003) and Cass County (FIPS code: 19029) in 2011, and export the statistical

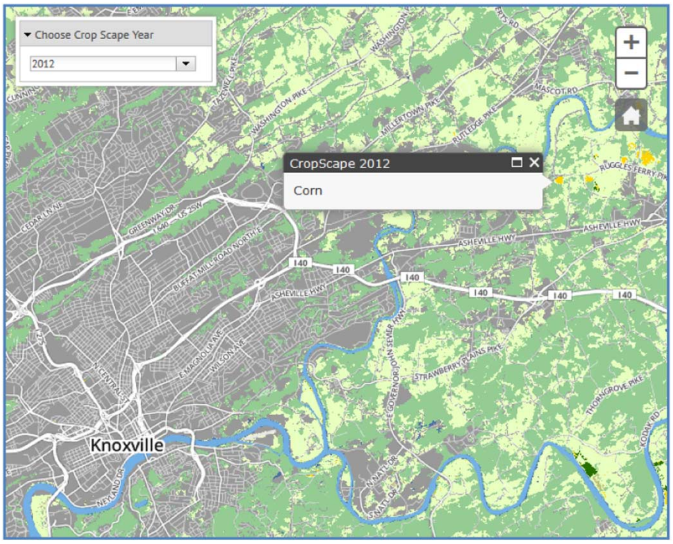

Fig. 3. Example of integrating CDL web services.

results in a CSV file. More examples can be found from the CropScape Developer Guide [23]. It should be indicated that these web services and analytical functions can be integrated with other geospatial applications, or be composed in a workflow to accomplish the specialized tasks.

IV. Use Cases

Geospatial Web services for CDL data and processing provide agricultural and other communities as well as the general public reliable and accessible interfaces to cropland information and useful analysis services. The following real world examples illustrate the practical usefulness of these services.

A. Geospatial Data Mash-Up

One successful application example is the use of CDL data service in the Bioenergy Knowledge Discovery Framework (KDF, https://bioenergykdf.net/). KDF lists the CDL WMS service as one important supporting dataset in the category of land cover of its library so that the CDL layer can be loaded with other bio-energy layers (e.g., bio-refinery locations) in the KDF online system for decision making support with regard to development options for biomass feedstock production and bio-refinery infrastructure [2]. Another example is integrating the CDL WMS layers and query service in the Plan East Tennessee (PlanET) project, a five-county (Anderson, Blount, Knox, Loudon, and Union) regional plan. A Web map application using ArcGIS JavaScript API is being constructed to integrate social, economic, and natural resources related data layers together in WGS84 Web Mercator Projection, as shown in Fig. 3. The CDL WMS layers for Tennessee are leveraged in the application to monitor changes in farming related land uses. *GetCDLValue* is called to display land cover category at the clicked point. The point coordinates in Web Mercator Projection are sent to *Project* operation of ArcGIS *Geometry* service and transformed to ones in USA Contiguous Albers Equal Area Conic (USGS version) Projection. Then the converted coordinates, along with the specified year, are delivered to *GetCDLValue* service in a PHP program. Next, the service response is parsed and the land cover category at the point is retrieved. Finally, the formatted information is displayed in a popup message box.

These CDL geospatial data services have been registered in Components and Services Registry (CSR) system of the Global Earth Observation System of Systems (GEOSS), and contributed to the development of the operational global agriculture monitoring and analysis system [24]. In the Fifth Phase of GEOSS Architecture Implementation Pilot (AIP), the CDL WCS and WMS services were utilized with other OGC data services for administrative districts, elevation data, and drought indices from various sources in the Agriculture Societal Benefit Area (SBA) demonstration of a real scenario of drought impacts on crops in the United States in 2012.[3]

The CDL WMS service has been collected in Spatineo's directory. The Spatineo (http://www.spatineo.com/) is a Web application of monitoring and reporting quality and performance of geospatial Web service. The Spatineo's latest monitoring report shows that the CDL WMS has a superior quality of service with 99.5% availability during the year of 2013.

B. Geospatial Processing Service Applications

CDLService was developed as a standard Web service for geospatial processing. It provides operations of cropland information query, customization and retrieval, statistics and reporting, change analysis, and thematic map creation. These operations are openly available and can be easily integrated with other geospatial analytical and simulation applications as part of product life cycle, and/or be composed in a workflow with other Web services to accomplish customized specific tasks.

1) Geospatial Modeling: Researchers at the Agricultural System Research Unit of the USDA Agricultural Research Service (ARS) requested to integrate crop statistical information with other inputs in a Web-service based tool to detect crop rotation patterns over years in several ecosystem response units (ERUs). To generate crop rotation input files for the component-based AgroEcoSystem-Watershed (AgES-W) model, which is a spatially distributed agroecosystem model, they developed a web service based Java tool called the Crop Rotation and Management Builder (CRMB) [21]. In CRMB, *GetCDLStat* of CDLService was first invoked to obtain main crop information for the specified years and the uploaded vector files of ERUs. *GetCDLFile* was then requested to get the CDL file for these ERUs. The CDL raster files were reclassified according to the ground truth data of crops. Next, the classified crops with the largest area for a specified year were selected to detect crop rotation for each polygon unit from the list of possible crop rotations such as corn-soybean or continuous corn cropping. The crop rotations were then mapped to the correct information in the Land Management and Operation Database (LMOD) of USDA NRCS using the LMOD web services. Finally, the LMOD information combined with crop information from CDLService were constructed into the input

[3]AIP-5 agriculture societal benefit area demo: http://www.youtube.com/watch?v=_zxJu1H3mPk#t=123.

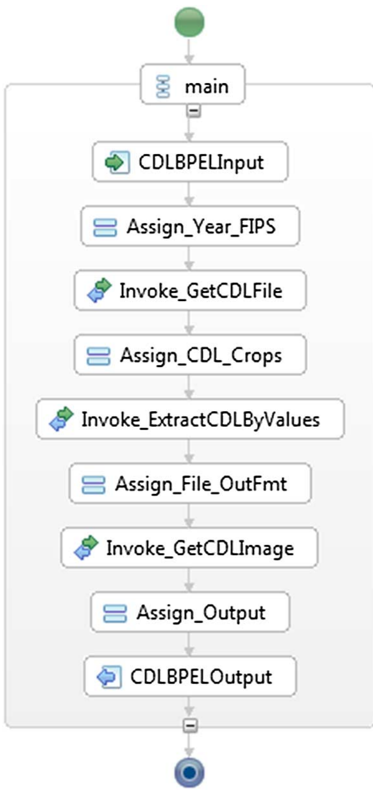

Fig. 4. Workflow design.

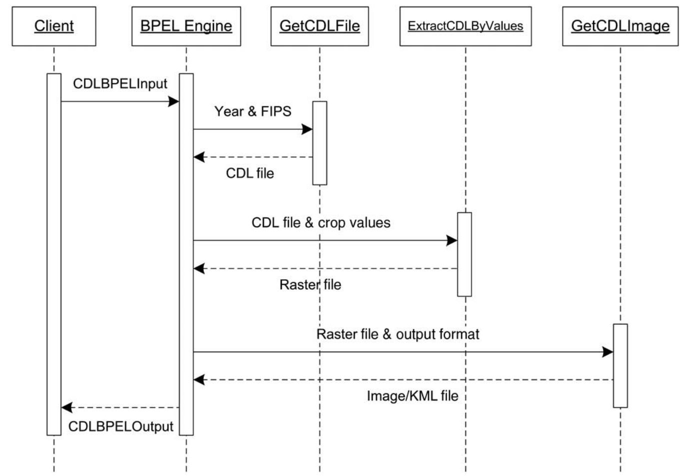

Fig. 5. Sequence diagram.

files of the AgES-W model. The proposed CRMB tool has been tested successfully in the Scott Field test site in eastern Colorado.

2) Scientific Workflow: Web services can be composed in a workflow with the Business Process Execution Language (BPEL) for Web Services. BPEL engine creates a new instance of process when receiving the SOAP message and starts to invoke each service in the process based on its WSDL [25]. The process defined in the BPEL is described in WSDL. Therefore, the composed workflow can be exposed as a new Web service. CDLService can be easily integrated in the automated process by complying with these Web service standards. As a demonstration of BPEL application, a use case of generating crop distribution map is given as follows.

Eclipse BPEL Designer was used to compose the related operations of CDLService in a workflow. This workflow was designed as a simple synchronous BPEL process. The components of actions were dragged and dropped on the diagram of workflow as shown in Fig. 4. And the sequence diagram of information flows is illustrated in Fig. 5.

CDLBPELInput (year, FIPS code, crop values, and output format) was created to handle the input parameters of the whole process. Year and FIPS code of *CDLBPELInput* were first assigned as the input variables of the *GetCDLFile* operation. Next, the CDL file link exported from the *GetCDLFile* operation and crop values of *CDLBPELInput* were assigned as the input variables of the *ExtractCDLByValues* operation, and this operation's output and the *CDLBPELInput* output format were then assigned as the input variables of the *GetCDLImage* operation. Finally, the output link of the *GetCDLImage* operation was assigned to *CDLBPELOutput*, the output parameter of the whole process. A new Web service of generating crop distribution map is consequently created from this service chain. This process can be deployed on workflow engine like Apache Orchestration Director Engine (ODE), or BPELPower that supports both W3C and OGC Web services [26], [27]. A SOAP message with year (here is 2012), FIPs code (i.e., 19163), crop values (1 means corn), and output format (i.e., KML) was configured to produce the 2012 corn distribution map of Scott County, Iowa, in a KML file from this process, as shown in Fig. 6(a).

U. S. Drought Monitor (http://droughtmonitor.unl.edu/) offers weekly drought information across the United States and Puerto Rico in multiple output formats, like ESRI Shapefile, GML, and KMZ. Drought data in Iowa on August 21, 2012 from US Drought Monitor can also be processed with a workflow and the results in KMZ can be exported into Google Earth. The workflow process generated Iowa crop data in KMZ file can be downloaded and exported into Google Earth with the above KML file to display which corn planting areas were affected by the 2012 Midwest drought [as shown in Fig. 6(b)] and convey useful information for further assessment and analysis.

V. Discussion and Conclusion

The geospatial Web service approach not only extends the reach of agricultural geospatial data products, but also promotes analytical and exploratory functionality related to agricultural research. The standards-based geospatial Web services can be easily discovered, accessed, and integrated by other geospatial applications to exploit the useful agricultural geospatial data products, and make agricultural geospatial data products such as the CDL data truly open and interoperable.

High-quality, efficient, and effective geospatial Web services for agricultural geospatial data are highly expected by end users. Taking the CDL data as an example, this paper presents the detailed information on geospatial data and processing services along with their practical applications to demonstrate its potential impacts on improving geospatial cropland information retrieval, analysis, visualization, dissemination, and integration. These standard geospatial services can be commingled with services from other sources to build new applications, or be used as one

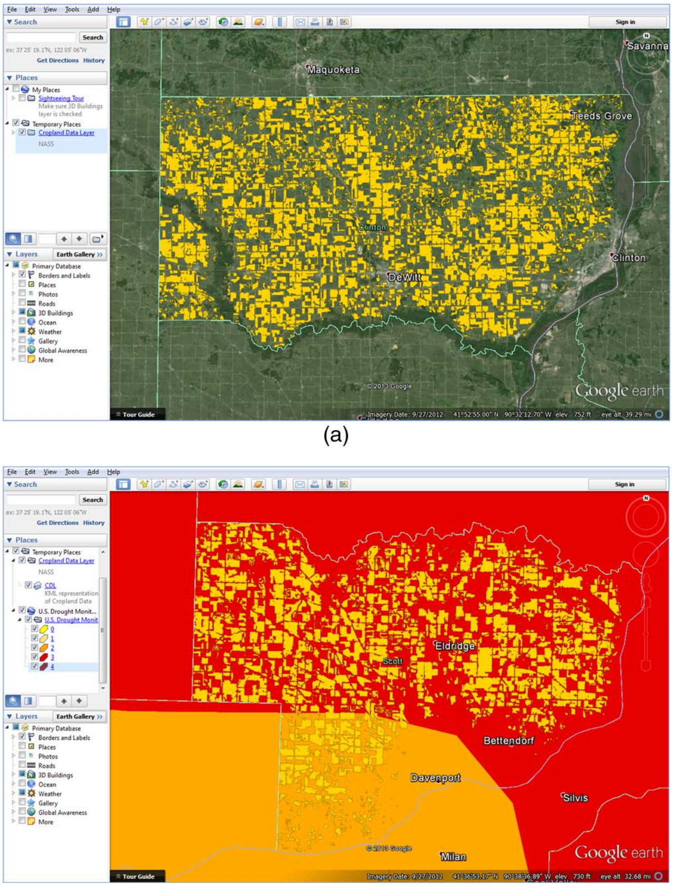

Fig. 6. 2012 (a) Corn distribution and (b) drought map in Scott County, Iowa.

part of a complex model in a workflow to accomplish particular tasks. As shown in this paper, using workflow to chain Web services together will greatly enhance the capability of Web service to accomplish complicated geospatial data processing, visualization, and analysis. It allows for integrating multiple online applications and different geospatial data sources via Web services.

The limited capacity of current servers and networks has become the bottleneck for further promoting CropScape along with its geospatial Web services in large scale real-time applications, especially for the concurrent requests of retrieving and analyzing geospatial data with large areas of interest. The feasibility of migrating them to a powerful server cluster or a cloud computing environment is being assessed. In addition, the lightweight RESTful Web services for geoprocessing are being developed to meet users' requirements and expectations. A general framework derived from CropScape for agricultural geospatial data products distribution and analysis is under development. Furthermore, publishing CDL data as a resource oriented service according to the principles of RESTful services and Linked Open Data is ongoing [28].

In summary, the geospatial Web service approach facilitates the interoperability among different Web Applications, enables automated delivery of agricultural geospatial data and analytical services to the agricultural community and general public, and facilitates better utilization of geospatial information in agricultural related decision making.

Acknowledgment

The authors thank the associate editor and two anonymous reviewers for their valuable comments and suggestions on this manuscript.

References

[1] Y. Yang, L. T. Wilson, J. Wang, and X. Li, "Development of an integrated cropland and soil data management system for cropping system applications," *Comput. Electron. Agric.*, vol. 76, no. 1, pp. 105–118, 2011.

[2] W. Han, Z. Yang, L. Di, and R. Mueller, "CropScape: A web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support," *Comput. Electron. Agric.*, vol. 84, pp. 111–123, 2012.

[3] N. M. Josuttis, *SOA in Practice*. Sebastopol, CA, USA: O'Reilly, 2007.

[4] H. Nezhad, B. Benatallah, F. Casati, and F. Toumani, "Web services interoperability specifications," *Computer*, vol. 39, no. 5, pp. 24–32, 2006.

[5] G. Percivall, L. Menard, L. Chung, S. Nativi, and J. Pearlman, "Geo-processing in cyberinfrastructure: Making the web an easy to use geospatial computational platform," in *Proc. 34th Int. Symp. Remote Sens. Environ.*, 2011.

[6] P. Zhao, L. Di, W. Han, and X. Li, "Building a web-service based geospatial online analysis system," *IEEE J. Sel. Topics Appl.*, vol. 5, no. 6, pp. 1780–1792, Dec. 2012.

[7] Infrastructure for Spatial Information in Europe (INSPIRE). (2013). *Data specification on agricultural and aquaculture facilities—Draft technical guidelines*. Available: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AF_v3.0rc3.pdf

[8] E. H. Fegraus, I. Zaslavsky, T. Whitenack, J. Dempewolf, J. A. Ahumada, and K. Lin *et al.*, "Interdisciplinary decision support dashboard: A new framework for a Tanzanian agricultural and ecosystem service monitoring system pilot," *IEEE J. Sel. Topics Appl.*, vol. 5, no. 6, pp. 1700–1708, Dec. 2012.

[9] P. Zhao, L. Di, W. Han, and Z. Yang, "Building geospatial web portal for use of national agricultural statistics in geographic context," in *Proc. 17th Int. Conf. Geoinformat.*, 2009.

[10] S. Shi and N. Walford, "Automated geoprocessing mechanism, processes and workflow for seamless online integration of geodata services and creating geoprocessing services," *IEEE J. Sel. Topics Appl.*, vol. 5, no. 6, pp. 1659–1664, Dec. 2012.

[11] D. E. Beaudette and A. T. O'Geen, "Soil-web: An online soil survey for California, Arizona, and Nevada," *Comput. Geosci.*, vol. 35, no. 10, pp. 2119–2128, 2009.

[12] J. J. Kerski and J. Clark, *The GIS Guide to Public Domain Data*. Redlands, CA, USA: ESRI Press, 2012.

[13] C. Boryan, Z. Yang, R. Mueller, and M. Craig, "Monitoring US agriculture: The US department of agriculture, National Agricultural Statistics Service, Cropland Data Layer program," *Geocarto Int.*, vol. 26, no. 5, pp. 341–358, 2011.

[14] J. B. Belden, B. R. Hanson, S. T. McMurry, L. M. Smith, and D. A. Haukos, "Assessment of the effects of farming and conservation programs on pesticide deposition in high plains wetlands," *Environ. Sci. Technol.*, vol. 46, no. 6, pp. 3424–3432, 2012.

[15] M. C. Hansen and T. R. Loveland, "A review of large area monitoring of land cover change using Landsat data," *Remote Sens. Environ.*, vol. 122, pp. 66–74, 2012.

[16] V. Chandola and R. R. Vatsavai, "A scalable gaussian process analysis algorithm for biomass monitoring," *Stat. Anal. Data Mining*, vol. 4, no. 4, pp. 430–445, 2011.

[17] A. J. Stern, P. C. Doraiswamy, and E. R. Hunt, "Changes of crop rotation in Iowa determined from the United States Department of Agriculture, National Agricultural Statistics Service Cropland Data Layer product," *J. Appl. Remote Sens.*, vol. 6, no. 1, pp. 063590-1–063590-16, 2012.

[18] C. Wang, F. B. Fritschi, G. Stacey, and Z. Yang, "Phenology-based assessment of perennial energy crops in North American Tallgrass Prairie," *Ann. Assoc. Amer. Geogr.*, vol. 10, no. 4, pp. 742–751, 2011.

[19] T. O. West, C. C. Brandt, L. M. Baskaran, C. M. Hellwinckel, R. Mueller, C. J. Bernacchi, V. Bandaru, B. Yang, B. S. Wilson, G. Marland, R. G. Nelson, D. T. Ugarte, and W. M. Post, "Cropland carbon fluxes in the United States: Increasing geospatial resolution of inventory-based carbon accounting," *Ecol. Appl.*, vol. 20, no. 4, pp. 1074–1086, 2010.

[20] U.S. Office of Management and Budget. (2012). *FY 2011 report to congress on the implementation of the E-Government Act of 2002*. Available: http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/fy11__e-gov_act_report.pdf

[21] H. Kipka, O. David, J. Lyon, L. A. Garcia, T. R. Green, J. C. Ascough, II, and K. Rojas. (2013). *A web-service tool to generate crop rotation management input files for spatially distributed agroecosystem models*. Available: http://hydrologydays.colostate.edu/Papers_13/Kipka_paper.pdf

[22] Apache Software Foundation. (2012). *Apache Axis2 user's guide*. Available: http://axis.apache.org/axis2/java/core/docs/userguide.html

[23] CropScape—Cropland Data Layer. (2012). *CropSpcae developer guide*. Available: http://nassgeodata.gmu.edu/CropScape/devhelp/help.html

[24] L. Di and G. Yu. (2012). *Agriculture scenario engineering report—GEOSS Architecture Implementation Pilot (Version 0.5)*. Available: http://www.ogcnetwork.net/pub/ogcnetwork/GEOSS/AIP5/documents/AIP5Docs/SBA/Agriculture/AIP_SBA_ER_Agriculture_0.5.doc

[25] G. E. Yu, P. Zhao, L. Di, A. Chen, M. Deng, and Y. Bai, "BPELPower—A BPEL execution engine for geospatial web services," *Comput. Geosci.*, vol. 47, pp. 87–101, 2012.

[26] N. Chen, L. Di, G. Yu, and J. Gong, "Automatic on-demand data feed service for AutoChem based on reusable geoprocessing workflow," *IEEE J. Sel. Topics Appl.*, vol. 3, no. 4, pp. 418–426, Dec. 2010.

[27] G. Yu, L. Di, B. Zhang, and H. Wang, "Coordination through geospatial web service workflow in the sensor web environment," *IEEE J. Sel. Topics Appl.*, vol. 3, no. 4, pp. 433–441, Dec. 2010.

[28] M. Roth and A. Bröring. (2013). *Linked open data in spatial data infrastructures*. Available: https://wiki.52north.org/pub/Projects/GLUES/2012-09-10_LoD_SDI_White_Paper_MR_AB.pdf

Weiguo Han (M'09) received the B.S. degree in applied mathematics from Tianjin University, Tianjin, China, in 1996, the M.Eng. degree in computer science and engineering from Huazhong University of Science and Technology, Wuhan, China, in 2002, and the Ph.D. degree in cartography and geographic information system from Chinese Academy of Sciences, Beijing, China, in 2005.

Currently, he is a Research Assistant Professor with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA. His research interests include geospatial data sharing and interoperability, Semantic Web, Web GIS, geospatial cyber-infrastructure, spatial analysis and modeling, land use and land change, vegetation dynamics, and remote sensing image processing and analysis. His research achievements have been extensively leveraged by the scientific community users and the general public around the world. DEM Explorer system developed by him has been adopted and reused by NASA Land Processes Distributed Active Archive Center to distribute ASTER GDEM to users worldwide. He also implemented the online system named CropScape for USDA National Agricultural Statistics Service to disseminate, visualize, query, and analyze Cropland Data Layer data. He has published more than 70 research papers in peer-reviewed academic journals, books, and conference proceedings.

Dr. Han won the 63rd Annual Secretary's Honor Awards of USDA for his significant contribution to CropScape in 2011.

Zhengwei Yang (M'02) received the Bachelor's degree in electrical engineering from Shanghai Science and Technology University, Shanghai, China, in 1982, the Master's degree in systems engineering from Shanghai Jiaotong University, Shanghai, China, in 1985, and the Ph.D. degree in electrical engineering from Drexel University, Philadelphia, PA, USA, in 1997.

He is an IT Specialist with Research and Development Division, National Agricultural Statistics Service. His research interests include remote sensing methods and GIS technology and their application in agriculture. His work spans conceptual, theoretical, and application research as well as system development.

Dr. Yang is the recipient of the 2011 USDA Secretary's Honor Award for Excellence.

Liping Di (M'01–SM'05) received the B.Sc. degree in remote sensing from Zhejiang University, Hangzhou, China, in 1982, the M.S. degree in remote sensing/computer applications from the Chinese Academy of Science, Beijing, China, in 1985, and the Ph.D. degree in geography from the University of Nebraska–Lincoln, Lincoln, NE, USA, in 1991.

He was a Research Scientist with the Chinese Academy of Science from 1985 to 1986 and with the NOAA National Geophysical Data Center from 1991 to 1994. He served as a Principal Scientist from 1994 to 1997 and a Chief Scientist from 1997 to 2000 with Raytheon ITSS. Currently, he is a Professor of geographic information science and the Director of the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, USA. His research interests include remote sensing, geographic information science and standards, spatial data infrastructure, global climate and environment changes, and advanced Earth observation technology. He is the Chair of ISO 19130: Imagery Sensor Models for Geopositioning.

Dr. Di is a Member of American Geophyscial Union, American Association of Geographers, and American Society of Photogrammetry and Remote Sensing. Since 2005, he has been the Chair of the Data Archiving and Distribution Technical Committee of IEEE Geoscience and Remote Sensing Society.

Bei Zhang received the B.Sc. degree in computer software from Beijing Institute of Technology, Beijing, China, in 2000.

She worked with Taiji Computer Company, Beijing, China, from 2000 to 2007, first as a Junior Software Engineer and then as a Senior Software Engineer. Currently, she is a Research Associate of remote sensing and geographic information systems with the Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA. Her research interests include sensor Web and Web GIS.

Chunming Peng received the B.E. degree in computer science and technology from Zhejiang University, Hangzhou, China, in 2008, the M.S. degree in geographic and cartographic sciences from George Mason University, Fairfax, VA, USA, in 2010, from where she will receive the Ph.D. degree in earth system and geoinformation sciences by May 2014.

In 2009, she worked as a Graduate Research Assistant with the Center for Spatial Information Science and System, George Mason University. Her research interests include agricultural drought, Web-based GIS, image processing, and software testing.