

# DEVELOPMENT OF A DISTRIBUTED SERVICE-ORIENTED GEOGRAPHIC INFORMATION SYSTEM FOR ECOLOGICAL RESOURCES QUANTITATIVE MEASUREMENT

Yansheng Dong; Xiufang Zhu; Yaozhong Pan; Hongping Chen; Jinshui Zhang

State Key Laboratory of Earth Surface Processes and Resource Ecology (Beijing Normal University); College of Resources Science & Technology, Beijing Normal University, Beijing, 100875. China - (goodboy, zxf ,pyz, zhangjsh, )@ires.cn

Commission WG IV/5

**KEY WORDS:** GIS Web Services; NPP; SOA; ArcGIS Server; Satellite Remote Sensing; Workflow

## ABSTRACT:

Service-oriented architecture supports data and model sharing and interoperability. The concept of such a service-oriented and distributed architecture is applied to build and integrate web-based service models of ecological resources quantitative measurement. The paper is to develop and integrate a GI web service by using an algorithm that calculates net primary productivity of terrestrial vegetation (NPP) of low spatial resolution satellite images. The system employs ArcGIS Server, which is a mature business platform, as the geographic information server to offer basic service functions of GIS and construct NPP web service. To achieve the overall result, two separate web services are developed by using the geographic information services specification for reference: a) A web map services, which handles the acquisition and manipulation of the satellite images. b) Service which implements NPP calculation algorithm using results from the first web service. The prototype which is developed by using ArcGIS Server is to integrate NPP model web services and publish NPP final results.

## 1. INTRODUCTION

Service-oriented architecture supports data and model sharing and interoperability (J.Horak, et al, 2007). Service providers, service agents and service requestor are three roles in the service-oriented architecture. Service providers register their web services and services description information. According to their business needs, the services requestors need to find the appropriate services in the service agent. By authorizing from service agent the requestors can bind these services and use them. Service agent manages the registration centre, public these services that have been registered and provides search services for requestors to find the necessary services easily (J.Bustos et al, 2003). In the application of geographic information system, it is necessary to integrate heterogeneous data and models of the existing algorithms to build professional models. Adopting service-oriented architecture is for the greatest degree of interoperability and share. The GIS web services initiative is mainly driven by the Open GIS Consortium (OGC) (OGC, 2004). ALL of these service-oriented concepts have been adopted by OGC. OGC has provided specifications for many web based applications like the OpenGIS Web Map Server and data descriptions like GML (Geography Markup Language). [4] (J.Horak, et al, 2007)

In some areas of application, there are many sophisticated models and algorithms, to format the corresponding spatial information model it requires a combination of geographic information, such as CA and its combination of remote sensing to study of urban land use change, distributed hydrological model and so on (Chunyang He, et al, 2005; Chunyang He, et al, 2005; Jin Chen, et al, 2002; Freeze R A, et al, 2007; ). These spatial information models require distributed computing, packaging, distribution and sharing. GI web services provide a simple but powerful way for addressing this task (Kotzinos D, et al, 2003). In a service-oriented architecture environment, these services can not only be a combination of individual

services, but also can be used alone. At the same time, in order to resolve the complex issues or parts of a complex problem, it can be consummated by a shared network services and it can also be used in a separate stand-alone environment. The advantage of GI web services is that it can be remote invoked by multi-requestors concurrently in the non-interference circumstances and by using the service, a complex task can be distributed to a number of nodes in order to reduce the strength of computer calculations(N. R. Brisaboa1, et al, 2007) . Service-oriented architecture can integrate data and model through loosely coupled approach. When the remote model services have been updated, the client does not need to update, thus reducing the burden of frequent updates, making updating and maintenance easily. When the complex operation, only the server computing needs, every step of calculating the data do not spread to the client, thus reducing the pressure on the network bandwidth (N. R. Brisaboa1, et al, 2007) .

Using GIS and remote sensing for ecological resources quantitative measurement, the most important is to obtain NPP data, which will involve a multi-source data and complex computing model, as required by the model parameters, all kinds of input data and complex data grid computing process(Wenquan Zhu, et al, 2004) . On the one hand, the need for the massive data management, sharing and integration; On the other hand, data is updated quickly, and NPP model is relatively stable, this will require data and model of loosely coupled, rather than tightly coupled, such a conducive to development of software reuse and sharing. The paper used a service-oriented architecture to integrate NPP model as an example, to achieve the quantitative measurement of ecological resources of the web-based GIS.

The remaining sections of this paper are organized as follows. Section 2 introduces NPP model algorithms and data pre-processing workflow. Section 3 designs the architecture of distributed geographic information system for integrating the

NPP model and the heterogeneous data based on ArcGIS server. The prototype and its initial applications are stated in Section 4. Conclusions and directions for future work are given in Section 5.

## 2. THE NPP MODEL

All within the study area of green vegetation in one year, NPP is one of the key indicators of using remote sensing technology for ecological resources quantitative measurement. NPP is a key component of the terrestrial carbon cycle which can reflect the condition of ecological resources in the study regions. By using multi-temporal remote sensing data, meteorological data related to vegetation and the ecological statistics, NPP estimation model has been improved and constructed by Dr. Zhu. This NPP estimation model has been improved from three areas: 1) the impact of NPP estimation by vegetation cover classification accuracy. 2) The use of meteorological data, combined with the existing regional evapotranspiration model to achieve water stress factor estimates, 3) According to the measured data, to simulate the largest solar energy utilization of various types of vegetation. (Wenquan Zhu, et al, 2004; ZHU Wen-quan, et al, 2004; PAN Yaozhong, et al, 2005; Yu Deyong, et al, 2007; ZHU WenQuan, et al, 2007 )

### 2.1 THE NPP ESTIMATION MODEL ALGORITHM

The NPP estimation model algorithm has been described as shown in Figure 1. The key point is to calculate the pixel area of the NPP value, and then calculated by multiplying the total area of the region. By this way we can estimate the region's NPP. In the model the estimation of NPP can be expressed by the two factors which are plant absorbed photosynthetic ally active radiation (APAR) and the actual utilization of solar energy ( $\epsilon$ ). Its estimates formula as follows:

$$NPP(x, t) = APAR(x, t) \times \epsilon(x, t) \quad (1)$$

In the Eq.(1),  $APAR(x,t)$  express absorbed photosynthetic ally active radiation of pixel  $x$  in the month  $t$ .( Units:  $MJ \cdot m^{-2}$ );  $\epsilon(x,t)$  express the actual utilization of solar energy of pixel  $x$  in the month  $t$ .(Units:  $gC \cdot MJ^{-1}$ ).

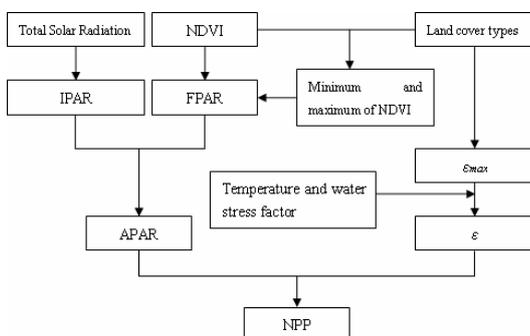


Figure 1 the NPP estimation model algorithm

According to APAR, temperature stress factor, water stress factor, the largest solar energy utilization rate the model can estimates NPP. The comprehensive algorithm of this model has been described in the literature (Wenquan Zhu, et al, 2004;

ZHU Wen-quan, et al, 2004; PAN Yaozhong, et al, 2005; Yu Deyong, et al, 2007; ZHU WenQuan, et al, 2007; ).

Although the calculation model has been able to achieve through IDL, the input data of this model is demanding and data preprocessing is tedious, such as the weather site of discrete points need to be interpolated by Kriging method and generate the meteorological factors raster data which is consistent with NDVI data. In the data pretreatment process, the different software is needed to complete different task. This process requires a number of operators in the computer environment to work together, and all the calculation should be distributed, asynchronous and remote sharing.

### 2.2 DATA-PROCESSING WORKFLOW

The input parameters of this model include a large number of remote sensing data, weather data and vegetation physiological and ecological data. The accuracy of these parameters directly affects the NPP estimates. In order to get precise model input data and parameters, it need to define a precise data-processing workflow for reduce the minimum to a common mistake. (Wenquan Zhu, et al, 2004; ZHU Wen-quan, et al, 2004; ZHU WenQuan, et al, 2007; ).

Data-processing workflow includes four different tasks that are data collection, data pre-processing, data accuracy test and NDVI computing (N. R. Brisaboa1, et al, 2007).

- 1) The data collection main task is to obtain long series of remote sensing images, such as the NOAA/AVHRR, and the classification of vegetation data. The meteorological sites data also need, that weather attribute information includes the monthly average temperature, the monthly total solar radiation, the monthly rainfall. The meteorological sites data should also include longitude, latitude and altitude of the geographic information. Some other necessary basic geographic data, such as 1:400 million of China's administrative boundaries, land use are needed too.
- 2) The Data pre-processing tasks mainly is to preprocess remote sensing image data, such as atmosphere correction, cloud inspection. To achieve request model, the vegetation classification data is need to consolidation and re-record.
- 3) The main task of testing the accuracy of data is to test all generated data to see whether or not to meet the requirements.
- 4) NDVI computing tasks are mainly based NDVI algorithm to generated NDVI data by using images.

## 3. ARCHITECTURE OF DISTRIBUTED INFORMATION SYSTEM

### 3.1 SYSTEM ARCHITECTURE

The distributed service-oriented architecture is a multi-layered, distributed and cross-platform interoperability of systems. Its architecture is shown in Figure 2. It is mainly composed of three parts, namely, the client application, registration services and application services, with the corresponding the services requestor, service agents and service providers in the architecture of web Service. Its implementation process as: through the service agents the service providers register their application web service in registration center and the requestors find appropriate services to bundled and consume them.

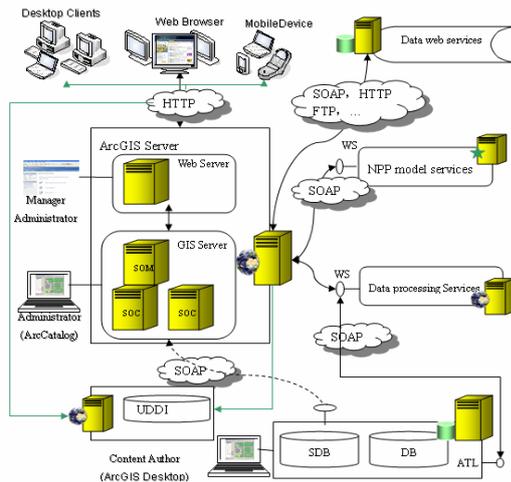


Figure 2 The overall system architecture

The client can be a Web browser, mobile users and ArcGIS Desktop, and so on. The users include decision makers, professional, ordinary users. Authorized users can implement various kinds of data operations and maintenance. It can direct remote access to data and services model, and register all the business web services as well to achieve sharing and interoperability.

The GI web service registration centre is the hub of the whole system. In this centre through geographical metadata it manages distributed geospatial data effectively. At the same time through the web services metadata it can manage distributed services resources. System administrators and authorized users can use various web services in this centre, such as query, retrieval, and maintenance.

The application services mainly include GI web services and all other related business services. The GI web services include the geographic data services and GI processing services that defined by OGC. The main ecological resources measurement GI web services model is the NPP application model services.

### 3.2 Analysis of key technologies

The application server analysis: The application server uses ArcGIS Server9.2 (ESRI, 2004), with the help of web servers communicates with client for handling service requests by the end customers. The application services are existed in the form of remote components, which offer web application service interface (WS-API) to interact with the application server.

The analysis of data server: all data which include NDVI data, meteorological data and so on are stored in the form of product data in the database. The product data is the monthly NPP data, the vegetation coverage data and the NPP per year. The system can use ArcCatalog to manage and update all of these spatial data and can also remotely access data services.

## 4. PROTOTYPE SYSTEM AND RUDIMENTARY APPLICATION

### 4.1 THE PROTOTYPE SYSTEM

The reasons using ArcGIS Server as the application server are the main consideration of convenience and more mature platform to integrated models. Here is the main model to achieve NPP estimation model to generate NPP eventually

products. First of all images web services are created and NPP model can bind these remote images web services to complete the model calculation and publish final NPP results. The whole process is related to publication, discovery, combination and binding of web services, which is actually a chaining of web services.

Using the map web services which offer by ArcGIS Server, the parameters required by NPP model have been established and published. The operator can not only easily select the required data, but also provides a default data, such as static parameters and all of image data by default.

According to the NPP model algorithms, the system can remote access image data by the image web services. The NPP model algorithm services interface is to access to these data and calculate, then produce a series of temporary images. All the calculations are completed in the distributed server-side, which can reduce the necessary time to download image data and reduce network bandwidth to minimize. In addition, to understand the practical significance of image, the image describe information is needed. Metadata can be used here to describe. The desirable approach is to use semantic information to describe these images because the semantic ambiguity can be eliminated (Kotzinos, et al, 2003; F. Fonseca, et al, 2002).

### 4.2 CLIENT APPLICATION

ArcGIS Server provides an ADF framework for the development (ESRI, 2004), which can be easily integrated web services. The prototype system provides a web-based development application. By using ADF development framework, it not only can quickly build and development web applications, but can also focus on the realization of model integration. We have adopted .NET framework to development web applications, and eventually realizes the interface as shown in Figure 3.

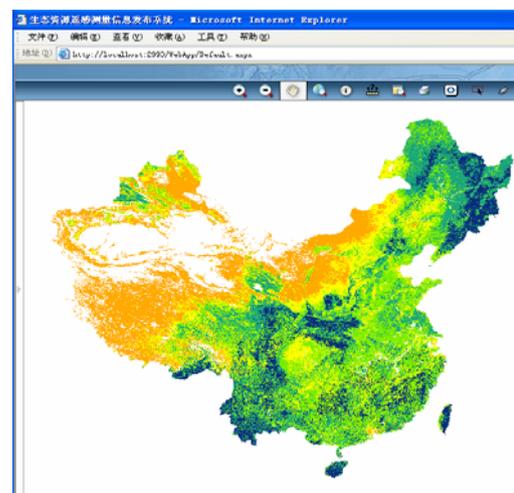


Figure 3 The map browser of NPP

## 5. CONCLUSIONS AND FURTHERWORK

The web service supports the development of web-based on applications. The ADF provided by ArcGIS Server can facilitate integrate the web services. The service-oriented architecture supports single model application, and also supports a complicated combination that consists of multiple services. This

paper introduced the NPP web service as the case. The process of model services is the combination of a series of simple services, which substantially represents a service chain.

The next step of this research focuses on some functions of web service such as auto-description, release and auto-discovery. Meanwhile, the implement scheme of integrated workflow, the interoperability of prototype and other systems are about to be completed in the future work.

#### ACKNOWLEDGEMENTS

The work is supported by The National High Technology Research and Development Program of China (2006AA120101).

#### REFERENCES

Chunyang He, Peijun Shi, Jin Chen, Xiaobing Li, Yaozhong Pan, Jing Li, Yuechen Li & Jinggang Li. 2005, Developing land use scenario dynamics model by the integration of system dynamics model and cellular automata model, *Science in China. Series D*, 48(11):1979-1989

Chunyang He, Qiaofeng Zhang, Yuechen Li, Xiaobing Li, Peijun Shi. 2005, Zoning grassland protection area using remote sensing and cellular automata modeling-A case study in Xilingol steppe grassland in northern China. *Journal of Arid Environment*, 63, 814-826.

D. Kotzinos and N. Chrysoulakis. (2003). "Design of GIS Web Services for Environmental Monitoring: Using Satellite Imaging to Calculate Vegetation Indices", *Proceedings of the International Symposium on GIS and Remote Sensing*, Volos, Greece, 7-9 November 2003.

ESRI. *Server Developers Guide*[R]. 2004,

Fonseca, F., M. Egenhofer, P. Agouris, and G. Camara. 2002. "Using Ontologies for Integrated Geographic Information Systems", *Transactions in GIS*, 6(3), 2002.

Geography Markup Language (GML) 2.0, <http://www.opengis.net/gml/01-029/GML2.html>

J. Horak, A. Orlik, J. Stromsky. 2007, "Web services for distributed and interoperable hydro-information systems",

*Hydrol. Earth Syst. Sci. Discuss.*, 4, 1879-1891, 2007

J. Bustos, K. Watson. 2003, "Beginning .NET Web Services using C#", Tsinghua University Press, Beijing, China

Jin Chen, Peng Gong, Chunyang He, Wei Luo, and Masayuki Tamura, Assessment of urban development plan of Beijing by using CA-based urban growth model *Photogrammetric Engineering and Remote Sensing*. 2002, 68(10): 1063 - 1073.

N. R. Brisaboa<sup>1</sup>, J.A. Coteló-Lema<sup>1</sup>, A. Fariña<sup>1</sup>, M. R. Luaces<sup>1</sup>, J. R. Paramá and J. R. R. Viqueira, Collecting and publishing large multiscale geographic datasets, *SOFTWARE—PRACTICE AND EXPERIENCE* *Softw. Pract. Exper.* 2007; 37:1319-1348

OGC (Open Geospatial Consortium) official web site <http://www.opengeospatial.org/>

PAN Yaozhong, SHI Peijun, ZHU Wenquan, GU Xiaohu, FAN Yida & LI Jing. 2005, Measurement of ecological capital of Chinese terrestrial ecosystem based on remote sensing, *SCIENCE IN CHINA (Series D)*, 48(6): 786~796

Wenquan Zhu, Yaozhong Pan, Haibo Hu, Jing Li, and Peng Gong. 2004, Estimating Net Primary Productivity of Terrestrial Vegetation Based on Remote Sensing: A Case Study in Inner Mongolia, China. *Geoscience and Remote Sensing Symposium*, 2004. IGARSS '04. *Proceedings. 2004 IEEE International Volume 1*, 20~24 Sept. 2004 Page(s): 528~531

Yu Deyong, Zhu Wenquan\* and Pan Yaozhong. 2007, The Role of Atmospheric Circulation System playing in Coupling Relationship between Spring NPP and Precipitation in East Asia Area. *Environmental Monitoring and Assessment*, Doi: 10.1007/s10661-007-0023-6.

ZHU Wen-quan, PAN Yao-zhong, HE Hao, WANG Ai-ling. 2006, Spatio-temporal Distribution of Net Primary Productivity Along the Northeast China Transect and its Response to Climatic Change. *Journal of Forestry Research*, 17(2): 93~98.

ZHU WenQuan, PAN YaoZhong, YANG XiaoQiong & SONG GuoBao. 2007, Comprehensive analysis of the impact of climatic changes on Chinese terrestrial net primary productivity. *Chinese Science Bulletin*, 52(23): 3250~3260.