A WEB-BASED SPECTRAL DATABASE FOR ENVIRONMENTAL APPLICATION

Jing Fang[1]*, Wenqing Liu[2], and Haibo Tan[3]

Key Laboratory of Environmental Optical & Technology, Anhui Institute of Optics & Fine Mechanics, Chinese Academy of Sciences, Hefei P. O. box 1125, Anhui, China
*[*1]Email:* fangjing@aiofm.ac.cn
[2]Email: wqliu@aiofm.ac.cnl
[3]Email: hbtan @aiofm.ac.cn

ABSTRACT

The administration and storage of environmental characteristic spectral data are highly relevant in many fields of environmental study, such as measurement of trace gases in the atmosphere and air quality estimation. For this reason, a web-accessible database has been developed, offering ready access to the main parameters of molecular absorption spectral data. Web-based and friendly interfaces allow for interactive queries as well as previews of plots and downloads of files of the resulting spectral data for thorough comparative analyses.

Keywords: Spectrum, Web-accessible, Database, Molecular absorption, Environmental monitor

1 INTRODUCTION

The optical absorption characteristics of trace gases in the atmosphere, which exist mainly in the form of molecular and particulate matter, can be regarded as their "fingerprint" (Nölle & Pätzold, 2004). The administration and storage of this environmental characteristic spectral information are key factors for a great variety of study fields: pollution gas identification, concentration of trace gases, discovery of unknown pollutants, air quality assessment, and so on. The availability of spectral data also supports further research in atmospheric pollution detection. The majority of existing spectral data are organized in file-based systems, such as the HITRAN standard atmospheric molecular spectroscopic database (NASA EOS) (Grant, Kagann, & McClenny, 1992) and the UVVIS spectra of atmospheric constituents (AUC DLR) (Leach, 2000). Most of these systems provide a logical organizational structure. Unfortunately, experience shows many inconveniences and limitations of this approach, especially for large amounts of data: slow sequential access, loss of data integrity and complexity, and strong coupling between data structure and programs. This approach also makes it difficult for investigators to do thorough, comparative analyses of the physically distributed data that comes from different experimental environments and instruments. In this paper, the design process of a web-based spectral database for environmental applications is presented. It allows for interactive queries about the main parameters of environmental spectral information, including line-by-line positions and molecular absorption cross-sections as well as previewing plots and downloading the resultant spectral data files. An underlying relational database management system ensures the independence of the file format and establishment of relationships among data as well as supporting an attribute-based search capability, common interfaces, and scalability.

2 METHODOLOGY

2.1 Web structure

An on-line spectral database has been designed for the environmental researcher to offer and share reference spectra and data. Considering that the resulting data requires some scientific calculation before being sent to the user and that these calculations consume certain CPU resources, a B/C/S (Brower/Client/Server) structure is implemented. With this, calculations and processing of requests can be undertaken by the client. This design has proven to be extensible, universal, and easy use. At the initial stage, a B/S structure can be built when the task is not too heavy. Along with the increase of system tasks, the primary B/C structure can be extended to B/C/S smoothly without modifying existing distribution and economizing costs. The structure is shown in Figure 1.

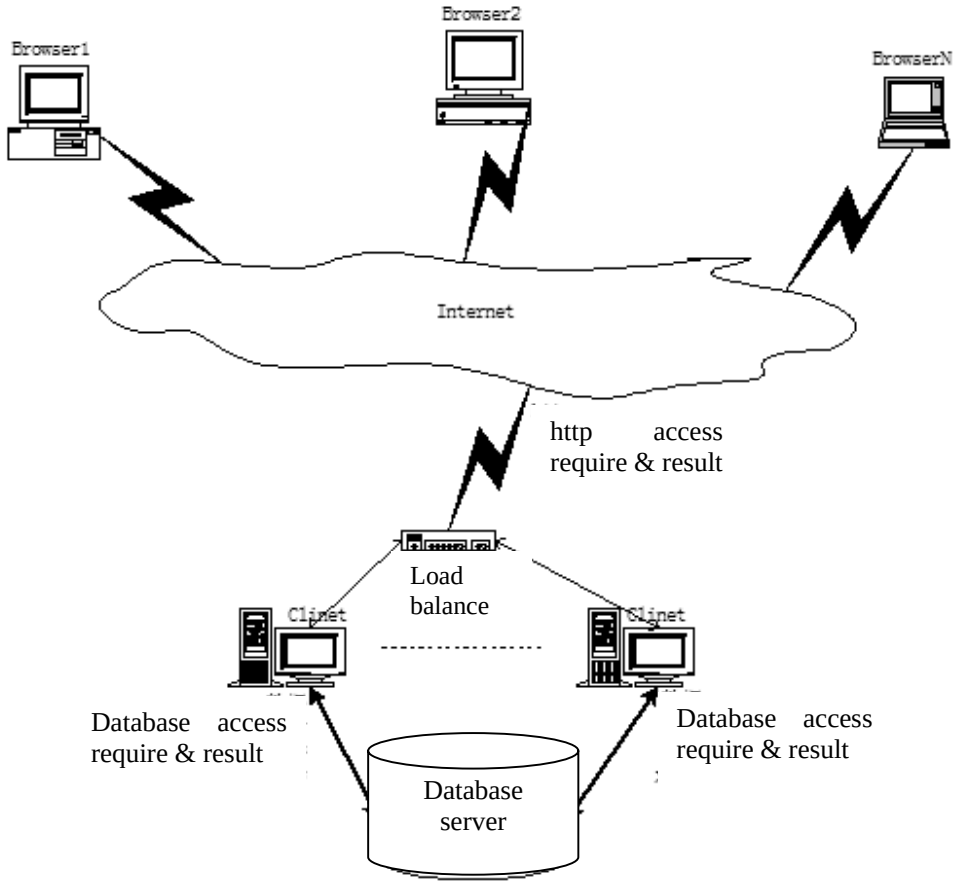

Figure 1. B/C/S structure of database.

The web link to databases is generally established by a programming interface that enables communication between a programming language and a web server. ASP and some extended common gateway interfaces (CGI) are chosen to interpret the structured query language (SQL) from database queries because of speed of implementation and ease of maintenance. Users supply queries by means of HTML forms, and these parameters are sent in standard HTTP to the web server, which calls the ASP program to interpret the request and establish a connection to the database. The query results provided by the database are encoded in HTML documents and returned to the client. Then users can select data sets according to their query for run-time-generated plots display in their local browser. Plotting is technically realized by a java script program that generates graphics. In practice, the IIS (Internet Information Server) is chosen as the web distribution platform with the RDMS SQL-SERVER 2000, which is particularly suited for reliable storage and administration of large amounts of data.

2.2 Function module

The web-based spectral database for environmental applications includes the following functions:

1) Management of user information: Administration of user registration, including setting or modifying password and user name.

2) Management of substance information: Maintain and query the main parameters of substances and their code table.

3) Maintenance of spectral data: Administration of characteristic spectral data collected from different bands and deal with routine processes, such as adding, deleting, and modifying data files.

4) Query of spectral data: Query characteristic spectral data collected from different molecular bands and supply multi-index queries. For example, users can search for absorption spectra using molecular parameter key words. It is also possible to search for all molecular absorption spectral data using optical source detection

key words.
5) Display of spectral data: Show the two-dimension plot on the client's browser to allow customizing the style of plots and legends.
6) Analysis of spectral data: Display two or more spectral data plot concurrently for comparative analyses.
7) Download and print of spectral data: Supply functions for downloading and printing according to users' requirements.

2.3 Data model

The data model is the logical structure design of the relational database. The entity–relationship diagram established by Erwin (Computer Associates) is a classical design and analysis tool that can help confirm entities, attributes, and relationships of data (Rothman & Barbe, 2003). It ensures the integrity of the data, access control, consistency, the recovery of data in case of hardware failure, effective retrieval of datasets by SQL statements and meets the requirements of 3NF (Third Normal Form). Furthermore, it can generate SQL code by the forward engineering construct function in SQL-SERVER. The ER diagram of the database is shown in Figure 2.

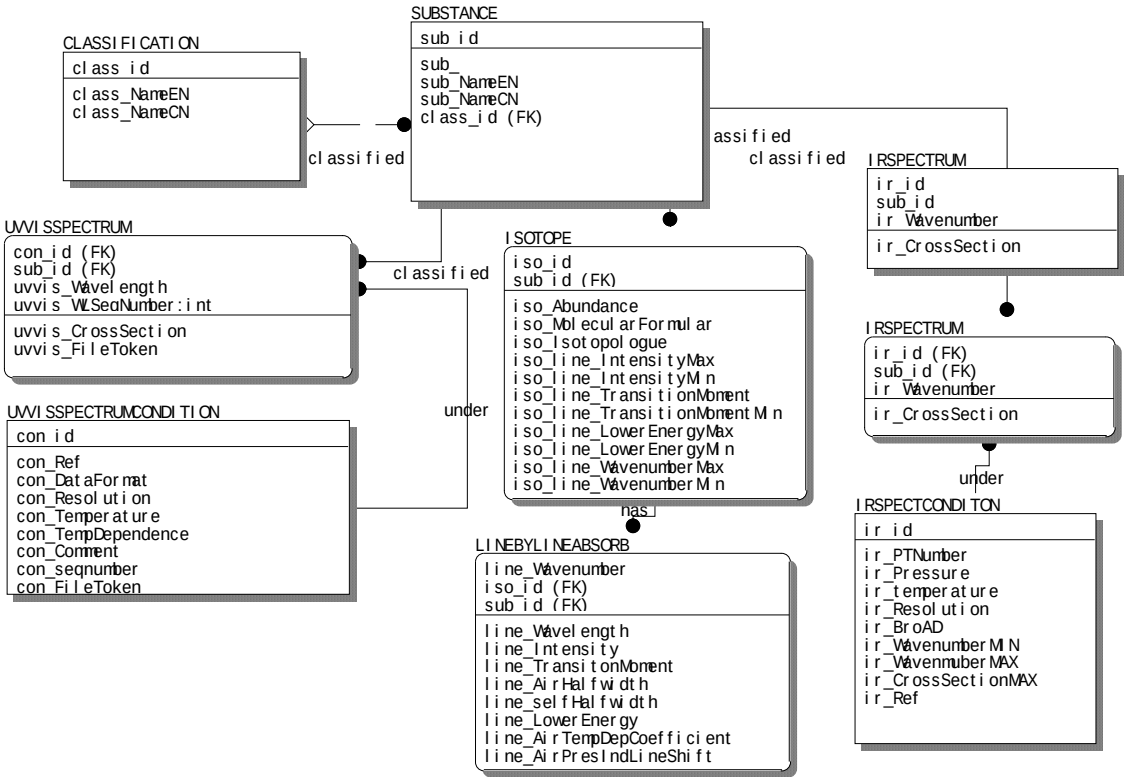

Figure 2. Entity– relationship diagram.

The web-based spectral database is divided into three sub-databases: an UVVIS (ultraviolet-visible) spectral database, an IR (infrared) spectrum database, and a line-by-line spectral database. There are more than 100 attributes of the absorption line; only the most commonly used are selected as attributes of the relevant entity. Because a large number of the collected spectral data are stored in text format, software, shown in Figure 3, has been developed to transfer the text data into the database recorders.

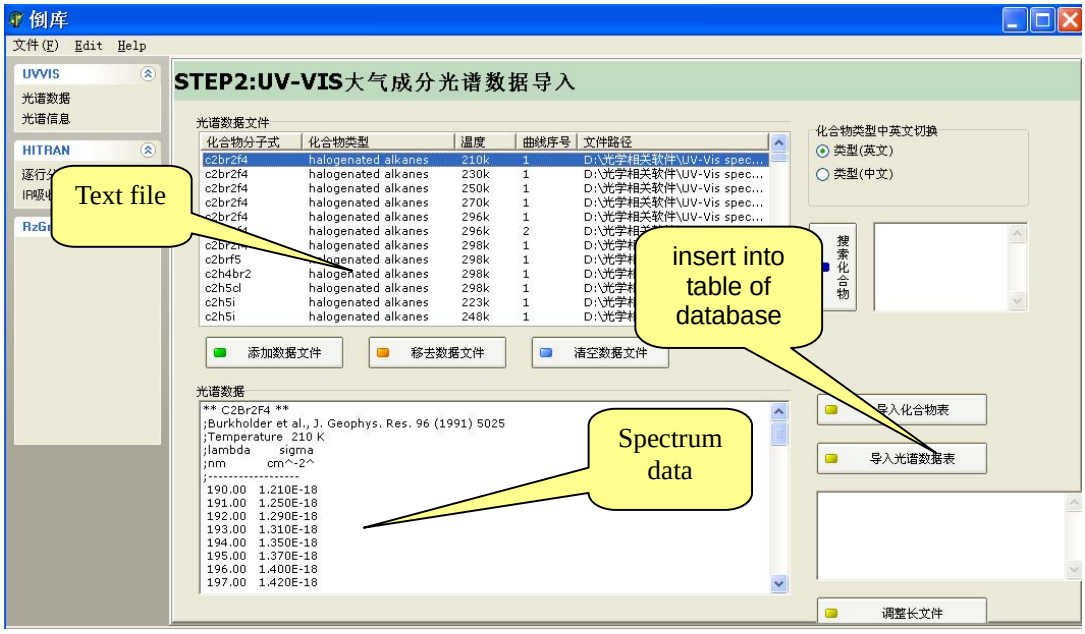

Figure 3. Transfer text data into recorders of table of database.

3 EXAMPLE AND APPLICATION

The web-based spectral database can be visited by using a common browser (MS IE 5.0). It is presently only in Chinese, but the English version will be completed in the future. This database is convenient for researchers to use and offers them an online query and visualization of spectral data, an online view of data content, and remote downloading of information. An example of the process of using the web-based spectrum database is described step by step in Figures 4 and 5.

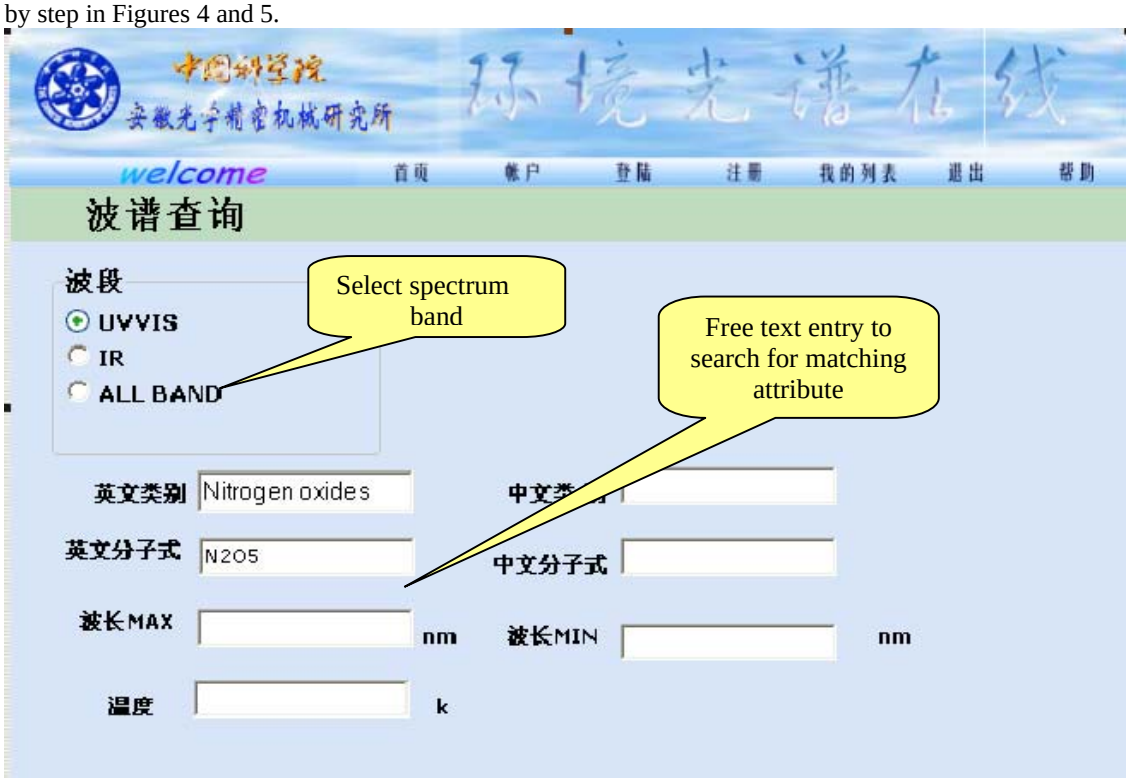

Figure 4. Query by input attribute

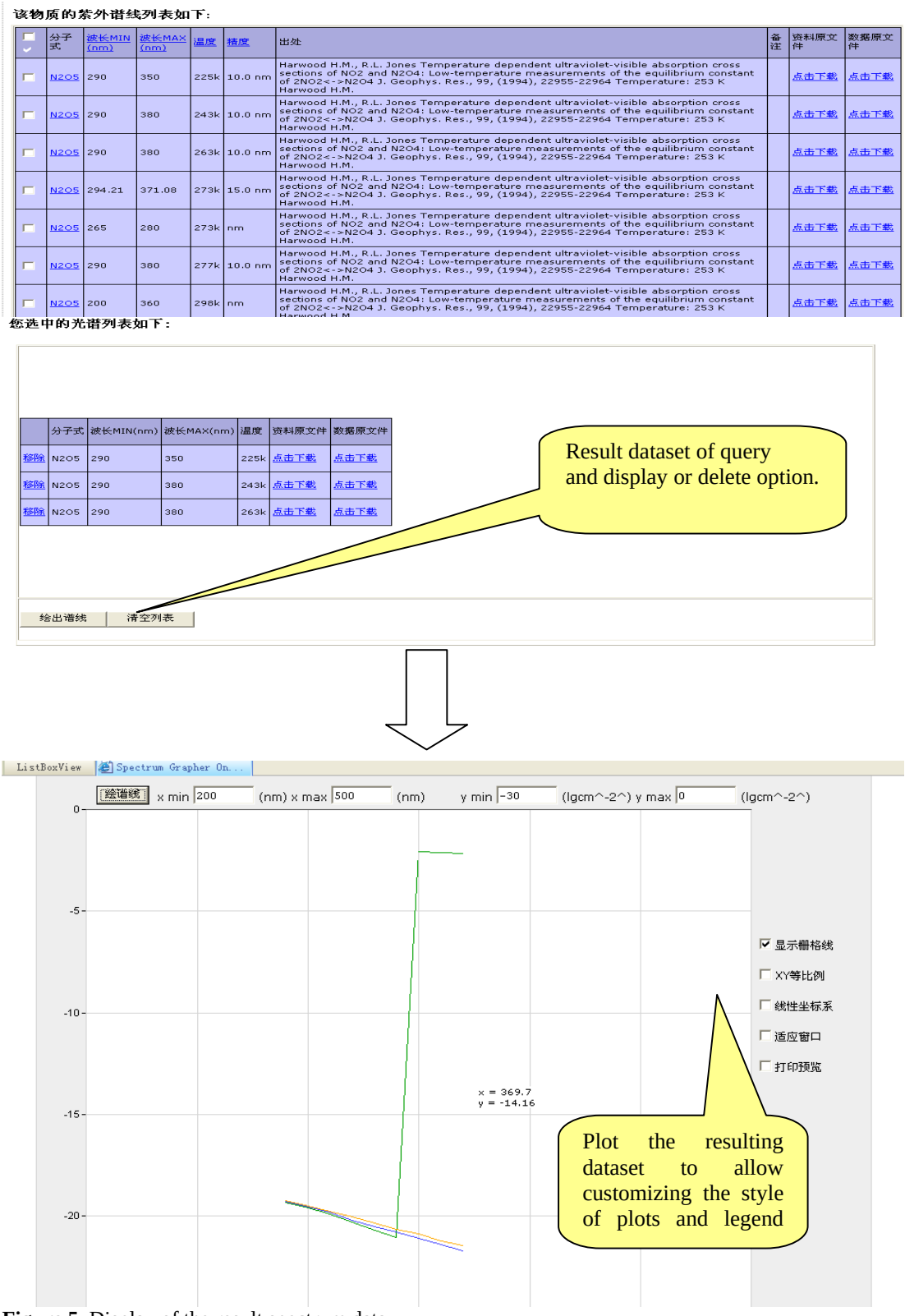

Figure 5. Display of the result spectrum data

4 CONCLUSION

This paper presents in detail the design process of a web-based spectral database for environmental applications. It applies effective queries of main parameters of characteristic spectral information of molecular absorption and data sources as well as previews of plots and compact downloads of the resulting spectral data. It provides a useful solution to existing low-efficiency file-based spectral data collection and information sharing.

5 ACKNOWLEDGEMENTS

We have greatly profited from hints, generously lavished in the course of correspondence, from Prof. Liu Wenqing and Prof. Li Xiaofeng. We also thank three anonymous reviewers for their valuable comments that helped improve the manuscript.

6 REFERENCES

Grant, W.B., Kagann, R.H., & McClenny, W.A. (1992) Optical remote measurement of toxic gases. *Air Waste Manage Assoc. 42(1)*, 18.

Leach, R. (2000) *Introduction to Software Engineering*, 428. CRC Press: Boca Raton, FL

Nölle, A. & Pätzold, F. (2004) UV / Vis Spectra of Atmospheric Constituents. Retrieved June 7, 2006 from the WWW: http://www.iup.uni-bremen.de/gruppen/molspec/DLR_spectra/index.html

Rothman, L.S. & Barbe, A. (2003) The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001. *Quantitative Spectroscopy & Radiative Transfer 82*, 5 – 44.