DEVELOPMENT OF THE WEB-BASED NIST X-RAY PHOTOELECTRON SPECTROSCOPY (XPS) DATABASE

Angela Y. Lee[1,*], Dorothy M. Blakeslee[1], Cedric J. Powell[2], and John R. Rumble, Jr.[1]

[*1]Standard Reference Data Program, National Institute of Standards and Technology, Gaithersburg, MD 20899-2310
Email: leea@nist.gov
[2]Surface and Microanalysis Science Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8370

ABSTRACT

The first Web-based version of the NIST X-ray Photoelectron Spectroscopy Database (XPSDB) is described. The current database, built from a relational database management system (RDBMS), contains critically evaluated data with over 19,000 line positions, chemical shifts, doublet splittings, and energy separations of photoelectron and Auger-electron lines. It is available free of charge to the public through the Internet at http://srdata.nist.gov/xps/.

Keywords: Journal, CODATA, Surface Analysis, X-ray Photoelectron Spectroscopy, Auger-Electron Kinetic Energies, Auger Parameters, Chemical Shifts, Core-Electron Binding Energies, Wagner Plots, Database.

1. INTRODUCTION

X-ray photoelectron spectroscopy (XPS) is one of the most commonly used techniques available for surface analysis (Brundle, 1992). This technique is utilized for the analysis of coatings, sensors, catalysts, biomaterials, electronic materials, ceramics, polymers, thin films, and composite materials for many purposes, ranging from failure analysis to the development of improved materials and processes (Briggs & Seah, 1990; Riviere & Myhra, 1998). In these applications, the composition of the outermost several atomic layers of a specimen material is determined from detection of electrons at characteristic energies that are emitted from a specimen material after irradiation with near-monochromatic X-rays. Measurements are typically made of core-electron binding energies (from emitted photoelectrons) and of Auger-electron kinetic energies with an electron energy analyzer. Elements (except for H and He) and their chemical states can be identified from the measured electron energies. Over the past 30 years, thousands of articles with XPS data have been published. In order to meet the demand for high quality data by XPS analysts, NIST established a Surface Analysis Data Center and released two PC versions of the XPS database in 1989 (Wagner, 1989; Powell, 1991) and 1997 (Wagner, Gale & Raymond, 1997; Powell, Rumble, Blakeslee, Dal-Favero, Jablonski & Tougaard, 1998), respectively. As part of our ongoing effort in critical data evaluation of the published literature, we have recently launched the third version of the XPS database. It gives free and convenient access to the energies of many photoelectron and Auger-electron spectral lines along with information for the newly added data about the specimen material, measurement conditions, and data-analysis procedure for each measurement. The latest version of the XPS database is a Web-based system that takes advantage of the recent revolution in Internet technology.

The new application system was developed on the typical three-tier Web application architecture (Figure 1).
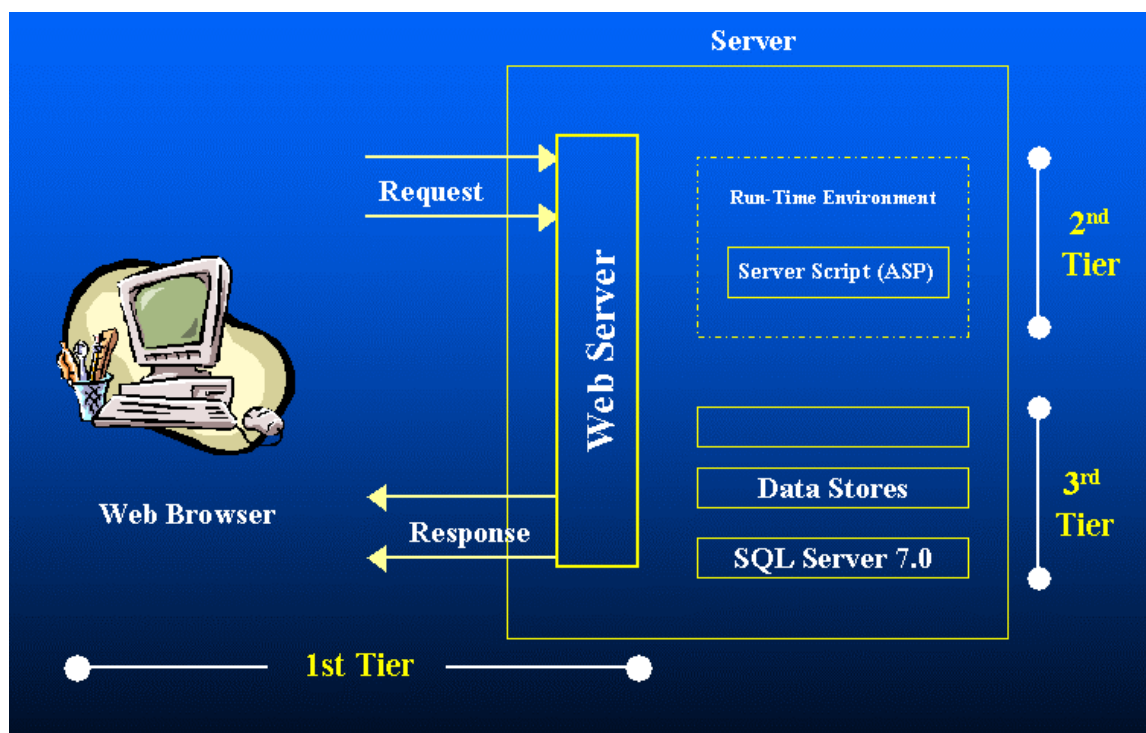
Figure 1. Simplified three-tier architecture for the XPS database system

The first tier is the presentation layer including not only the Web browser, but also the Web server, which is responsible for assembling the data into the desired format. The second tier is the application layer, which usually consists of scripts or programs. The data layer is the third tier, which provides the second tier with data. For example, the Web application collects data from the user (first tier), sends a request to the Web server, runs the requested server program (second and third tiers), packages the data to be presented to the Web browser, and sends them to the browser for display (first tier). For the server side of this system, we chose Windows NT 4.0 as the operating system and Internet Information Server (IIS) 4.0 as the Web server (see References and Notes section). Scientific data are stored in a relational database management system (RDBMS), namely SQL Server 7.0. Active Server Pages (ASP), a scripting environment located on the server, is used to process the requests from the end-user, access the database, and then create a dynamic Web page that is tailored for the particular user's request. The sophisticated Web interface, which is built from HTML pages, provides options for users to identify unknown spectral lines, retrieve data for selected elements or compounds, display Wagner plots (in which a measured Auger-electron energy is plotted against a measured core-electron binding energy for the same compound) and retrieve data for a scientific citation. This paper describes the new features of our data system as well as its design and implementation.

2. DISCUSSION

The Web version of the NIST X-ray Photoelectron Spectroscopy Database (XPSDB) (also called the NIST XPS database), version 3.0, contains new data and new software (Figure 2.). Development of the Web-based XPSDB involves five major components:

- Data compilation and evaluation
- Database design
- Graphical user interface (GUI) design
- Search strategies
- Implementation.

NIST X-ray Photoelectron Spectroscopy Database

Contents

NIST Standard Reference Database 20, Version 3.0 (Web Version)

Data compiled and evaluated
by
Charles D. Wagner, Alexander V. Naumkin, Anna Kraut-Vass,
Juanita W. Allison, Cedric J. Powell, and John R. Rumble Jr.

Web software written
by
Angela Y. Lee and Dorothy M. Blakeslee

Introduction

Database Search Menu

Data Field Definitions

Version History

Disclaimer

Acknowledgments

Contact Information

Distributed by Standard Reference Data Program of the National Institute of Standards and Technology.
NIST reserves the right to charge for access to this database in the future.

Figure 2. Graphical user interface for the front page of the XPS database

2.1 Data compilation and evaluation

The current Web-based XPSDB contains critically evaluated data originating from scientific papers published between 1968 and 1992. The data compilation and evaluation process involves the following four steps: first, a literature search of published papers identified by keywords such as "X-ray photoelectron spectroscopy" and "XPS"; second, review of the identified papers and selection of those that have adequate information on energy-scale calibration, as well as adequate information on the physical and chemical states of the specimen material; third, checks for data consistency, classification of the chemical compounds, and entry of the data into the database; and finally, data verification from checks of all numerical data and chemical information.

Compared to the previous release of the NIST XPS database, the PC version 2.0 with 16, 000 data records, the current version of the database has a total of 19,000 line positions, chemical shifts, doublet splittings, and energy separations of photoelectron and Auger-electron lines. Newly added data contain additional information about the specimen materials, the XPS measurement conditions, and the methods used to analyze the XPS measurements, as well as comments about the specimen morphology and the specimen treatments, if available. In addition, data are also included on surface and interface core-level shifts.

2.2 Database design

The entity-relationship diagram for the logical data model is given in Figure 3. The database schema of the Web version has been completely redesigned from the earlier PC versions and normalized to the third normal form (Codd, 1970), except for the chemical information table (CHEM_INFO).
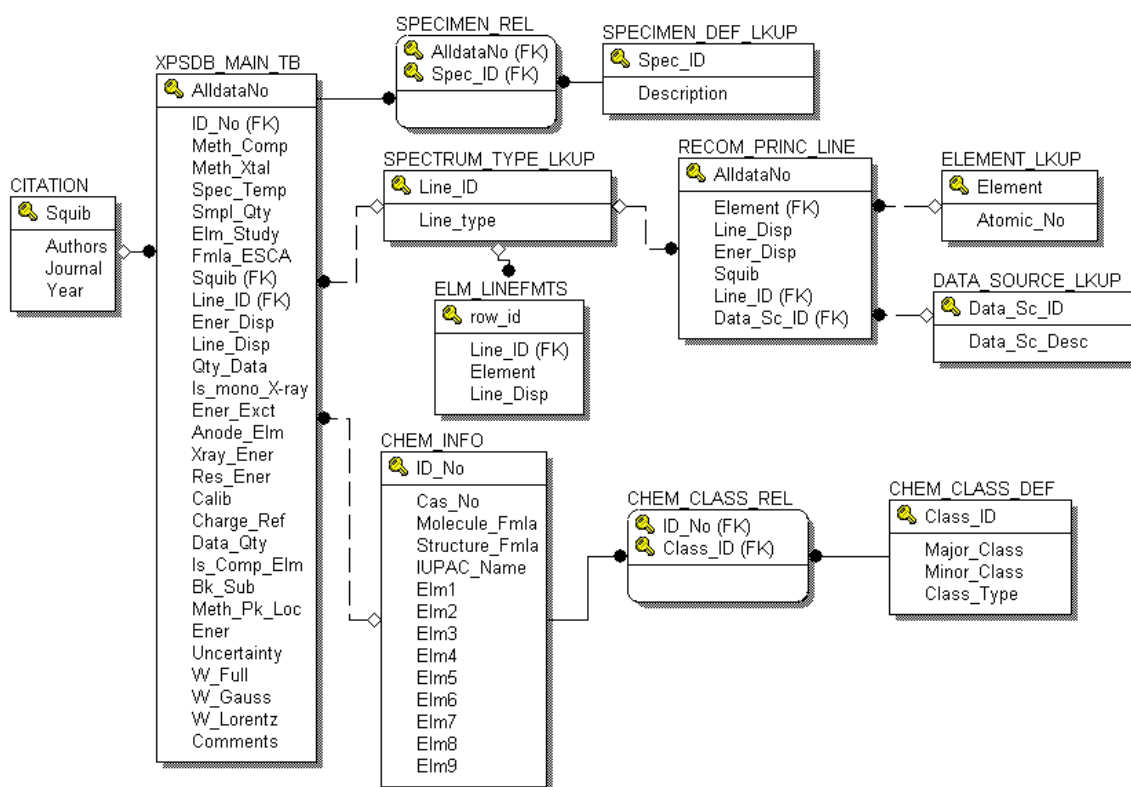
Figure 3. Logical data model for the Web version of the XPS database

The database contains the following entities:

- *Chemical information (CHEM_INFO) entity*: identity No (ID_No), CAS Registry Number (Cas_No), molecular formula (Molecule_Fmla), structural formula (Structure_Fmla), IUPAC name (IUPAC_Name) and element symbol 1-9 (Elm1-9) of the molecule, where element is sorted in an alphabetical order and assigned as Elm 1 to 9.

- *Chemical class definition (CHEM_CLASS_DEF) entity*: class identity (Class_ID), major class (Major_Class), minor class (Minor_Class) and class type (Class_Type).

- *Chemical class relationship (CHEM_CLASS_REL) entity*: identity No (ID_No) and class identity (Class_ID).

- *Citation entity*: citation identity (Squib), a list of authors (Authors), journal + volume number + page number (Journal) and publication year (Year).

- *Specimen definition look-up (SPECIMEN_DEF_LKUP) entity*: specimen identity (Spec_ID) and definition of the specimen code (Description).

- *Specimen relationship (SPECIMEN_REL) entity*: record identity (AlldataNo) and specimen identity (Spec_ID).

- *Spectrum type look-up (SPECTRUM_TYPE_LKUP) entity*: spectral line identity (Line_ID) and the type of the spectral line (Line_type).

- *Data source look-up (DATA_SOURCE_LKUP) entity*: data source identity (Data_Sc_ID) and data source description (Data_Sc_Desc) for the recommended principal lines.

- *Element look-up (ELEMENT_LKUP) entity*: element symbol (Element) and atomic number (Atomic_No).

- *Recommended principal line of element (RECOM_PRINC_LINE) entity*: record identity (AlldataNo), element (Element), spectral line designation (Line_Disp), energy (Ener_Disp), citation identity (Squib), spectral line identity (Line_ID) and data source identity (Data_Sc_ID).

- *Element spectral line formation (ELM_LINEFMTS) entity*: row identity (Row_id), spectral line identity (Line_ID), element symbol (Element) and spectral line designation (Line_Disp).

- *XPS database main table (XPSDB_MAIN_TB) entity*: a record identity (AlldataNo), identity No (ID_No), method of determining specimen composition (Meth_Comp), method of determining specimen crystallinity (Meth_Xtal), specimen temperature (Spec_Temp), sample quality (Smpl_Qty), Element of study (Elm_Study), XPS formula (Fmla_XPS), citation identity (Squib), spectral line identity (Line_ID), energy (Ener_Disp), spectral line designation (Line_Disp), data quality (Qty_Data), use of X-ray monochromator (Is_mono_X-ray), excitation energy source (Ener_Exct), element of anode (Anode_Elm), energy of X-ray (Xray_Ener), overall energy resolution (Res_Ener), calibration (Calib), charge reference (Charge_Ref), quality of data processing (Data_Qty), element or compound (Is_Comp_Elm), background subtraction method (Bk_Sub), peak location method (Meth_Pk_Loc), energy uncertainty (Uncertainty), full width at half-maximum intensity (W_Full), Gaussian width (W_Gauss), Lorentzian width (W_Lorentz) and comments about the specimen morphology and the specimen treatments (comments).

2.3 Graphical user interface (GUI) design, search strategies and implementation

The rest of the paper will focus on the GUI design, search strategies and their implementations, which are closely interrelated. Based on the needs of the XPS analyst, mainly the identification of elements from observed spectral features and the determination of chemical state from small energy shifts of photoelectron and Auger-electron peaks, we have implemented the XPSDB with the following search options:
- Identify unknown spectral lines
    - Binding energy
    - Auger kinetic energy
    - Auger parameter
    - Doublet separation
- Retrieve data for selected elements
    - Binding energy (for photoelectron lines)
    - Auger kinetic energy (for Auger lines)
    - Surface/Interface core-level shift (for photoelectron lines)
    - Chemical shifts (for photoelectron lines, Auger lines, and Auger parameters)
        - Retrieve measured data
        - Calculate data using elemental data as reference
        - Both of the above
    - Elemental reference data
        - Recommended binding energies for principal photoelectron lines (85 elements)
        - Recommended kinetic energies for principal Auger lines (18 elements)
        - All data for one element
- Retrieve data for selected compounds
    - Selected groups of elements
    - Selected element in a compound

- o Chemical name
- o Chemical classes
  - ▪ Inorganic
  - ▪ Organic
  - ▪ Ligand
  - ▪ Other
- Display Wagner plot
- Retrieve data by scientific citation

For the search option, identification of unknown spectral lines, a user selects one of the four energy types (binding energy, Auger kinetic energy, Auger parameter or doublet separation) from the select menu and then enters a desired energy value along with its uncertainty for the search (Figure 4).

Figure 4. Graphical user interface for identification of unknown spectral lines

Up to four unknown spectral lines may be requested for each search, and the data are validated by a Javascript program before the user request form is submitted to the server. If matches are found for the specific search, the element, spectral line, chemical formula and energy value as well as a hyperlink for additional detailed information are provided for each match (Figure 5).

⊘ Matches (Total: 7) from unknown spectral lines search

energy type   =   binding energy

line energy (eV)   =   8

energy tolerance (eV) =   ± 1.00

Element	Spectral Line	Formula	Energy	Details ?
Hg	5d5/2	Hg0.87Fe0.13Se	8.15	Click
Hg	5d5/2	Hg0.86Mn0.14Se	8.14	Click
Hg	5d5/2	HgSe	8.17	Click
Hg	5d5/2	HgTe	7.805	Click
Hg	5d5/2	Hg/Cu	7.81	Click
Hg	5d5/2	Hg/Cu	7.86	Click
Hg	5d5/2	Hg/Cu	8.19	Click

⊘ Sorting by:

◉ Element   ◉ Spectral Line   ◉ Formula   ◉ Energy

[Start Sorting]   [Help Menu]   [Go Back]

Figure 5. An example of a summary table showing matches from a search

A sort function is provided to arrange the matching results in a selected order. The additional detailed information is classified into six categories: i.) general (element, formula, XPS formula, IUPAC name, Chemical Abstracts Service (CAS) Registry Number, and compound classification); ii.) measurement information (use of X-ray monochromator, excitation energy, overall energy resolution, calibration, charge reference and energy scale evaluation); iii.) specimen information (specimen, method of determining specimen composition, method of determining specimen crystallinity, specimen temperature and sample quality); iv.) citation (author name(s) and journal citation information); v.) data processing (data type, line designation, quality of data, energy, energy uncertainty, background subtraction method, peak location method, full width at half-maximum intensity, Gaussian width and Lorentzian width); and vi.) comments.

Five options are provided for retrieving data for a selected element: (1) binding energy, (2) Auger kinetic energy, (3) chemical shift, (4) surface/interface core-level shift, and (5) elemental reference data. Since the first four options share the same user interface, the GUI for binding energy is used for the illustration (Figure 6a). A simple HTML frame is the primary vehicle for the GUI. Users can enter the desired elemental symbol in a "Select element" box and press the "Submit" button on the left frame. Available spectral lines for the desired element are listed. In order to arrange the available spectral lines in spectroscopic order (Figure 6b), the row identity in the ELM_LINEFMTS entity (see Figure 3) is introduced. Up to four selected elements and their spectral lines can be shown in the display box on the right frame and submitted for a further search. The search results are organized into a table similar to Figure 5, which provides the same kinds of information. On the other hand, retrieving data in the elemental reference data option has three suboptions: i. recommended binding energies for the principal

photoelectron line of 85 elements; ii. recommended kinetic energies for the principal Auger line of 18 elements; and iii. all data for one element. The four underlying relation entities (SPECTRUM_TYPE_LKUP, RECOM_PRINC_LINE, ELM_LKUP, and DATA_SC_LKUP) in Figure 3 are applied here. The principal photoelectron and Auger lines searches both result in a table with the same format containing columns for element, atomic number, spectral line, energy, reference and a detail hyperlink that gives further information about data type, energy value, line designation, data source and literature reference for the selected element. For the third suboption of retrieving all data for one element, a periodic table is chosen to be the GUI. An underlined elemental symbol in the periodic table gives access to results in the same style of data display table as shown in Figure 5.

Figure 6a. The common graphical user interface for retrieving data for a selected element

Figure 6b. Available spectral lines for the element Cu

In addition to retrieving data for selected elements, one can retrieve data for selected compounds. The first option is to search for data by selected groups of elements in a chemical compound (Figure 7).

Figure 7. Graphical user interface for retrieving data by element composition of a compound

The user first selects one or more elements (up to 9) from the drop-down lists and then chooses one of the two search options (either specified elements only or specified elements plus any other). The "Start Search" button sends the user's request to the Web server. In order to sort the selected elements in a desired order, a series of stored procedures were written and stored in the SQL database to process the user's data before query of the database table called CHEM_INFO. The search results are presented in the same style of table as shown in Figure 5. The second choice is to retrieve all data for a selected element in a compound. To serve this purpose, the periodic table is a perfect choice for the GUI. Again, an underlined elemental symbol in the periodic table gives search results in the same style of display table as shown in Figure 5. The third option is to search data by text strings in a chemical name (Figure 8a). If any matches are found, IUPAC names of the compounds are sorted alphabetically and tabulated in a table with thirteen rows per page (Figure 8b). If the search results have more than thirteen records, the user

can go to the successive page(s) by clicking the link in the list on the page counter below the table. The last choice is to retrieve data by chemical classes: inorganic, organic, ligand, and other for which there are subcategories for each class. For example, the inorganic class allows the user to access the inorganic compounds by either anion or cation. If the user clicks on the acetate anion, for example, all the compounds belonging to this family are displayed. A hyperlink gives access to all of the spectral data for a selected compound.

Figure 8a. Graphical user interface for retrieving data by compound name

Figure 8b. The matching results from compound name search

A periodic table is also used as the GUI for the option to retrieve an element for displaying a Wagner Plot (Wagner, Gale & Raymond, 1979). After an element is selected, a list of compounds containing that selected element is displayed. The user can select up to twelve compounds for a Wagner plot (Figure 9), which is graphed by a Java Applet. For each compound in the plot, the user can also select a compound for data display.
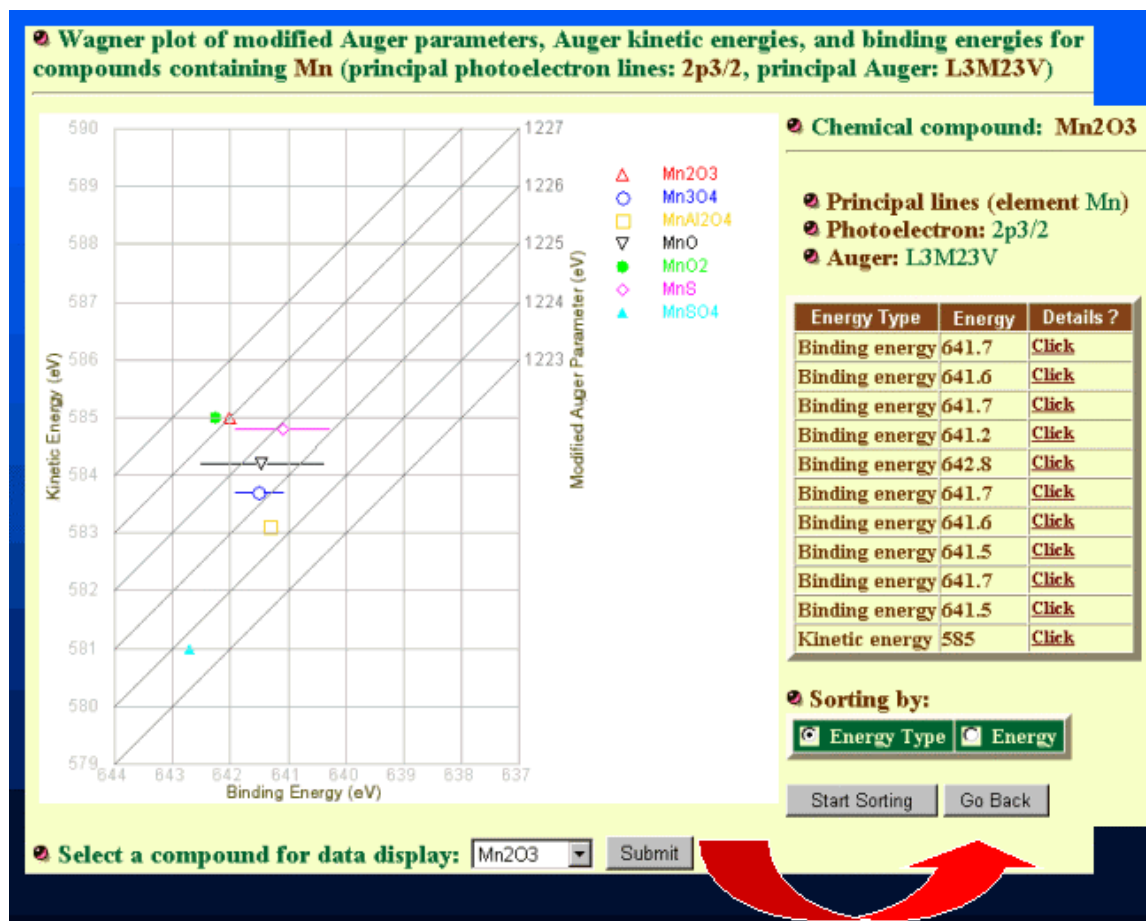
Figure 9. Wagner plot

Finally, the last option allows the user to retrieve data by scientific citation with search options of author, publication year and scientific journal. The search submission yields a summary table for the publication year, author name(s) and reference link for each match. Clicking on the reference link gives the same format of display as illustrated in Figure 5.

3. CONCLUSIONS

The recently launched Web version of the NIST X-ray Photoelectron Spectroscopy Database has already had a positive impact on surface analysis by XPS. The experience we have gained will lead us to continue the development of this system. Future versions of the database will be cross-referenced with other NIST Web data systems and will include additional critically evaluated data with improvement in the user interface and enhancement in the functionality.

4. ACKNOWLEDGEMENTS

5. REFERENCES AND NOTES

Briggs D. & Seah M.P. (Eds.) (1990) Auger and X-ray Photoelectron Spectroscopy in *Practical Surface Analysis*, *Vol. 1*. (2nd Edition) Chichester: Wiley.

Brundle, C.R. (1992) X-Ray Photoelectron Spectroscopy. In Brundle, C.R., Evans, Jr. C.A. & Wilson S. (Eds.), *Encyclopedia of Materials Characterization: Surfaces Interfaces, Thin Films* (pp 282-299). Boston: Butterworth-Heinemann.

Codd, E.F. (1970) A relational model of data for large shared data banks. *Communications of the ACM* 13, 377-87.

Riviere J.C. & Myhra S. (Eds.) (1998) *Handbook of Surface and Interface Analysis: Methods for Problem-Solving*. New York: Marcel Dekker.

Powell, C.J. (1991) Formal Databases for Surface Analysis - The Current Situation and Future Trends, *Surf. Interface Anal.* 17 (6), 308-314.

Powell, C.J., Rumble J.R., Blakeslee, D.M., Dal-Favero, M.E., Jablonski, A. & Tougaard, S. (1998) The NIST Surface Analysis Data Center. In D. G. Seiler, A. C. Diebold, W. M. Bullis, T. J. Shaffner, R. McDonald, and E. J. Walters (Eds.), *Characterization and Metrology for ULSI Technology. American Institute of Physics Conference Proceedings No. 449* (pp 887-891). Woodbury, New York: American Institute of Physics.

Wagner, C.D. (1989) *NIST X-Ray Photoelectron Spectroscopy Database, Version 1.0.* Retrieved from NIST Standard Reference Data Program website: http://www.nist.gov/srd/.

Wagner, C.D., Gale, L.H. & Raymond, R.H. (1979) 2-Dimensional Chemical-State Plots - Standardized Data Set for Use in Identifying Chemical-States by X-Ray Photoelectron Spectroscopy, *Analytical Chemistry, 51*(4), 466-482.

Wagner, C.D., Powell, C.J., Allison, J.W. & Rumble J.R., Jr. (1997) *NIST X-ray Photoelectron Spectroscopy Database, Version 2.0.* Retrieved from NIST Standard Reference Data Program website: http://www.nist.gov/srd/.