

CasaXPS



User Guide

CasaXPS User's Manual

for Version 2.0, October 2001

Product design: Casa Software Ltd.

Documentation: Acolyte Science.

Manual version 1.2

CasaXPS and the User's Manual are copyright © 2001 Casa Software Ltd.

All rights reserved.

Information in this document is subject to change without notice and does not represent a commitment on the part of the copyright holder. The software described in this document is furnished under a license agreement. No part of this publication may be copied, reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photo reproduction, recording or otherwise without the prior written consent of Casa Software Ltd.

Acknowledgements

CasaXPS wishes to thank all those connected with the production of this manual and the software system it describes. In particular, thanks are due to those who provided experimental data, results, and interpretations which form the bulk of the examples here. Notable among these contributors are: Morgan Alexander, Graham Beamson, David Briggs, Peter Cumpson, Kevin Harrison, Len Hazel, Fran Jones, Simon Page, Roy Paynter and Martin Seah

All third party Trade Marks are acknowledged and used without prejudice.

Any references to company names in sample output and templates are for demonstration purposes only and are not intended to imply any endorsement, or refer to any particular existing organisation or equipment.

CasaXPS and the User's Manual are copyright © 2001 Casa Software Ltd.

Contents

Introduction	5	Experimental Variable	22
Getting Started	6	Transmission Correction and Quantification	22
Installing CasaXPS	6	Transmission Functions	23
Starting CasaXPS	6	Quantification	24
Terminology	6	Data Display and Browser	25
A Quick Tour of CasaXPS	7	Understanding the Data	25
Step 1: Load the experiment file	8	Selecting the ISO 14976 Blocks	26
Step 2: Select a spectrum for processing	8	Zooming the Data	26
Step 3: Identify peaks.	9	Tiles of Spectra	26
Step 4: Annotate peaks using element markers	11	Tile Display	27
Step 5: Create quantification regions	11	Colours	28
Step 6: Adjust region start and end points	12	Fonts	29
Step 7: Print the results	14	Display Parameters and Scrolled Tiles	29
Data Display and Browser Windows	15	Graph Annotation	31
Element Library	16	Quantification Tables	31
Graph Annotation	17	Moving Annotation and the History Mechanism	32
Processing	17	Peak Labels	32
Quantification	17	Text Annotation	33
ISO 14976 (“VAMAS”) Files	17	Processing	34
ISO 14976 File Format	19	Smoothing	34
Partially Encoded Format Versions	19	Differentiation	35
File Structure	20	Integration	35
Experiment header	21	Energy Calibration	35
Data Blocks	21	Intensity Calibration	36
Binding vs Kinetic Energy	22		

Contents

Processing History	36	Asymmetric Line-Shapes	61
Element Library	38	Modifications to the Doniach Sunjic function .	62
Library File Structure	38	A New Line-Shape	64
Loading an Element Library	40	Background Subtraction	65
Identifying Peaks	41	Adjusting the universal cross-section	68
Quantification	42	Simulating Spectra using Tougaard methods ..	69
Regions	42	Using the Calculator and Comparing Spectra ...	71
Synthetic Components	44	Creating a Comparison File	71
Quantification Calculation	45	The Calculator	72
Quantification Report	45	Principal Component Analysis	80
Optimisation of the Peak Parameters	47	Introduction	80
Propagating Quantification	49	Theory of Principal Component Analysis	80
Data Editor	49	Residual Standard Deviation (“Real Error”) ..	82
Derivatives and Peak Envelopes	50	Chi square	83
Quantification by Example	51	Target Factor Analysis	83
Report Files and Excel	54	Target Testing	84
Line Shapes and Backgrounds	55	Principal Component Analysis by Example ...	84
A List of Line-Shapes	57	Principal Component Analysis and Real Data .	90
Gaussian/Lorentzian Product Form	57	PCA and CasaXPS	94
Gaussian/Lorentzian Sum Form	57	Viewing the Data in Factor Space	96
Alternative Asymmetric Line-Shapes	58	Monte Carlo Methods	
Line-Shapes Based upon Backgrounds	59	- Uncertainties in Intensity Calculations	97
Line-Shapes Available in CasaXPS	59	Monte Carlo Data Sets	98
Further adjustments to the basic shapes	60	Error Matrix	98

Contents

Monte Carlo: A Simple Example	99	SSI M-Probe Files	129
Quantification	102	PHI MultiPak ASCII files	130
Monte Carlo End Point Determination	102	Command Summary	131
Integration Region Limits	102	Main menubar	131
Intensities Determined by Peak-Fitting	104	File	131
Summary	106	View	132
Analytical Applications	107	Window	132
Organic Polymers and Curve Fitting	107	Options	132
Quantification using “Tagged” Regions	111	Help	133
Introduction	111	Toolbar	133
“Tagging” regions in a survey scan	112	Options Bar	134
“Tagging” for different line shapes	113	Processing Dialog windows	136
Trend Analysis for Metal Oxidation	114	Page Layout	136
Introduction	114	Tile Display Parameters	136
PCA for Aluminium Oxidation Sequence ..	116	Quantification	138
Synthetic Model for an Oxidation Sequence ..	117	Processing	140
Peak Fit Using Doniach Sunjic Line Shapes ..	118	Annotation	141
Adjusting a Custom Quantification Report ...	120	Library	141
Using Different File Formats	123	Appendix 1: ISO 14976 format files annotated ...	143
Kratos DS800 Binary Files	123	Appendix 2: ISO 14976 and the World Wide Web	148
VG Eclipse Files	124	Appendix 3: Definitions and Formulae.....	149
VGX900 (Ron Unwin)	125	Glossary of terms	149
“Dayta” System Files (Bristol IAC system) ...	127	Formulae: Peak shapes	149
Kratos Vision 1.x/2.x ASCII files	128	Gaussian/Lorentzian Product Form	149
		Gaussian/Lorentzian Sum Form	149

Contents

Doniach Sunjic:	150	References from the Text Footnotes:	151
Appendix 4: References and other Resources	150	Appendix 5: Quick Reference Card	151
World wide web links:	150		



Computer Aided Surface Analysis for X-ray Photoelectron Spectroscopy (CasaXPS) has been designed and written for analysts and research scientists who use XPS for surface characterisation.

In contrast to the special purpose control systems that form part of commercial XPS instruments, CasaXPS offers a compact, portable, efficient and user-friendly processing system to anyone with an IBM compatible (Pentium) PC running Microsoft Windows 95 (or later) or suitable emulator. It incorporates much of (in many cases more than) the processing functionality of the instrument linked packages without recourse to unfamiliar operating systems or hardware or proprietary file formats. It is designed from the outset on the basis of the ISO 14976 Surface Chemical Analysis Standard Data Transfer Format and so by design has a universal, “cross platform” approach, independent of anything other than a reasonable adherence to the ISO standard. Spectra collected in the standard format may be selected, viewed and processed in a simple yet powerful way and the results of CasaXPS data reduction, presented in a variety of graphical and tabular formats, are available for incorporation into

and use as data by other popular Microsoft packages such as “Word” or “Excel”. Additional spectrum input filters for other commercial file formats are also available, supplementing the ISO standard and broadening its usefulness.

Casa XPS has been written entirely in native C++ using the Microsoft Developers Studio programming environment. It employs Microsoft Foundation Class (MFC) libraries to provide a standard graphical user interface as well as much of the data management required by the system.

The software suite as a whole may be regarded as being structured from five interacting components:

- Data display and browser (an “experiment window”)
- Element library (a “dialog window” as are the others)
- Graph annotation
- Data processing
- Quantification.

This manual provides a detailed description of the system, its operation, command and function set, and its use in practice in “real life” analytical applications.

Getting Started

Installing CasaXPS

You must install the application from the CasaXPS CD onto your hard disk. “Drag and Drop” installation is all that is required for Windows 98. Instructions are available, if required, in the “Read Me” file on the installation CD. Installation should occupy no more than 5MB on your hard drive. You will need perhaps an equivalent space for storage of data files. Make sure that you have your serial number to hand when you install the application, or else the package will revert to “Demonstration Mode” and will not allow you to save the results of your work. The serial number is available from your registration documents (which may be e-mailed to you or included with the CD). You can “validate” a registration at any time after installation, turning a “demo” version into a fully licensed product, by using the “About” button in the Toolbar.

If you are using an emulator package on a processor other than an Intel Pentium, ensure that you obey the emulator’s instructions for installation of applications packages, that you understand the limitations (if any - e.g. mouse button availability) of the system, and that you have sufficient space and processing power for the convenient use of CasaXPS.

Starting CasaXPS

Simply click on the desktop short-cut (if, as recommended, you installed one) or double click on the programme icon in “Windows Explorer”. Ensure that you have stored your

ISO 14976 data files in an easily accessible directory (or have the demo data on the installation CD available) Ensure also that you have an appropriate element library available (one needs to be installed before the first use of “Library” system - see “Element Library” on page 38 - a simple “generic” version is available on the installation disk, but check with your instrument operator which is the most appropriate file for the data acquired by that particular instrument). At some point you may wish to install and use the on-line “help” files (see note on your CD inlay) and you should have a “.def” directory prepared if you wish to store synthetic components and other items for future use

Terminology

CasaXPS component windows appear on screen as shown in Figure 1, where the annotation defines their names as used throughout this manual.

It is assumed throughout that the user is familiar with the “normal” IBM PC and Microsoft Windows operation and terminology.

The following section should be used as a “hands on” introduction to CasaXPS to provide an overview of key features. Subsequent sections enlarge briefly on the role of the system modules, and the remaining chapters deal in depth with specific topics and applications.

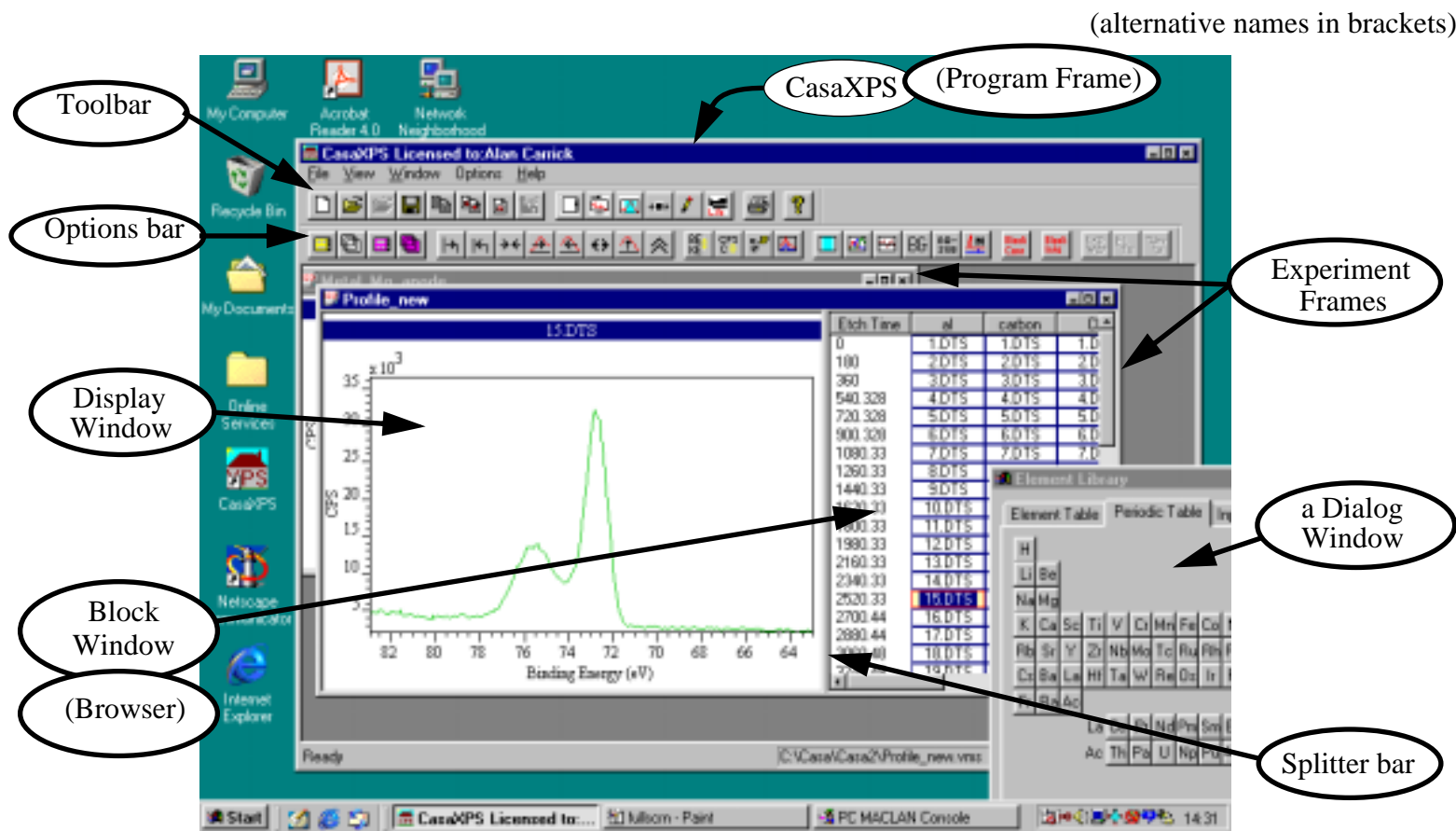


Figure 1: A Windows 98™ screen showing CasaXPS in operation

A Quick Tour of CasaXPS

This section describes one of the many possible ways to

quantify a survey spectrum in CasaXPS. The intention is to explain how the interacting components of the system aid the analyst rapidly to produce a quantification table suitable for inclusion in a customer's report.

Step 1: Load the experiment file

CasaXPS is designed for ISO-14976 formatted XPS spectra. A data file containing spectra stored in the ISO format may be selected via the Open VAMAS file dialog window shown in Figure 2. To select a VAMAS file press the toolbar button indicated in Figure 3 to invoke the Open VAMAS file dialog window

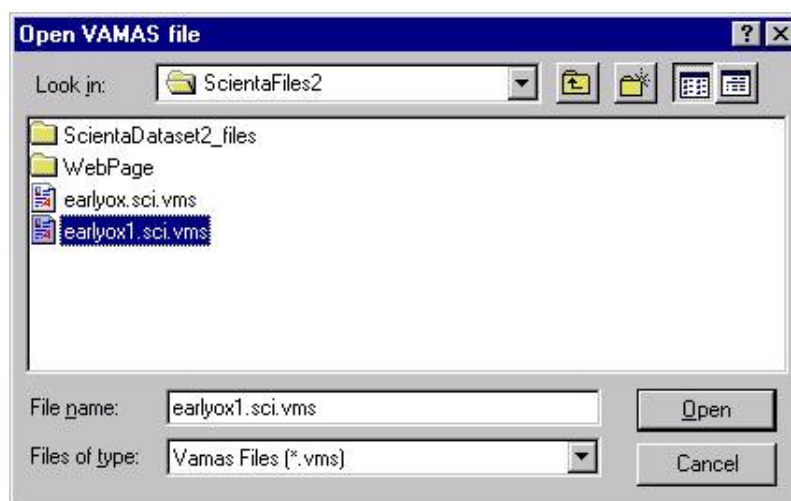


Figure 2: Open VAMAS file Dialog Window

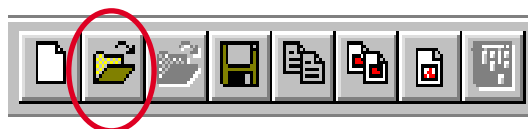


Figure 3: Toolbar. The Open file toolbar button is circled.

Step 2: Select a spectrum for processing

Once the data file has been selected via the Open VAMAS file dialog window, the spectra held within the file will be displayed in a logical array where each data region is labelled by the VAMAS block identifier string and these block labels are arranged according to the element/transition strings recorded in the VAMAS file. Figure 4 shows the state of the CasaXPS windows after a spectral data file has been loaded.

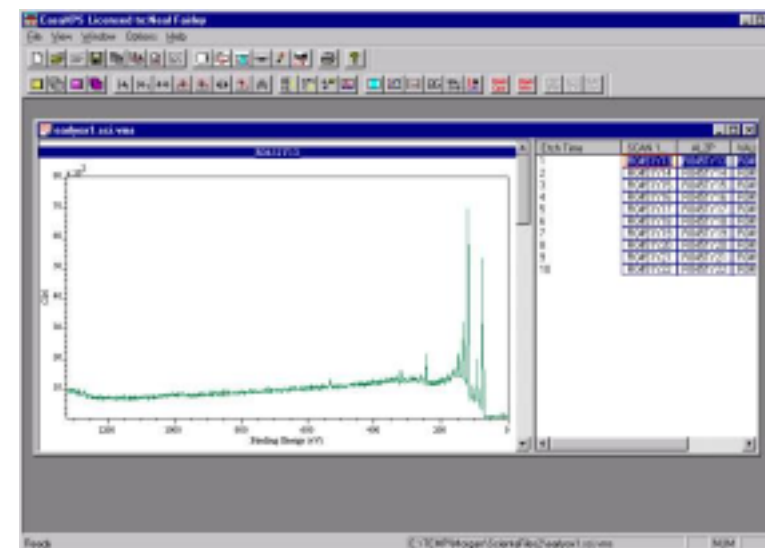


Figure 4: CasaXPS frame windows after a file has been loaded

Note the Multiple Document Interface MDI architecture. When opened, the VAMAS file appears as sub-frame window, which is managed by the CasaXPS main frame and it

in turn displays the file as a logical set of spectral blocks. These logical blocks labelled by name allow spectral regions to be selected and displayed in the left-hand-side of the sub-frame (in the Display window). The current selection is that set of labels highlighted blue (default colour scheme), and when the file is first opened the top row of spectra regions will be highlighted and displayed in the left-hand sub-frame. If more than one spectrum appears on the first row of the Browser, then the spectra will be displayed in the Display using a scrolled list of display tiles.

The left-hand mouse button and the cursor are used to make a new selection. Point the mouse at a block label within the array of labels on the right-hand-side and left-click. Any blocks currently selected will be deselected and the block under the mouse will become the current selection. Press the toolbar button indicated in Figure 5 to display the selection in the scrolled list on the left-hand-side of the sub-frame.

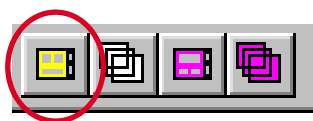


Figure 5: Toolbar button to show spectra in the Display window (left hand sub-frame).

In this example, the spectrum of interest is the last in the sequence of survey spectra recorded in the data file. Figure 6 shows the state of the CasaXPS windows after a specific survey spectrum has been selected and displayed. Also note in Figure 6 that the user has maximized the Experiment Frame within the main frame of CasaXPS, hiding others below.

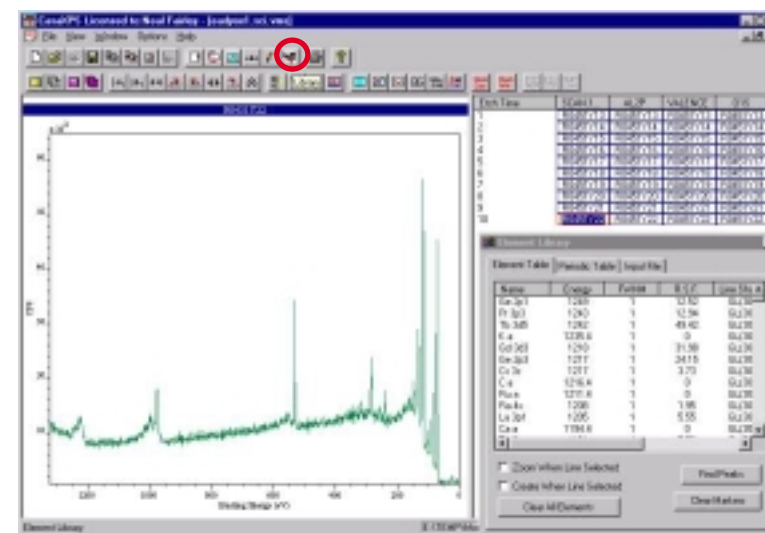


Figure 6: CasaXPS ready to identify spectral features.

Step 3: Identify peaks.

CasaXPS will load an element library called CasaXPS.lib located in the same directory as CasaXPS.exe. The element library is used to identify the peaks seen in the data and is central to rapidly creating quantification regions for a survey spectrum.

The Element Library dialog window is invoked or brought to the front of any CasaXPS dialog windows by pressing the toolbar button indicated in Figure 6. The first property page of the element library dialog window is a scrolled list of line positions where the list is ordered by the energy of the entries. If the mouse is left-clicked on the spectrum displayed

in the left-hand-side of the sub-frame shown in Figure 6, then the scrolled list of element library entries will scroll to the energy indicated by the mouse click. Therefore, by pointing at the peaks in the spectrum via the mouse, the names for the various features can be brought into view and if the name field displayed in the scrolled list is selected,

again using the mouse, a set of markers will appear on the spectrum positioned at the energies at which spectral lines should appear for the chosen element. Figure 7 shows the spectrum after the O 1s and C 1s lines have been activated via the element library scrolled list.

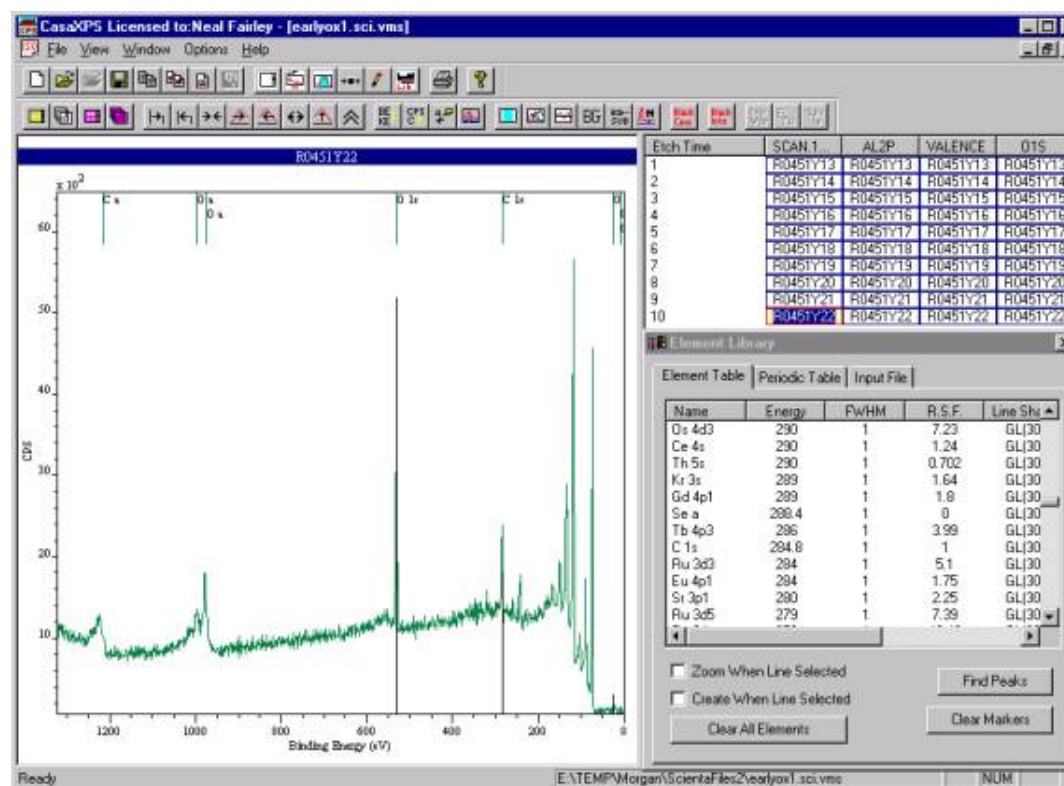


Figure 7: Peak identification via the element library.

If the name field of an element library entry is selected a second time, the element markers are removed from the display. The current set of active element markers can be view on the Periodic Table property page also located on the element library dialog window. The element markers may be activated and deactivated via the Periodic Table as well as the scrolled element list.

Step 4: Annotate peaks using element markers

Once all the spectral lines are identified and markers have been placed on the displayed spectrum, the peaks can be annotated using the names stored in the element library. This step not only creates a labelled spectrum, but these annotation peak labels are used to create a set of quantification regions at the press of a button.

The annotation dialog window includes a property page entitled Peak Labels (Figure 8). All the peak markers active on a spectrum are listed on the Peak Labels property page and may be selected using the mouse and the control key (to add to the selection). When the Apply button is pressed those labels selected in the list of names on the Peak Labels property page become annotation labels on the spectrum. The peak labels are positioned on the spectrum according to the location of the nearest peak, so it is important that the spectrum is calibrated before this procedure is performed.

Step 5: Create quantification regions

The Quantification Parameters dialog window is available from the top toolbar button indicated in Figure 9. Integration regions are created and adjusted via the Quantification Parameters dialog window, where a scrolled list of region en-

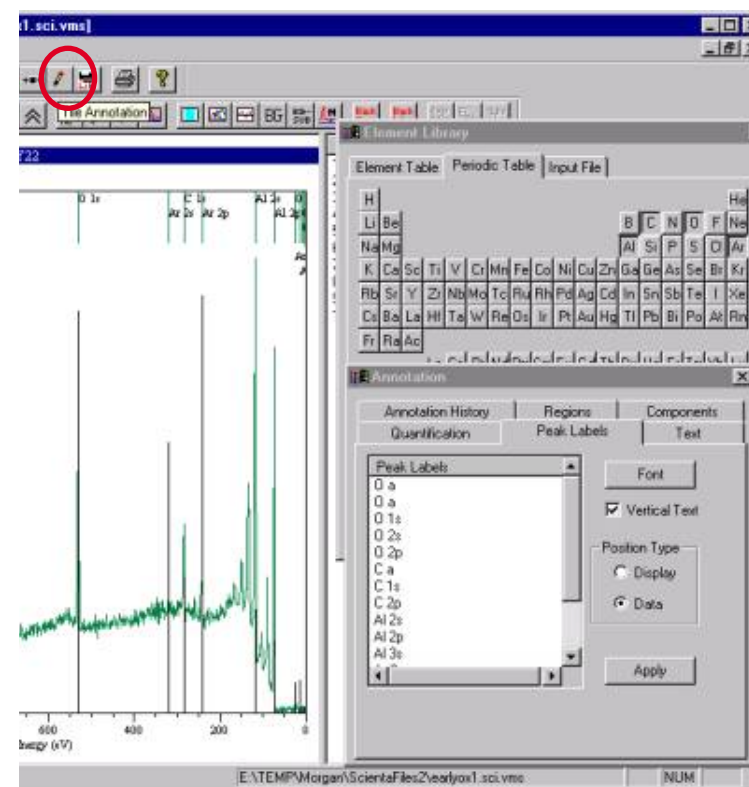


Figure 8: Peak Labels Property Page on the Annotation Dialog Window

tries is offered. Each column in the scrolled list represents the set of parameters that define a quantification region.

The annotation for the peaks is used by the “Create from Labels” button indicated at the bottom of the Quantification Parameters dialog window in Figure 9. Regions created in this way have the name and RSF set from the entries in the element library. The background type for each of the newly

created regions is determined from the last background type entered on the Regions property page.

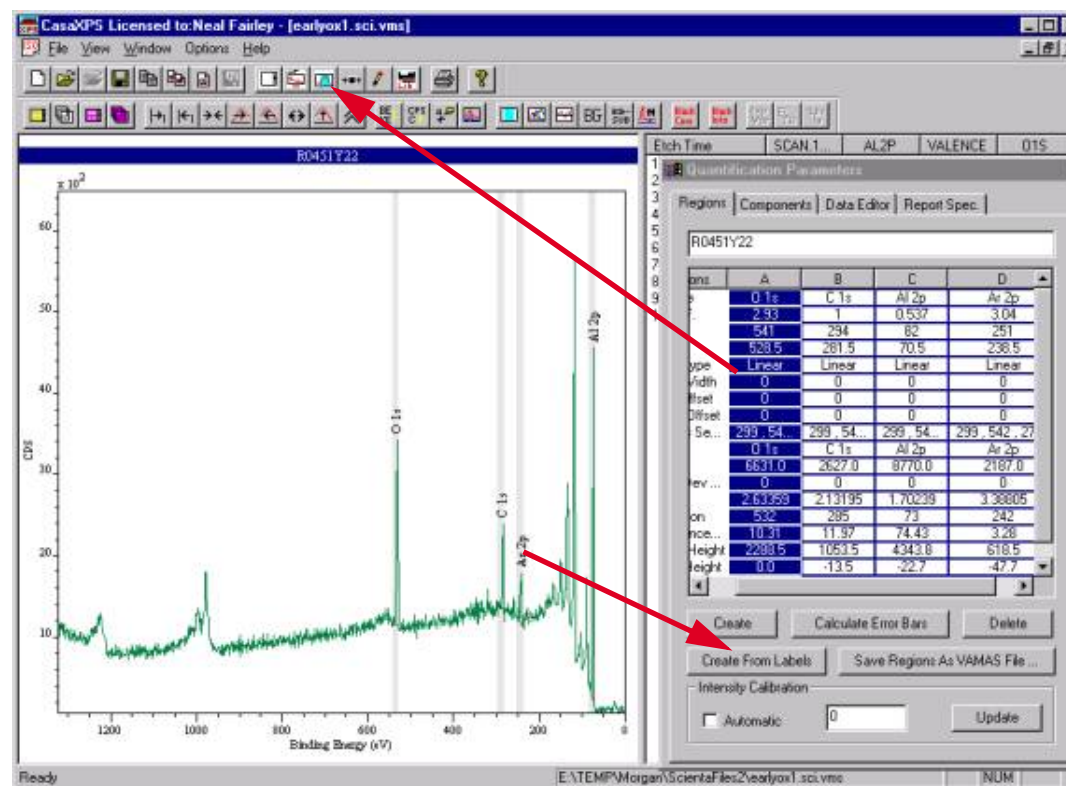


Figure 9: Creating Quantification Regions

Step 6: Adjust region start and end points

The Regions Property Page is used to manually adjust the values for the integration regions. The table of parameter fields offered in the scrolled list may be adjusted by select-

ing a field within the table using the left-hand mouse button and the cursor position. If a field can be altered, then selecting that field with the mouse will both highlight the column of parameters for that region and cause the selected field to become a text-edit field. Any changes are accepted only after the return-key is pressed.

Although manual input is achieved through the Regions property page, an easy way to check the start and end points for each integration region is to use the Zoom List toolbar buttons and adjust the integration limits under mouse control. Provided the Regions Property Page is the visible page on the Quantification Parameters dialog window, the inte-

gration regions on the active spectrum will be marked by vertical bars. The presence of these bars indicates that the limits for the regions can be adjusted under mouse control. If the cursor is dragged starting at the position of one of these vertical bars, the corresponding region end point will be adjusted when the mouse button is released.

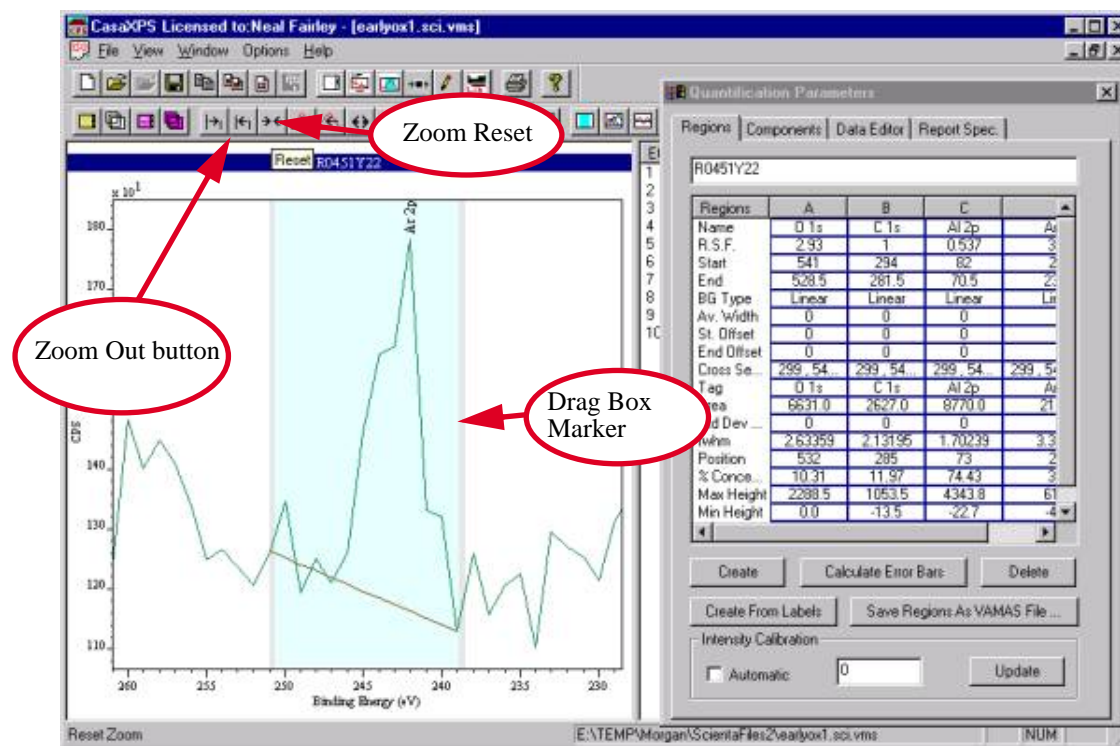


Figure 10: Cycle through the integration regions using the Zoom Reset button followed by the Zoom Out button.

The integration regions on a survey spectrum are typically too narrow to be adjusted in this way without the use of a

zoomed display. To facilitate the use of the mouse, CasaXPS allows all the current integration regions to be entered onto

the Zoom-List and then using the Zoom Out toolbar button (or Ctrl right-click of the mouse) the current set of integration regions can be zoomed into and adjusted.

Press the Zoom Reset toolbar button indicated in Figure 10. The spectrum will be displayed showing the full set of data and any integration regions currently defined for the spectrum will be entered onto the Zoom-List. Now pressing the Zoom-Out toolbar button also shown on Figure 10 will cycle

the set of zoom states and therefore sequentially display the regions.

Step 7: Print the results

Once the elemental composition has been defined using the integration regions, the results may be added to the spectrum display. This is achieved via the Annotation Dialog window (Figure 11), where the Regions property page allows a table

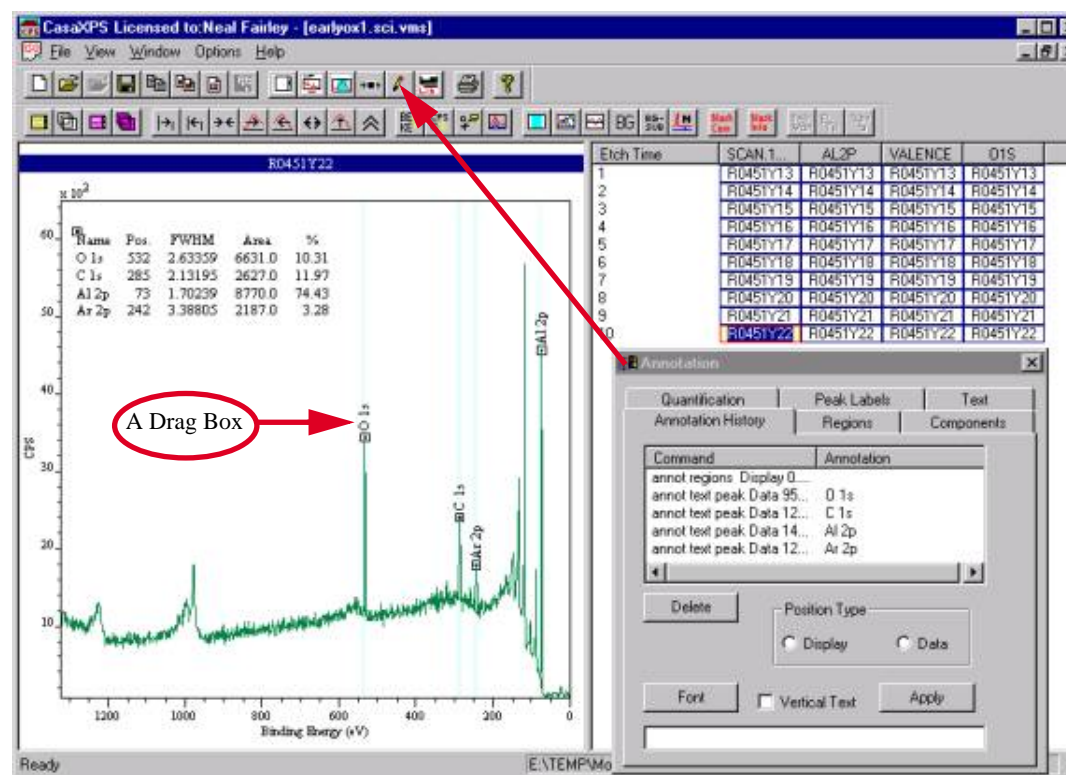


Figure 11: Annotating the spectra using the Quantification results.

of atomic concentrations to be positioned on the spectrum. The entries in the quantification table are updated when

changes are made to the set of integration regions so there is no problem about out-of-date quantification results being re-

ported over a spectrum. If the Annotation History property page is the active page on the Annotation dialog window (Figure 10), then each annotation item will be displayed with a small position box attached. The annotation may be repositioned using these boxes. This is achieved by pointing the cursor at a box and then, with the left mouse button held down, the cursor is dragged to a new position. The annotation will move once the mouse button is released.

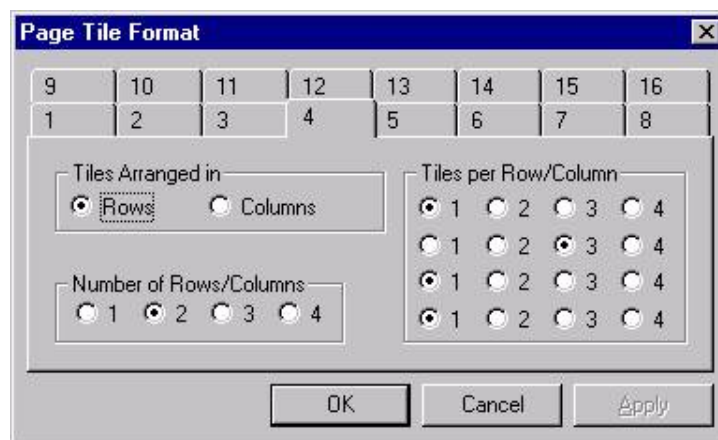


Figure 12: Page Layout dialog window.

The spectrum shown in Figure 11 is now ready to be printed. Any spectra visible in the left-hand-side of the experiment frame (in the Display window) will be printed on a single page. If the Display window contains a scrolled list of spectra then each tile within the scrolled list will be printed as a separate page. If more than one spectrum is required on the same printed page then the Page Layout dialog window offers the means of arranging up to sixteen tiles on one page (Figure 12 and Figure 13).

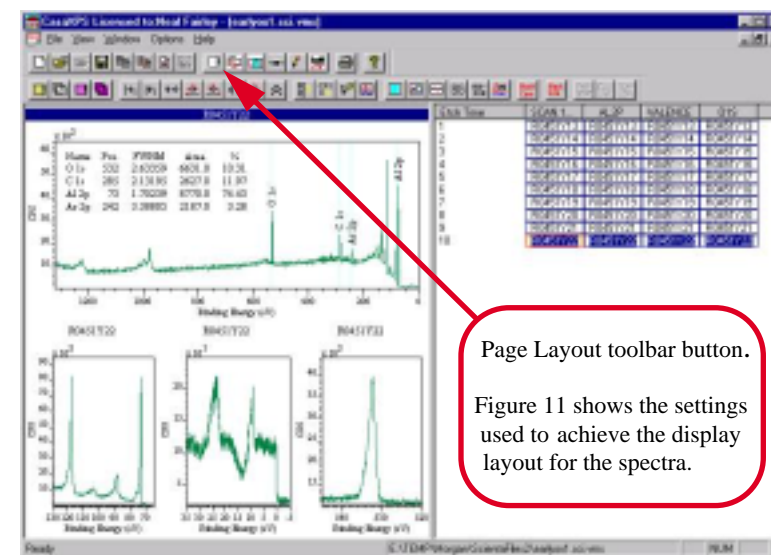


Figure 13: “Tiling” a printed page is achieved using the Page Layout dialog window

Now that we’ve completed “the tour”, we’ll consider all the CasaXPS modules in outline in the summaries below and in much greater detail in the chapters that follow.

Data Display and Browser Windows

Modern surface analysis instruments produce data at a high rate that makes efficient data reduction an important consideration when assessing the sample throughput for a laboratory. The time spent processing the data can become a

serious problem when an automated acquisition sequence has produced hundreds of spectra and spectrum regions all of which fall into sets with related features. The data reduction often consists of the same processing/quantification operations tediously repeated for each of the spectra within a data set. The results of these calculations often have to be graphed or tabulated to show the variation of concentrations as a function of some experimental variable (e.g. etch time or sample tilt).

CasaXPS has been designed based on the ISO 14976 standard format. The structure imposed by this data format permits acquisitions to be saved with the ancillary information necessary for understanding the context of the data, and so processing software now has the opportunity to operate on the data with the same degree of automation that is available to the acquisition routines.

The Browser in CasaXPS presents the ISO blocks in a tabulated (matrix) form, in which each row of the table represents data acquired with the same experimental variable (e.g. ion gun etch time). The columns are spectral regions that have the same chemical species label (element symbol or formula) and transition (detection energy or level notation).

Individual spectra or complete sets of them can be selected and displayed in a scrolled list, and all processing and quantification operations can be applied to a current selection. The Browser thus forms the basis for automatic processing of the data.

It is also through Browser operation that quantification reports are generated again reducing the time to complete the analysis of both simple and complex samples.

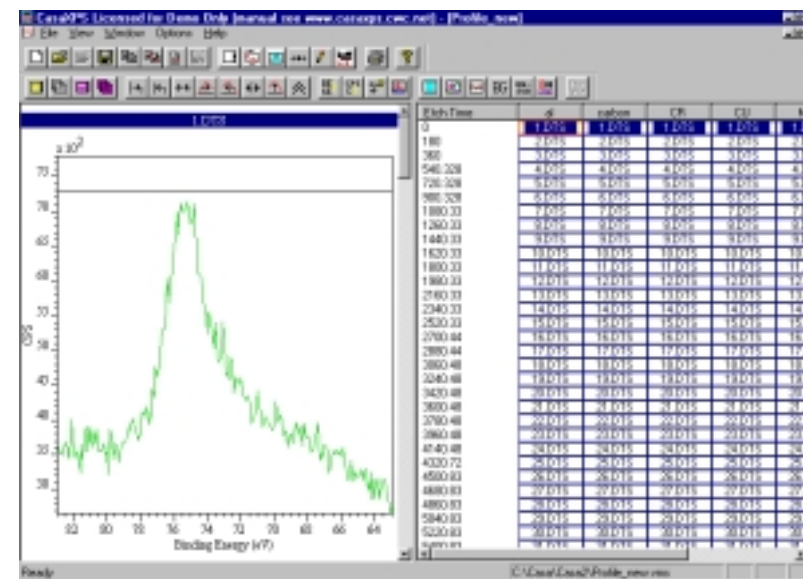


Figure 14: The CasaXPS Program Frame, containing the Toolbars and a single Experiment Frame, which shows both display and browser (splitter) windows.

Element Library

An element library is maintained by the CasaXPS system. It allows the user to manage the information that describes the X-ray induced transitions, such as binding energy, peak widths or relative sensitivity factors. These quantities, once entered, can be accessed through either a periodic table interface or an energy-ordered scrolled list. Information se-

lected via these mechanisms can be used to annotate the spectra or create quantification objects such as regions or synthetic peak shapes.

Graph Annotation

Although Casa XPS provides Microsoft enhanced metafiles for displaying spectra via Word and Excel, XPS specific annotation and other important information can be added to the display through the Casa XPS options. These include:

- Tabulated results and summaries extracted from the quantification regions and components.
- Peak labels derived from the element library.
- Basic text overlays.

Annotation may be positioned on the graphs either with respect to the axes or the data. This allows annotation to be positioned at a fixed location relative to the axes, independent of zoom-state, or to follow the data as a peak is expanded to fill the display area.

Processing

Understanding the data envelope - the peak shape for any transition - can sometimes be assisted by smoothing, differentiating or integrating the data. Techniques for providing these alternative views of the spectra are made available to the users of CasaXPS through the processing option. The spectra can be calibrated with respect to energy and intensity. The latter requires the transmission characteristics of the

instrument to be included in the ISO 14976 file as one of the corresponding variables within a data block. Some manufacturer's data acquisition systems routinely provide output in this important form.

Principal Component Analysis and Target Factor Analysis are also available as a processing options, as is a Monte Carlo method for uncertainty analysis.

Quantification

Spectra can be divided into quantification regions which form the basis for estimating the background shape associated with the recorded data. A sample may be analysed based on these user defined energy ranges or fine structure within a data envelope investigated using synthetic components. Intensities for the elemental composition can be combined from regions, components or via an arithmetic expression that involves any combination of the region or component intensities.

Tables generated from the quantification process are saved to disk in a TAB separated ASCII file. Spreadsheet programs such as Excel will read these files and provide many options for printing the data in either tabulated form or as a graphical profile.

ISO 14976 ("VAMAS") Files

The ISO Surface Chemical Analysis Standard Data Transfer Format (ISO 14976) is a prescription for presenting the ac-

quisition parameters and data for a range of instruments and techniques. CasaXPS is only concerned with X-ray photoelectron spectroscopy (XPS) and as such whenever the term ISO spectrum or “VAMAS file” is used in this manual it is intended to mean, that subset of ISO 14976 corresponding to the fields used by XPS normally stored in ISO format. The “VAMAS” appellation often used loosely and interchangeably with ISO 14976, arose because the early work on standard data transmission format was sponsored and encouraged as part of VAMAS, the Versailles project on Advanced Materials And Standards¹ from the early 1980’s - see for example “VAMAS Surface Chemical Analysis Standard Data Transfer Format with Skeleton Decoding Programs” W.A. Dench, L.B.Hazell, M.P.Seah, and the VAMAS Communi-

1. <http://www.vamas.org> - TWA-02 is the Surface Chemical Analysis work area with secretariat in the USA

ty, *Surf. Interface Anal.* **13**, 63 (1988). (The present ISO standard is closely related to the VAMAS original, but not quite the same). It is not intended to mean here that every variation of ISO XPS format is supported within CasaXPS - there are many allowed variations even though the basic skeleton structure is well defined and invariant. While every effort has been made to support a range of XPS formats (VAMAS “flavours” as interpreted by particular systems) it has not been possible as yet to accommodate the entire set of combinations and permutations that are allowed by the complete file specification.

Many of the processing techniques rely on equally spaced abscissa. For this reason, irregular ISO files are not supported. That is to say, although the full set of ISO file formats can be read by CasaXPS (excluding the annotated form), only a subset can be displayed and manipulated.

The ISO 14976 Surface Chemical Analysis Standard Data Transfer Format² was designed to enable uniformly encoded transmission of data originating from a variety of instruments, which in turn might employ a variety of techniques, to any appropriate processing system, local or remote, for its interpretation within specific applications. A desire to unify descriptions for SCA data and so reduce the number of programs necessary to manage the increasing number of proprietary formats, led to a specification of how data files should be ordered (formatted) and encoded. This work, identified and steered by a Technical Working Party of the VAMAS project (see “ISO 14976 (“VAMAS”) Files” on page 17) in the early 1980’s, was taken up by the International Standards Organisation (ISO - Technical Committee TC 201, Sub-committee 3) and an internationally agreed standard was produced.

2. ISO 14976:1998 - Surface Chemical Analysis - Data Transfer Format (ISO/TC 201) ISO Marketing Services, Case Postale 56, CH-1211, Geneva 20, Switzerland.
<http://www.ISO.org>

One natural consequence of defining a standard data format independent of data origin is that the file must contain all the information required to process (interpret) the data. An ISO 14976 file is therefore much more than just a convenient way of uniformly encoding output from surface analysis instrumentation; it is also a blueprint for anyone designing a data acquisition system.

The strength of the ISO file format lies in its ability to store all the information needed to process the data at a later time. This is also a weakness since its completeness makes the detailed description formidable to the uninitiated. Nevertheless, in the fullness of time, anyone designing a file format for XPS data will surely move towards a structure that essentially holds the same information as ISO 14976; albeit perhaps in a different order or layout.

Partially Encoded Format Versions

ISO 14976 is based on the earlier VAMAS format but with three major changes:-

- the spectrometer geometry description is now referenced to a right-handed rather than left-handed co-ordinate system,
- the number of entries in the “parameter inclusion or exclusion list” has been set to zero in order to simplify the format,
- and the line separator sequence has been changed from the single 7-bit ASCII character “carriage return” to the more commonly used two character sequence of “carriage return” followed by “line feed”

The ISO format thus may be a sub-set of the VAMAS definition.

Although both the VAMAS and ISO formats are very general, CasaXPS is only concerned with files generated from X-ray Photoelectron Spectroscopy (XPS) data. That is to say, while CasaXPS will read files in the format defined by ISO 14976, display and processing within the system are strictly only available for a subset of the Format. This should pose no significant difficulty for virtually all practical situations.

Partially encoded versions of the Format - essentially “templates” for the most common (and most useful) instances - were developed as annexes to the VAMAS definition and other early reports: e.g. Appendix 2 of the National Physical Laboratory Report DMA(A) 164 (July 1988). This is maintained in the full ISO standard (Annex C) The first of the partially encoded formats (C2) is particularly appropriate to CasaXPS. It describes the fields for defining “an experiment involving a number of regular-scanned spectra or spectral regions for one technique as a function of one experimental variable, the analysis not being at a specifically-addressed

point on the sample”. This partially encoded format is suitable for experiments such as XPS depth profiles (either using angle resolved methods or ion bombardment etch sequences) or surface chemical analysis over time with or without temperature variations (see “Appendix 1: ISO 14976 format files annotated” on page 143).

File Structure

What is defined by ISO 14976 is an *experiment*. This consists of a series of parameters which apply to the measurement procedure as a whole, followed by a number of *blocks* of data (often called data blocks), followed by an experiment terminator. Each block consists of a set of parameters that only apply to that block, followed by a series of ordinate values which may represent a curve (e.g. a depth profile), a spectrum, or a map. The experiment header contains the context for the acquisition sequences (blocks) that follow, while the information held in the blocks is specific to a particular determination (e.g. a spectrum region) recorded in the course of an experiment (Macintosh computer users may see some resemblance to the “resource fork/data fork” file structures of MacOS here).

An experiment described in this way is suitable for both complex and simple data sets. The header and block structure may seem perhaps over-elaborate for a single spectrum experiment yet even the most complex profiling analyses are well supported by the same organisation in the data; the only difference lies in the number of blocks within each file.

Experiment header

The parameters which apply to the experiment as a whole occupy the first section of a file written in the ISO format. They fall into the following groups in the order in which they appear:

- identity of the experiment in its (analytical) environment
- optional comments
- experiment mode
- number of blocks and how they are arranged
- pointers to manually-entered parameters
- future-upgrade parameters

The layout and interpretation of these parameters is straightforward.

Data Blocks

The parameters which apply to a block (the “business end” of the format) fall into the following groups, in the order in which they appear:

- identity of the block in the experiment
- identity of the sample
- date and time
- optional comments
- technique
- analysis
- analyser

signal recording parameters

sputtering

sample orientation

additional parameters and future-upgrade parameters

(followed by the corresponding data values)

Each block thus provides the acquisition parameters together with a set of “corresponding variables”. The corresponding variables may contain more than just recorded counts. Additional corresponding variables can include values such as transmission function (correction factors) for adjusting the recorded counts with respect to a reference spectrum. Use of a second corresponding variable is the method preferred by at least one commercial manufacturer to record transmission function data in a “VAMAS” format. The transmission characteristics for their instruments accompany the data point by point within each block.

Although blocks contain a comprehensive set of parameters, it is permitted to replace many by a value that indicates no information is present (“1E37”(10³⁷) is the defined value). Use of this feature can reduce the ability of a target system to process the data, but it can help to limit the complexity of writing a valid ISO file for an acquisition system.

Comment lines may be included in each block. These are useful for summarising technique specific information associated with a particular instrument. For example, names corresponding to lens modes (magnification) could be entered as a comment line. Operational parameters could accompany such text, where the state of the charge neutraliser or any other relevant descriptive information may be added.

CasaXPS provides an option on the “Tile Display” dialog window for the display of block comments as a header for

the graphical display. Suitably formatted comment lines may be used to annotate the display with, e.g. acquisition parameters for a spectrum.

Binding vs Kinetic Energy

An important piece of information that is not explicitly stated in the ISO format is the nature of the energy scan. CasaXPS needs to know whether the energy is in binding or kinetic energy. The ability of the display to switch between binding and kinetic energy is dependent on the “analysis source characteristic energy” containing the correct value and the correct assignment for the energy type.

A REGULAR scan is defined by “abscissa start” and “abscissa increment” energy values. To differentiate between the two types of energy the associated “abscissa label” field is used by CasaXPS as an indicator of which type has been used to describe the data. If the label includes one of the substrings “B.E.”, “BE”, “Binding” or “binding”, then the energy type is assumed to be binding energy. Otherwise the energy type defaults to kinetic energy.

Experimental Variable

The numerical value for the experimental variable is recorded in each block. Associated information for the value entered in the block is specified in the experiment header section of the file, where a label and the units for the variable are found.

A proper use of the experimental variable is essential when multiple blocks are recorded in a file with the intention of quantifying trends in the data. Reports generated by

CasaXPS rely on these values to organise the blocks so that profiles are meaningful

Transmission Correction and Quantification

Modern XPS instruments are relatively complex devices, characterised by numerous operating modes and methods. An ability to control and configure analyser input lens operation enables use of a range of energy resolutions, spatial resolutions, and magnifications, but for each operating mode of an instrument, transmission characteristics will differ and also perhaps vary in a complex fashion within any particular scanning mode, causing energy-dependent differences to occur between spectra taken from the same sample but using different operating conditions. These differences in data translate into quantification results that depend on the instrumental settings as well as the sample itself and so an intensity calibration procedure is required if consistent results are expected from the same instrument operating within its range of appropriate settings.

The most common method for comparing peak intensities is to use tabulated Relative Sensitivity Factors (RSF) to adjust the calculated values for physical effects. Without intensity calibration procedures based upon transmission characteristics these RSF values are only appropriate for a specific operating mode. It is therefore essential to map out the transmission functions that accompany the lens modes and pass energies before consistent results can be obtained from a particular sample.

Transmission Functions

Transmission functions are dependent on instrumental settings that can be adjusted by the user and other settings that change with age. A useful calibration procedure should therefore be easy to perform and apply. Moreover since transmission functions vary with time, the transmission function for a particular spectrum should be included as part of the data file.

State-of-the-art instruments typically record the measured intensities and extract some functional form for the transmission characteristics; both pieces of information are saved to disk. The methods adopted vary: for example, a piecewise linear approximation is used by one manufacturer, where the nodes may be determined either from peak intensities or from the inelastically scattered background. Another manufacturer uses a polynomial form for the logarithmic shape of the transmission as a function of retard ratio, where the coefficients of the polynomial are determined using peak area measurements rather than background information. A third manufacturer again uses peak area measurements to determine two analyser-dependent constants a and b as follows:

$$T(R) = \left[\frac{a^2}{(a^2 + R^2)} \right]^b \dots\dots\dots \text{Eqn. 1}$$

R is the retard ratio⁽³⁾.

NPL offers an empirical intensity calibration (standardisa-

3. Tanaka A., J. Surf. Analysis, 1 189 (1995)

tion) procedure where the output takes the form of a rational function of two polynomials. The background of inelastically scattered electrons measured from pure gold, silver and copper provides a vehicle for comparing the response function of an instrument under test with that of a well characterised “metrology spectrometer”. Given spectra acquired at the various operating modes of the instrument to be calibrated and standard spectra from the reference machine, then relative transmission functions are readily constructed and consistent quantification becomes possible.

It is clear that there is no absolutely correct way of representing transmission characteristics for a given set of intensities, although the importance of this information should not be underestimated. Proper intensity calibration opens up the opportunity for consistent quantification not just across operating modes of a single instrument but also for instruments from the same manufacturer and possibly instruments from different manufacturers. Analysts reliant on data taken from a range of instrumentation will be aware of this problem.

Although there is no prescription for transmission functions within the ISO 14976 file format standard, the mechanisms are in place that enable recording of more than one corresponding variable per ISO block and therefore the energy dependence of the transmission can easily be saved along with the energy dependence of the signal. ASCII files have come-of-age with the increasing size of hard disks and “zip” technology. There is no longer a serious concern about the size of (XPS/AES) spectroscopic files and the benefits of recording open and explicit data outweighs any drawbacks that may accompany ASCII format spectra.

Quantification

Historically transmission correction has been performed via adjustments to the RSF values used to quantify the peak intensities. While at first sight this would seem a reasonable approach, the nature of transmission functions means that errors are introduced by this procedure. The RSF attempts to allow for the set of factors that cause peak intensities to vary for reasons other than sample composition. The influences of the transmission function are usually assumed to be constant across a peak, but this assumption is often false in the case of the relative transmission functions. The response of the instrument can vary rapidly across a peak width, particularly when portions of spectra are compared at opposite ends of the resolution or energy scale.

Empirically determined RSF values necessarily compensate for instrumental response in the reference mode. Alternative modes need to be adjusted relative to the reference mode by dividing spectra point by point by the transmission function rather than choosing a representative value for the transmission (typically evaluated at the peak maximum). The difference between intensity values determined by these two

approaches can be outside the precision estimates for the quantitative values. A consistent approach to transmission correction therefore goes hand in hand with proper procedures for monitoring the response of an instrument.

Provided a transmission function is included in the data file, CasaXPS offers both methods for adjusting spectra with respect to the transmission function. The default method is to use a single point in the energy scale to evaluate the transmission function. This single value is then used to scale the RSF value before applying this value to the calculated peak area. The main reason for offering this as the default mode is that most manufacturers systems perform the quantification by this means. The results from CasaXPS in the default mode for transmission correction therefore will agree with these other systems, although it should be stated that correcting spectra point by point prior to calculating the intensity is probably the better method.

An example of the way in which transmission can be encoded within the ISO 14976 file format (and thus used by CasaXPS) is given in Appendix 1 (see page 143).

CasaXPS uses a multiple document interface (MDI) as defined by Microsoft. That is, it conforms to the same style as programs such as Microsoft Word or Excel and therefore presents each file (Document) with a document view. The main frame of the program manages a set of document/file frames. These file frames may be minimised or maximised, or arranged using the Window menu allowing the frames to be tiled with respect to one another or to create additional views of the same file.

Understanding the Data

The ISO 14976 files used by CasaXPS are viewed through a “splitter window” called the Experiment Frame. The right-hand section of the Frame (the Block window) represents the logical structure of the experiment; showing the relationship of the spectra to each other and to the experimental variable. The “splitter” terminology is a reminder that the relative sizes of the two components of the frame may be adjusted by “dragging” the central divider - the “splitter bar” inside the

frame, allocating a larger or smaller area to the graphical display (the spectrum) or the block structure, as appropriate. Data held within a VAMAS (ISO) file appears in rows and columns in the Block window, each row corresponds to data acquired under the same experimental variable value while the columns correspond to spectra (regions) resulting from measurements on the same chemical species or transition (energy range). The mouse is used to select a set of one or more blocks from the right-hand side (the Browser) and these data blocks are displayed by selecting one of the two leftmost toolbar buttons shown in Figure 15

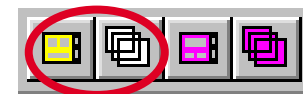


Figure 15: Spectrum Display Options.

The first icon will display a scrolled list of spectra (or regions), one per tile. The second of the two allows (many) spectra to be overlaid in the same tile. In both cases the x and y scale values are determined by the maximum range values for the spectra or regions requested. The corresponding but-

tons on the right of the toolbar provide display and overlay within the already set scale values. If there are no points in the range, then a message indicating “No Data Points in Window” appears (this facility enables the overlay of (detailed) regions onto a wide scan, for example).

Selecting the ISO 14976 Blocks

Each ISO file requires at least one Experiment Frame. Experiment Frames are created using the “New” option on the file menu. Once an empty Frame is selected, the “Open” option from the same menu offers a File Dialog window for choosing the ISO file.

The Browser view (right hand pane of splitter window) offers the data as described above. Selecting the data is achieved via the mouse in combination with the “Shift” and “Ctrl” keys. Positioning the cursor over a name in the browser view then clicking the left-hand mouse button selects that block. If a range of contiguous blocks is required (as seen in the browser view), then select the first block in the range, followed by holding down the “Shift” key and selecting the last item for the range. All blocks between the first block and the last will be selected. The range selection works across columns as well as within rows or columns. That is to say, rectangular sets of block can be selected by this mechanism. Additions to the current selection require the use of the “Ctrl” key. If the “Ctrl” key is held down and a new name is selected, the existing selection is retained in addition to the block just chosen. If the block indicated with the mouse is already selected then it is removed from the current selection. The “Shift” key and the “Ctrl” key when used together allow rectangular sets of blocks to be added or removed

from the current selection.

Selecting a block without the use of the “Shift” key or the “Ctrl” key clears the current selection and replaces it by the indicated block.

Zooming the Data

Once a spectrum has been entered into the left-hand side scrolled display window the data can be examined under mouse control. In order to zoom into a range of the data, the mouse is used to mark a rectangular region and then the third button is pressed (Figure 15) The result of this action is to cause the graph axes to be computed from the defined area and the data is redisplayed. Zooming may be performed a number of times and the system maintains a list of zoom states. These zoom states can be revisited using the fourth toolbar button. To reset the list of zoom states press the fifth toolbar button. The axes are returned to the state when the data was first displayed in the scrolled list.

Tiles of Spectra

A printed page corresponds to the visible area associated with the left-hand scrolling window. Each click within the scroll bar moves the display by one page whilst dragging the scroll button allows the pages to be moved at will. The arrow buttons on the scrollbar permit fine adjustments to the position of the spectra in view.

Pages can be divided into sub units referred to as tiles. The format for the page is defined using a dialog window available from the main frame menu headed “Options”. If “Page Layout” is selected, the dialog box presents property pages that allow the user to choose the number of tiles per page as

well as how many rows and columns of such tiles should appear on the page. These pages represent an initial format for the number-of-tiles per page; however, the user is at liberty to adjust these as seen fit.

There is a maximum of sixteen tiles per page. The layout is specified first by selecting whether the tiles are to be displayed in “rows” or “columns”, (this permits the layout to be transposed at a click of a button) and then by specifying the number of rows or columns, and finally by indicating, using the “radio button” array, the number of tiles in each row or column (see “Page Layout” on page 136). Sixteen predefined layouts are provided, corresponding to arranging spectra in symmetrical rows up to the maximum (16) allowed. All these predefined formats may be changed to provide a “custom option” library of preferred layouts, as required.



Figure 16: Main Toolbar

Figure 16 shows the Main Toolbar. The dialog boxes available from the main menu “Options” are also displayed using buttons on this Toolbar (see, e.g. “Options Bar” on page 134). Move the cursor over a toolbar button; a hint describing the action associated with the button will pop up and a slightly longer description will be displayed in the main window status bar (see page 132).

Problem: The first release of CasaXPS limits the number of pages that can appear in the scrolled view. Although the modern PC hardware is 32-bit, some 16-bit code still remains in MS Windows software, in particular the class used by the routine that sets the scroll bar attributes, does not like the logical scrolled

area to exceed the range of a signed 16-bit number (32767). Sorry! If more tiles are required to view a set of spectra, then the page layout mechanism can be used to reduce the overall size of the scrolled list.

Tile Display

Many of the attributes for displaying the spectra can be adjusted. The menu item labelled “Tile Display” on the “Options” menu displays a dialog window for modifying the appearance of the spectra.

Ranges for the axes, labels and fonts used for the labels can be adjusted on the first two property pages found on the “Tile Display” dialog window. The user may also choose between displaying the spectra using binding energy or kinetic energy, as well as counts per second or recorded counts.

An option on the “Y Axis” property page allows overlaid spectra to be offset with respect to one another. A check box enables the offset mechanism and a numeric-field permits a percentage to be entered which changes the separation of the traces. Other check boxes offered on the “Y Axis” property page alter the appearance of the display. The axis label and scale may be toggled on or off; also a spectrum may be plotted with the value of the experimental variable displayed at either side of the graph area, and the spectrum background may be subtracted from or included with the display (see detail of command options on page 136).

An alternative way of displaying overlaid spectra is via the “Geometry” property page. Here the user can select between

2-D and 3-D views of the spectra. The latter option plots the spectra as a function of the experimental variable. The first spectrum is plotted in the front plane of a cube while the last spectrum in the selection is plotted at the back plane. All spectra in between appear at a plane that is determined from the value of the experimental variable for the VAMAS block.

The 3-D plot is adjusted using parameters that shift and scale the two planes associated with the front and back of the cube. Figure 3 shows a set of spectra plotted against etch time. The front plane is 75% of full size in both X and Y directions, while this reduced axes area has been shifted left by the maximum amount by specifying -100% for the shift. Note that the front plane can only be shifted around in X while the back plane may be shifted in both X and Y.

A third option for data display allows a set of spectra to be viewed in terms of abstract factors. This involves performing a Principal Component Analysis (PCA) on the spectra. Please see the section on Principal Component Analysis for a description of PCA and the use of this display option.

Other information drawn on a tile can be changed through the “Display” property page. The font and text for the title of a tile may be adjusted. The optional header information font may be changed similarly. A range of optional display items can be toggled on/off allowing the spectrum to appear, for example, with or without quantification regions present.

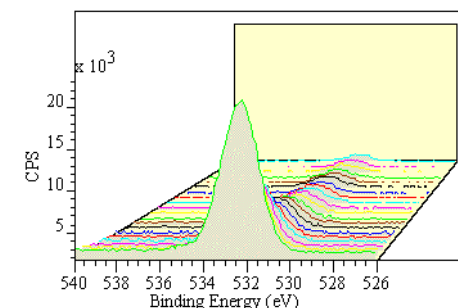


Figure 17: 3-D plot.

Colours

The “Colours” property page provides the means to set the colours used throughout the system. Spectra, backgrounds to spectra, synthetic components, regions and residual plots may have the colours adjusted. Fill colours used to display 3-D plots are also changed through the options on this page. Selecting a button on the “Colours” property page brings up a dialog window that allows the existing colours to be viewed and new values set. In the case of spectra, sixteen colours may be chosen. These colours are displayed in the “Custom Color” section of the dialog window and changed using the “Define Custom Colors” expanded form of the dialog window. To change a colour within the “Custom Color” set, first click on the colour you wish to change, then select a new colour, either from the default palette or using the colour values on the expanded section of the window. On pressing the button labelled “Add to Custom Colors” the

colour square first selected is set to the colour just defined. Each colour subsequently specified and the “Add to Custom Colors” button pressed will cause the next custom colour cell to change. The cells are updated in a top to bottom, left to right order. However, the colours assigned to the graphs are assigned on a left to right, top to bottom order. Some colour selections are for a single colour. In these cases, clicking on any colour cell, so that the focus box surrounds the intended cell, followed by selecting the “OK” button on the dialog window will activate the “Apply” button on the property sheet for the tile parameters. The colour changes only take effect when the “Tile Display” window is applied, either by the “OK” button or the “Apply” button.

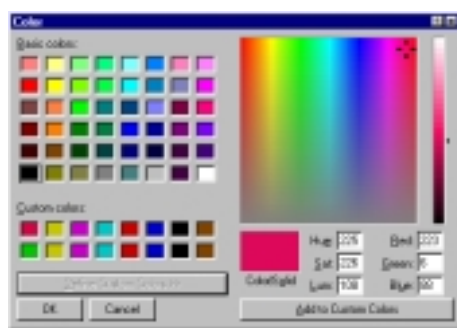


Figure 18: Colour definition window

Fonts

Typefaces (“fonts”) are managed in a similar way to colours. Buttons are provided for fonts associated with axes labels, the title and header text. The procedure for adjusting a font is to select one of the buttons labelled “Fonts”. The title, for example, is located on the “Display” property page. To adjust the font, press the button next to the text entry field that offers the current value for the title; a font dialog window appears. Then select the font parameters from the dialog window and press the “OK” button. Once again, like the colour dialog window, this action will activate the “Apply” button on the property sheet for the tile parameters. The font will only be changed if the “OK” or the “Apply” button is pressed.

Display Parameters and Scrolled Tiles

Tile display parameters can be applied to all the tiles in a scrolled list. A property page headed “Global” allows a toggle to be set that causes the settings currently active to be transferred to all tiles when the “Apply” or “OK” button is selected.

Two additional toggle buttons enable the transfer of ranges for the X-axis and Y-axis. It is not always desirable to include these display parameters when the other characteristics are globally applied, but on occasion it can be useful. For example, setting the Y-axis range to that of a specific size for a set of similar spectra can provide an interesting

visual effect when scrolling is used to view the data. The variation in intensity as a function of an experimental varia-

ble is highlighted in this way.

An important feature of CasaXPS is the ability to write Enhanced Metafiles. This permits data to be visualised through the data system, then exported in to other programs in order to provide the final form used in a document or report. Enhanced Metafiles supply the drawing information used to display the data in CasaXPS to a word processor such as Microsoft Word, which can then be used to edit the drawing in remarkable detail. The button to save these graphics files is indicated below (red circle).



Figure 19: EMF & Annotation buttons in Main Toolbar

Although the tile display options permit text, such as the title, to be adjusted and enhanced metafiles allow sophisticated visual editing, XPS spectra can sometimes best be annotated with information derived directly from the data. The most obvious being tables related to the quantification of the chemical state. CasaXPS therefore provides some options that allow the inclusion of such information as part of the annotation.

The options for adding annotation are found on the main menu “Options” or via the main toolbar (Figure 19 - magenta square).

Quantification Tables

After quantification regions have been defined for a spectrum, a tabulated form of the information gathered from data may be added to the graphical display. The procedure to include region information is similar to that used for synthetic components. Both types of tables are added using the “Region” or “Component” property pages found on the “Annotation” dialog window. The font used for the table may be selected before pressing the "Apply" button on the page of interest. A table of the current values derived from the source in question will appear in the display area, and it is then repositioned via the "Annotation History" property

page found on the same dialog window.

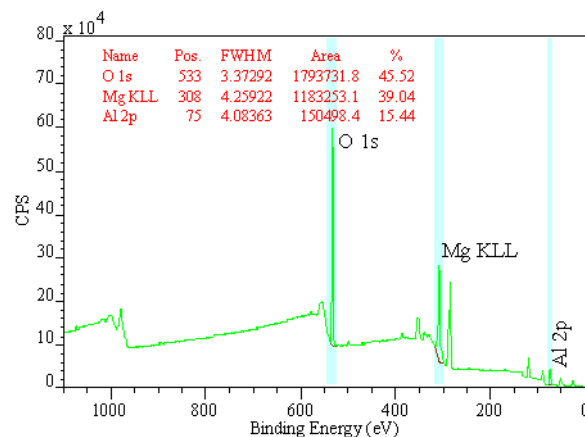


Figure 20: Annotation using region information

Moving Annotation and the History Mechanism

Annotation, either tables or text, can be moved using the list of annotation entries currently defined for the active spectrum. The spectrum display will be updated with the annotation now at the new location. Position for annotation items may be referenced to the display tile frame or the data itself (for convenient placement e.g. when multiple blocks are displayed in overlay) and labels may be orientated vertically or horizontally. Positioning targets (small squares with central dots) may be set with the mouse anywhere within the display tile and disappear when the annotation window is dismissed. Other attributes for the annotation can be adjusted using the history list. Select the line describing the annotation for

which the change is required. Then adjust the settings using the options found below the history list. Press "Apply" to see the result of the changes.

Text lines may have the orientation adjusted in addition to altering the contents via the text edit box. Fonts and colour are reselectable through the history list and the edit options on the "Annotation History" property page.

Entries on the "Annotation History" property page are in the format used to save the annotation in the ISO 14976 file. A block of data within the file includes a number of comment lines. It is within these comment lines that processing, annotation and quantification information is stored. The idea of offering the full description used to save the data to file is that the parameters can be understood (if desired). It is then possible to save the data and include CasaXPS information from the start.

Peak Labels

Spectra can be labelled using the names from the element library. The element markers currently displayed are made available for annotating the spectral features on the "Peak Label" property page of the "Annotation" dialog window. The element marker text is entered in the scrolled list, found on the page of interest. To annotate the spectrum, the element names must be selected within the scrolled list and the desired attributes chosen. The annotation text entries are added to the annotation history list when the "Apply" button is pressed. Again, the exact location for the annotation will need to be assigned through the normal positioning proce-

dure.

Note that “Peak Labels”, derived from the entries in the Element Library, are different from “Peak Markers”, generated by the “Find Peak” button in the Element Table tab of the Element Library module, (see page 141). These markers serve only as a guide to the presence (or absence) of a feature in the spectrum (a peak maximum) which you may want to identify, for example, by using the Element Library routines, as described on page 9. Markers may be set and cleared independently of Peak Labels (use the “Clear Elements” button for the Labels)

Text Annotation

Peak labels are merely multiple instances of text annotation. Individual lines of text can be added to the display using the "Text" property page. A text edit box allows the desired text to be entered and any attribute, such as, vertical/horizontal, font or colour may be chosen before pressing the "Apply" button.

An XPS spectrum represents an envelope that is derived from many unresolved peaks. The relative intensity of these peaks may vary as a function of the kinetic energy imparted to the electrons being recorded as well as the characteristics of the electron optics (lens mode/pass energy). In addition, the position of the peaks may differ from the expected value due to a combination of instrument energy calibration and charging effects on the sample. CasaXPS therefore provides a number of processing options that assist in understanding the data envelope and allow adjustments to the spectrum to enable compensation for these analysis variables.

The processing options are available from the "Options" main menu item or from the main toolbar button shown in Figure 21.



Figure 21: Main toolbar processing options.

Smoothing

Data smoothing may be performed using a range of Savitzky-Golay algorithms or by weighted averaging the data using a set of weights that are a normalised Gaussian distribution.

Smoothing algorithms should always come with a statistical health warning. Smoothing can seriously affect your curve fitting! The act of fitting synthetic peaks is not only the best form of smoothing, but also makes assumptions about the distribution of errors within the data as well as the statistical independence of the measurements. Smoothing the data before curve fitting invalidates these assumptions.

A Savitzky-Golay filter is derived by approximating the data with a polynomial of degree less than or equal to the chosen number of $2n + 1$ channels. The coefficients for the selected polynomial are determined in the linear least square sense and the datum at the centre of the $2n + 1$ channels is replaced by the value for the polynomial at that point.

To perform a smoothing operation using Savitzky-Golay

methods it is necessary to select the degree of the polynomial and the width of the data-channels over which the approximation is to be made.

Gaussian smoothing merely needs the width of the data-channels to be used when averaging the spectrum.

Differentiation

Differentiation of spectra is also achieved using the Savitzky-Golay polynomial. This is appropriate since noise interference would cause problems to an algorithm used for analytical functions, and so the implied smoothing operation involved in Savitzky-Golay differentiation, would explicitly be required with other techniques.

The parameters used in differentiation are, therefore, identical to those for smoothing the data. To get a feel for what was actually done to the spectrum during the differentiation procedure, a smooth operation with the same parameters, applied to the original spectrum can provide an insight into the shape of the data that produced the derivative.

Integration

A more novel use of the Savitzky-Golay polynomial is in calculating the integral for a spectral range. Integration by any means has implied smoothing involved since it is essentially an averaging process. However, when noise is involved there is little benefit in using sophisticated Newton-

Cotes or other quadrature methods, as these generally involve forcing a function to take on the data values at the corresponding nodes. The virtue of putting a quadratic through three points containing noise is somewhat doubtful. Integrating the data using a least square fit of a quadratic to more than three points seems better, however from a practical perspective, if there is a significant difference between these two operations, then there's trouble somewhere!

To summarise, integration is performed using the Savitzky-Golay polynomial in which a quadratic is used to approximate 5 data channels. From a user's point of view, integration is performed simply by pressing the "Apply" button on the "Integration" property page, with the single option of being able to reset the zero point of the integral, if required.

Energy Calibration

A spectral line may appear at an energy position that is not the expected value for a transition. The reason for this may be due to spectrometer calibration or sample charging, however, for presentation purposes, the energy scale can be adjusted using parameters on the "Calibration" property sheet. Two values can be entered, the recorded position of the line and the required energy for the peak. On pressing the "Apply" button a shift is computed and the display is updated. Note that if preferred, the spectrum can be shifted by an amount by setting the measure value to zero and entering the required shift in energy into the field for the true energy.

Intensity Calibration

A spectrum recorded on a particular instrument represents only the electrons as a function of kinetic energy that the instrument was capable of sampling. The transmission characteristics for an instrument are important if the spectra are used to quantify the chemical composition of a sample. This extends even as far as calculating the background for the data, where techniques presented by Tougaard can be only applied after accounting for such variations in the intensity.

The "Intensity Calibration" property page allows the intensity to be adjusted using a function of a constant power of the energy. The same page on the "Processing" dialog window allows the inclusion of a relative transmission function to be specified.

The ISO 14976 file format does not specify how the transmission function for an instrument can be incorporated with the data, however CasaXPS does allow the use of a second corresponding variable per block to define the transmission behaviour for the instrument (See "Appendix 1: ISO 14976 format files annotated" on page 143.) This is particularly important when a single relative sensitivity factor (R.S.F) is used to quantify results from a range of magnifications and pass energies. Without first correcting for the differences due to operating modes of the instrument, a single R.S.F. would not yield consistent quantification results for the same sample.

If a valid corresponding variable index is specified, the act of calibrating the intensity modifies the data by dividing the raw spectrum by the indicated transmission function.

Processing History

Each VAMAS file data-block can be processed using any of the available options. The "Processing History" property page provides the means of viewing the operations that have been applied to the data. Further, the processing may be reversed entirely or selectively. The "Reset" button removes all the processing operations and restores the data to that originally supplied in the ISO 14976 file. Alternatively, by selecting a set of lines within the scrolled list, then pressing the "Apply Selection" button, the indicated set of processing operations are performed on the original data. (Selection of items follows the normal Microsoft conventions - left mouse button for a single item, left click then shift left click for an inclusive list, or control click for separated items). Thus the previous processing history is replaced by only those actions selected prior to pressing the button.

Processing for a block of data can be globally applied to a set of ISO blocks. Select these blocks within the browser view, then right click on the corresponding display view to reveal a dialog window for propagating operations such as curve fitting and processing. Choose the required actions then press the "Apply" button. A second window appears showing the progress of the propagation and offers a "Stop" button. If the "Stop" button is pressed, the propagation will terminate following the completion of the current action. Note that some actions, for example curve fitting, may take a significant period of time to complete.

Not all actions for processing the data may be applicable to

all the selected blocks. To prevent, for example, the energy calibration for the current block from being propagated to other blocks, the history item can be flagged so that it is not propagated along with other appropriate actions.

To exclude a processing operation from a propagate action, select the item in the history list, then press the button labelled "Propagate Flag". The entry in the list for the selected processing option will change to indicate that the flag for propagating the action is set to "false" ("don't perform this"). Repeat the procedure (toggle) to remove the exclusion.

Please note that although the "**Calculator**", "**PCA**" (Principal Component Analysis), and "**Test Data**" tabs appear (rightly) in the Processing module, they are of sufficient importance and power to merit their own topic sections (see "Using the Calculator and Comparing Spectra" on page 71, "Principal Component Analysis" on page 80, and "Optimisation of the Peak Parameters" on page 47).

The same is true of the **Monte Carlo** procedures, the entry to which appears under the "Components" tab of the "Quantification" module. (see "Monte Carlo Methods - Uncertainties in Intensity Calculations" on page 97).

The "Element Library" dialog window provides the means of managing the information used to identify spectrum features and then create regions and components with predefined values. Strictly it is not an "element" library at all, but a list of spectroscopic lines which, while derived from characteristic transitions within the electronic structure of atoms of a particular element, can also reflect the exact chemical environment of those atoms - that is provide chemical compound specificity. Three property pages within the module offer two views of the data held in the element library file, and a means of loading and merging data into the current session.

The module is launched from the main toolbar, as shown below



Figure 22: Library button in Main Toolbar.

In order to function correctly, CasaXPS should have an element library present in the system (although some basic features will function without one) - but for many commands and operations (involving qualitative and quantitative anal-

ysis) a library file *must* be present. The file may be changed at any time using the controls described below (so that several library files may be held on disk), but only one library (of unlimited size) may be active at any time.

CasaXPS is normally shipped with a relatively simple generic library. Instrument or analysis specific libraries must be generated or imported as required by the user. the module recognises both CasaXPS and Kratos (proprietary) formats.

Library File Structure

The element library file is an ASCII (plain text) file. As such it is open to simple modification or change by the user, so long as changes are made within the (casaXPS) format outlined below.

The file is a series of single item lines, each terminated by a (single character) "return".

The first line is the version number.

Each subsequent entry in the file has items that are included one per line as follows.

1.	Element	GL(30)
2.	Transition	2.000000e-004
3.	Label/Name	Mg
4.	Mass (Daltons)	He
5.	Energy Type (B.E. or K.E.)	1s
6.	Energy (eV)	He 1s
7.	F.W.H.M.	0.000000e+000
8.	Line shape (e.g. GL(30))	BE
9.	Relative Sensitivity Factor	2.500000e+001
10.	Ionisation source (X ray anode)	1.000000e+000

Photoelectric transitions should be entered with item 5 (energy type) set to "BE"; this allows the same energy to be used for different X-ray anodes. Auger lines always appear at the same kinetic energy independent of the excitation source and these values should be specified with energy type "KE". Note that the ionisation source is also specified so that source specific information can be offered for spectra that match the acquisition characteristics found in the library. Line shapes are specified using the abbreviations listed in "Line Shapes and Backgrounds" on page 55

The beginning of the "default" (CasaXPS) library file is shown below:

```

0
H
1s
H 1s
0.000000e+000
BE
1.400000e+001
1.000000e+000

```

```

GL(30)
8.200000e-003
Mg
. . . .
. . . .

```

An entry allows the species and transition to be specified but also permits a name other than these two labels to be also entered. For example, a carbon line corresponds to a chemical state arising from PVC is identified by entering, "C" for the species, "1s" for the transition and "C 1s PVCa" for the name. It is the name that appears as the first item in the scrolled list of transitions.

Although the expected energies for X-ray induced transitions are universal constants, quantification data will be (in practice) specific to an instrument. An element library can be customised through the energy-ordered scrolled list on the property page labelled "Element Table". Right clicking on an item name will produce a dialog window that allows the items associated with the indicated line to be updated. New entries can be created using this mechanism but for

large numbers of new entries a text editor with cut and paste can be the best method of modifying the element library.

Items which are specific to an instrument or instrument configuration include the relative sensitivity factors (R.S.F.), full width half-maximum (F.W.H.M.) and the line shape. Instrument manufacturers will supply some of this information, but individual experience will determine appropriate values for items such as the F.W.H.M. and line shapes. These parameters clearly vary from analysis to analysis and specific figures are present here only to provide “sensible” initial values for the creation of synthetic components - values that are sometimes useful!

Loading an Element Library

The property page labelled "Input" (see “Library” on page 141) allows the session element library to be modified by reading different element library files. A browser button offers a file selection dialog and the chosen file may be loaded or merged with the current session element library. If the "Load" button is selected, the element library in the current session is replaced by the contents of the specified file. Alternatively, the "Merge" button causes the contents of the file to be added to the existing information for the session. A user can maintain a set of element library files and load/merge only those appropriate for a type of sample or operating mode of the instrument. In the absence of transmission function data this mechanism allows the user to maintain

different R.S.F. values for different acquisition conditions.

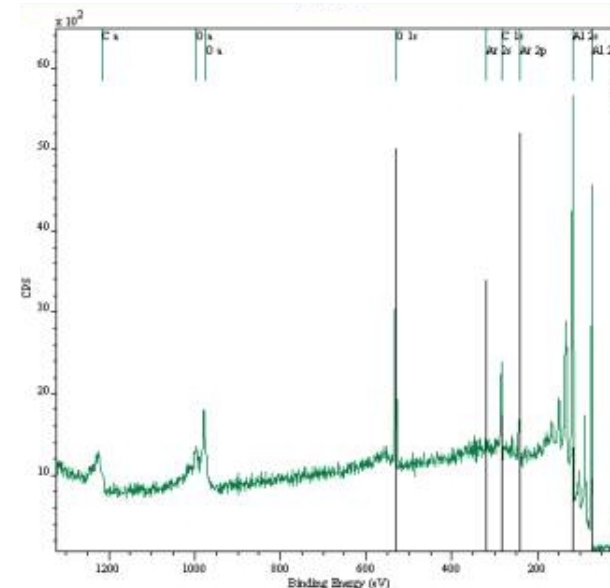


Figure 23: Element Markers.

The **Element Table** window provides a scrolled list of the contents of the Library File: clicking on a name (toggle) transfers a labelled position marker to the display for the selected item. Position markers may be removed from the display with the **Clear All Elements** button. **Find Peaks** adds coloured markers to the display for all major peaks in the displayed block, according to a defined algorithm: the **Clear Markers** button removes these. Take care to note the distinction between element markers (from the library) and peak markers (from the Find Peaks routine)

Identifying Peaks

Peak identification is performed using the scrolled list of transitions and the "Periodic Table" property page. A spectrum may be viewed with markers added at energies where peaks might be expected (Figure 23 - note that the position line sizes (heights) are scaled to reflect the RSF values stored in the library) To add and remove these markers, select an element through either the periodic table interface or from the element table. The markers are toggled on and off by repeatedly selecting the same element.

An alternative method for examining a spectrum is to have the scrolled element table visible and then run the cursor over the tile display window with the left mouse button held

down. The energy order list will scroll with the energy position of the cursor. Release the mouse button when over a peak of interest and then use the table to toggle the peak markers until a match is established.

A more detailed example of how to approach peak identification is given in "A Quick Tour of CasaXPS" on page 7

Peak and element markers are maintained within a session on a spectrum by spectrum basis. Altering the markers on one spectrum does not change those previously chosen for a different spectrum (block) even within the same experiment. Simply select the tile used to display the data and the periodic table will update to reflect the state of the markers chosen for that spectrum.

Data can be quantified at a number of levels. The most basic form of quantification is using regions applied to a wide scan spectrum; the results are added to the display using an annotation option. Alternatively, both regions and synthetic components can be prepared over a set of narrow scan spectra and the results specified through a set of expressions involving the names used to label the objects. A report generated in this way can be viewed within CasaXPS or exported as a TAB separated ASCII file for use with a spreadsheet program such as Excel. Figure 24 shows the main toolbar button for the dialog window used to perform these two extremes plus a number of reporting mechanisms in between.

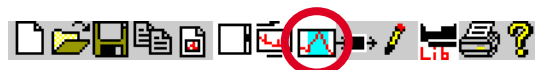


Figure 24: Main Toolbar.

The three property pages used to quantify spectra are labelled "Regions", "Components" and "Report Spec".

Regions

The first step in quantifying a spectrum is defining one or more energy regions. These specify the data channels that are to be included in the intensity calculation.

Each region requires the following information:-

- **Name** The name given to a region is significant as a quantification object. A report is specified in terms of these names and intensities are summed for regions with identical names.
- **R.S.F.** Relative Sensitivity Factor for the data within a region. The integrated data are divided by this value.
- **Start** Start point for the integration interval.
- **End** End point for the integration interval.
- **Background** Tougaard, Shirley, Linear, None
- **Average Width** Number of channels used to fix end points for the backgrounds.

Armed with the values described in the above table, a region is created by either pressing the "Create" button on the "Region" property page or by clicking the column buttons on the scrolled list of regions.

The start and end points for a region are taken from the zoom parameters currently in effect when the region is created. Before pressing a create button, zoom into the spectral feature until the display shows the range which characterises the intensity in question. The new region created at this stage will have start and end energies defined by the display.

After creation, the limits can be adjusted under mouse control. To perform this operation it is necessary that the "Quantification" dialog window is visible and the active-page corresponds to the "Region" property page. Position the cursor near an end point for a region, then hold down the left mouse button. A box will appear showing the extent of the region. The end under the cursor is altered when the mouse is moved whilst continuing to hold down the mouse button. Should the same procedure be followed, but instead of initially positioning the cursor towards an end point, the mouse button is pressed when the cursor is towards the middle; the result is both end points adjusted simultaneously. A box outlining the region is displayed and moving the mouse causes the box to shift both end points for the region.

Backgrounds to spectra can be selected from Tougaard, Shirley, Linear or None. To specify a new background type, it is sufficient to type the first character for the background name. For example, to change from a Tougaard to a Shirley, replace the name "Tougaard" by "S" then press return. The regions will be recomputed with the new background type. The initial and final values for the background are tied to the

end points of the region. Noise in the data often makes the value at the end point unrepresentative of the desired background limits. To reduce this problem, a number of channels may be used to define the background limits. The "Average Width" parameter defines the number of data channels to be averaged when calculating the tie-in points for the background.

The backgrounds are calculated using algorithms based on those presented in a number of published articles. The "Linear" case is simply a straight line between the end points of the region. "None" is just a constant set to the lowest data channel in the region. The more complicated backgrounds are those due to Shirley and Tougaard.

The procedure due to Shirley is essentially a weighted-average of the background limits chosen to tie in with the spectrum at the end points of the region. The weighting is determined from the area between the background and the data. Since the weighting is determined using the quantity being computed, a sequence of iterations are required to arrive at the desired result.

Tougaard has extensively studied the subject of backgrounds to XPS spectra, however the background generated by CasaXPS is simply calculated using the Universal loss function. For this algorithm to work it is necessary to remove all instrumental contributions from the spectral shape before calculating the background. This is seldom possible for practical situations. To allow the procedure to provide a background under less than ideal conditions, an adaptive procedure has been adopted that attempts to fit the background to the given spectrum. The result is a background that "looks" plausible in situations when the data has not re-

ceived the necessary pre-processing and equal to the background proposed by Tougaard for practical applications when the appropriate adjustments have been made to the data.

To delete a region, first select the region in the scrolled list of regions. Then press the "Delete" button.

Each region defined for a spectrum has a number of statistics gathered from the data and displayed in the scrolled list on the property page in question. The raw intensity (CPS eV) between the background and the data, an estimate for the F.W.H.M. and the position of the maximum count rate recorded are all listed below the parameters that define the region. In addition, a percentage concentration is shown at the bottom of each column. The raw intensity is calculated directly from the data, however, the percentage concentration includes R.S.F. adjustments. The results of quantification based upon a wide scan and a set of regions are therefore available through the "Regions" property page.

Synthetic Components

The property page labelled "Components" manages the creation and optimisation of synthetic components.

Synthetic components are specified by name, line shape, R.S.F., position, F.W.H.M. and area (CPS eV).

The name is the means of identifying a component. A quantification report specified using expressions makes use of the name to define how the intensities are to be combined. Further, components with the same name are summed together before an expression is evaluated.

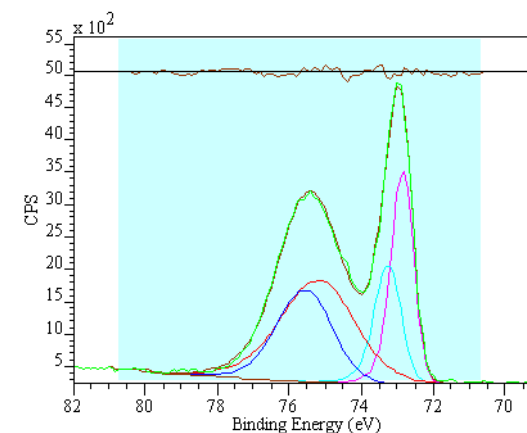


Figure 25: Region plus Synthetic Components

Line shapes and R.S.F. values are stored in the element library and these values are used when a new component is created. A new component is created using the buttons above the list on the "Components" property page. If the element library dialog is active and a transition is selected from the "Element Table" property page, then when a "create" button is pressed the values for the synthetic component are read from the transition so specified. If the element library page is not active then the ISO species/transition label is used to pick out the parameters for the component from the element library.

Line shapes that are available at present are either Gaussian-Lorentzian product or sum functions augmented by an asymmetric tail shape (for more detail, see "Line Shapes and Backgrounds" on page 55). These are specified as follows, $GL(n)$ where $0 \leq n \leq 100$ for a product function or $SGL(n)$ for a sum function. A value of " $n = 0$ " provides a pure Gaus-

sian while "n = 100" results in a pure Lorentzian. A tail is introduced by appending T(x) where $x > 0$ and is typically between 0.5 and 10.0. For example, an asymmetric line-shape for Al 2p might be "GL(30)T(2.0)", see Figure 25 for an example of such a component.

Clicking on the corresponding item in the scrolled list and entering new values can specify peak positions, widths and areas. Alternatively, when the "Components" property page is on top of the dialog window, the mouse can be used to pick up a component and move it to a new position, area or width. If the cursor is located near the top of a peak then the position and area are adjusted. If the cursor is located near the side of the peak then the width will change. Note that the height also changes when the width is adjusted using the mouse. These two parameters are modified simultaneously so as to preserve relationships between peak intensities.

Constraints may be specified for position, width and area. These may be either an interval (for example "0.5,1.3") or a constraint that is measured with respect to another component. In the latter case, the parameter for one peak can be fixed with respect to another by entering the column label into the constraint box. For example, to fix the position of a peak in column "B" with respect to a peak in column "A", type "A" for the position constraint in column "B". The system responds by setting the constraint to be "A + x", where x is equal to the initial offset between the two peaks. The value for x can be modified by hand. To provide an offset of 2 eV, column "B" constraint item should read "A + 2". Area may be constrained similarly, but instead of plus use times, e.g. "A * 0.666". Read, column B component is 0.666 times the area of column A.

Quantification Calculation

Percentage atomic concentration (X_A) calculations are performed using the formula

$$X_A = \frac{(I_A E^\alpha)/(R_A T(E))}{\sum (I_i E^\alpha)/(R_i T(E))}$$

Where R_i is the relative sensitivity factor for the measured intensity I_i and $T(E)$ is the transmission function for the instrumental operating mode used to measure the intensity I_i at kinetic energy E. The term involving raising the kinetic energy E to a power α allows traditional adjustments for analyzer behavior for those who wish to quantify spectra according to a manufacturer's specification. The value for the exponent a is entered on the Regions Property Page of the Quantification Parameters dialog window and is used whenever the accompanying check box is ticked. If the check box is ticked then both the exponent and the transmission function located in the VAMAS block will be automatically used during quantification, however a transmission function must be present for either of these corrections to be included in the calculation. Note that intensity calibration for a spectrum can be performed using the "Intensity Calib" processing option where an exponential adjustment to the data can be made regardless of the presence or absence of a transmission function

Quantification Report

There are several types of quantification report. The property page labelled "Report Spec" allows peak intensities to be

compared to one another either via information derived from regions, components or both regions and components. In addition, a set of named formulae allows intensities to be mixed and matched in any way that an arithmetic expression can define.

A report generated from regions and/or components is presented in the form of a scrolled list. For each experimental variable from the ISO file, the set of quantification items is listed. Information including peak position, F.W.H.M., R.S.F. and raw-area appear in separate columns. The final column is the percentage concentration for each set of quantified items with the same experimental variable.

When the report is based upon regions and/or components, each item is included in the report on a separate line even if a region has the same name as another region or component. This allows the statistics associated with each quantification item to be included in the list. Items with the same name are treated in a different way when formulae are used to generate the report.

If a report is based on regions and/or components, and not all the regions are to be included in the calculation for the percentage concentration, then an R.S.F. of zero should be entered for those items that are to be excluded from the results.

The second type of report is generated using the table of named formulae (see, for example, "Adjusting a Custom Quantification Report" on page 120). Each formula takes the form of an arithmetic expression. Variables in these expressions may include names given to the regions and components, such as the data blocks which are currently selected

in the Browser view of the active data file. The list entitled "Quantification item names" shows the set of names currently available for generating a report.

Clicking on the column header buttons labelled "Name" or "Formula" creates a new name/formula pair. The entry is adjusted via a dialog window. Right click over the name item for which changes are required. The dialog window enables the name and/or the formula to be modified or deleted.

Each entry in the "Custom Report" corresponds to the experimental variable in the active data file. The columns are entered two per named formula, one for the raw intensity calculated from the expression and one for the percentage concentration. Thus a table as follows (Table 1) would result in a report with five columns. One for the experimental variable and two pairs headed "Oxygen CPS eV", "Oxygen %" and "Aluminium CPS eV", "Aluminium %".

Table 1:

Oxygen	O 1s region
Aluminium	Al 2p + Al 2p Oxide

The variable names derive from an Oxygen region "O 1s region", and three synthetic components, two "Al 2p" (from a doublet of aluminium metal peaks both named "Al 2p") and one "Al 2p Oxide". The "Al 2p" name was repeated so that the intensities calculated for both peaks for the metal doublet were implicitly summed before the explicit sum defined via a formula computes the total contribution from the aluminium.

Optimisation of the Peak Parameters

Fitting synthetic components to a data envelope is probably the most useful tool in XPS data reduction. It is also one of the more difficult procedures to perform owing to the complex nature of the underlying peak structure and the question of how to account for the background structures in the data. The line-shapes used to characterise a peak are only approximations and there can be no doubt about the uncertainty associated with the different background algorithms. It is therefore entirely possible to fit a set of synthetic components with good statistics, yet without any chemical or physical meaning.

To provide a guide to peak fitting, an option on the processing dialog window labelled "Test" can be used to replace the true data using one of a set of known peak structures. These structures derive from work presented by Seah and Brown. The relative intensities and separations are those presented in Table 4 in the Seah and Brown publication⁴; a GL(50) line shape has been used to generate the components. It is useful to exercise the fitting procedure using these structures with different options for the synthetic components, especially when a background has been added.

The following table provides the characteristics of the artificial peak structure. Note that the data replaced determines step size, energy position and count rate.

4. M.P. Seah and M.T. Brown, J. Elec. Spec., **95** (1998) 71-93

Table 2: Artificial Peak Structures.

Carbon C 1s peaks are used to form data envelopes with the offsets and relative sizes list below

	PMMA	PVA	PVC	PIB
Number of peaks	4	4	2	3
peak offset (eV)	2	2	2	2
factor	1.238	1.04167	1.0	2.0
peak offset (eV)	3	3		3
factor	1.238	1.04167		1.0
peak offset (eV)	4	4		
factor	2.4762	1.08333		

It is very useful to try to fit the artificial peaks without any constraints with respect to the other peaks and then gradually include the relationships shown in the above table. The chi square value for the curve fit should go to zero. This is the case since the artificial data does not contain noise. Ordinarily, in the presence of noise, the chi square value should be about equal to the number of degrees of freedom (also shown on the "Components" property page).

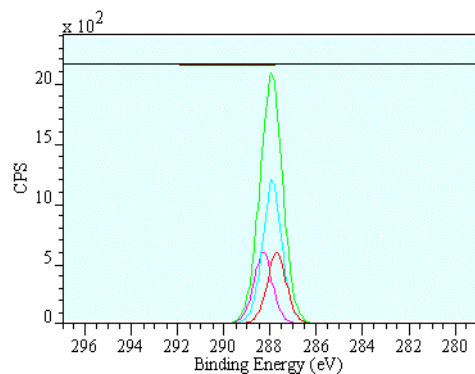


Figure 26: PIB artificial peak structure.

A good exercise is to attempt to reproduce the peak envelope for PIB. What should be observed is how difficult it is to produce a good fit for PIB. The three peaks are very hard to identify unless additional information is given to the fitting procedures. If the offsets are first provided then the fit improves, but the exact match does not materialise until both the correct intensity ratios are supplied as well as the relative positions.

Another useful feature of fitting this artificial data is that the strengths of the two fitting algorithms on offer can be seen. PIB is a stern test for the Marquardt method. The uncertainty as to which of the many combinations of similar peak parameters presents a plateau in the parameter space that fails to give any good direction towards the optimum values. On the other hand, the Simplex algorithm, once close and sufficiently constrained, will march straight to the exact fit. The Marquardt method works best when it has a clear view of the target such as with PVC (Figure 27).

The option for switching between the two-optimisation methods is on the "Components" property page.

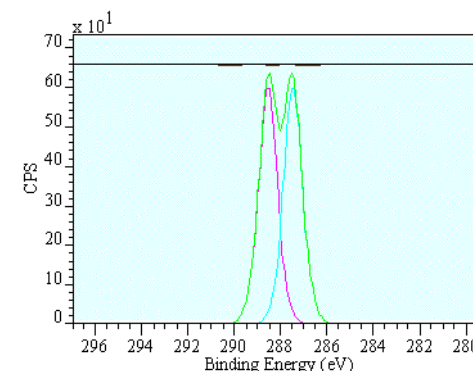


Figure 27: PVC artificial structure.

The optimisation procedure labelled "Marquardt" is actually not a pure Levenberg-Marquardt method. Linear and non-linear parameters are separated. Each step of the optimisation procedure includes the solution of sub-problems introduced by the presence of constraints. The Marquardt method is used to establish the next set of non-linear parameters for the current set of optimal linear ones. Although no optimisation procedure should be viewed as universally applicable, these procedures provide a robust method for determining the peak parameters.

The Marquardt method uses information about both the function and its derivative. There are situations where the informed steps produced by this method are slower than the "lets guess here" approach of the Simplex method. In fact, once the Marquardt method stops making significant improvements in the chi square value there is no harm in

switching to the Simplex method for one last try. This can sometimes push the parameters away from a local minimum that has trapped the Marquardt method or on occasion serve to confirm the optimum has been found.

Marquardt method with constraints can help with the progress towards physically significant peak parameters. However, if the constraints prevent the synthetic peaks reaching an obvious optimum without an alternative available then the algorithm tends to be slow. That is, if no natural optimum lies within the range of the constrained parameters and yet an optimum is visible outside the parameter range, then the algorithm will labour. The problem lies in the numerous probes outside the range that will cause backtracking to the boundary values. Understanding this fact is useful. A well-posed problem will tend to converge quickly, while an inadequate model may result in sluggish behaviour.

The Simplex method on the other hand works based upon maintaining a set of function values at the vertices of an N-dimensional simplex. One of a set of prescribed transformations for the simplex is employed depending on what function value is founded at a probe point.

As it is prescribed, there is very little that can go wrong with the Simplex algorithm. However, convergence to the optimum set of N-parameters is not guaranteed especially when a non-smooth function is used (i.e. when constraints are introduced). It does seem to work well for the situations typically found by CasaXPS, although it has been pointed out that for some optimisation problems, the Simplex method fails to converge when the number of parameters exceeds a problem-dependent value (even for smooth functions). The conclusion is, don't try to fit too many peaks using the Sim-

plex method unless the parameters are already close to the optimal values.

Propagating Quantification

Once a representative spectrum has been modelled using regions and components, the next question is how to transfer this model to similar spectra within the same ISO 14976 file and other files. CasaXPS provides a propagation mechanism to accomplish this task.

The spectrum for which the component model has been constructed must be displayed in the active tile view. Any ISO blocks that require the equivalent model must be selected via the Browser View for each of the files involved. Once these conditions are met, the mouse cursor is placed over the active tile and the right mouse button pressed. A dialog window appears for propagating the quantification objects.

Check boxes are offered on the "Propagate" dialog window. These allow a choice of what actions are propagated through the selected ISO 14976 blocks. To transfer the quantification model the "Regions" and the "Components" check box should be ticked. Then press the "OK" button. A progress dialog window appears that allows the propagation process to be terminated (see "Processing" on page 34).

Data Editor

Curve fitting assumes a statistical model for the noise recorded with the data. This model can sometimes be invalidated by the presence of unrepresentative spikes that are due to the detector system and have nothing to do with the true electron yield. It is best to remove such artefacts before attempting to optimise the parameters for the synthetic com-

ponents.

A data editor is provided on the quantification dialog window. The abscissa and ordinates for the data displayed in the active tile are listed on a property page labelled "Data Editor". The value for an ordinate can be changed. By right-clicking the mouse whilst the cursor is over the corresponding abscissa a dialog window is brought up that allows the ordinate to be edited.

A check box allows the user to specify that an ordinate value should be permanently altered. That is to say, if the data is written back to disk then the adjusted value will be used in the file. It should be noted that some actions in the processing window can cause edits made to the data to be "undone". This occurs if the check box is not ticked and the processing history is used to change the state of the processing (either by resetting or applying a selection of processing to the data). The history mechanism always refers to the unprocessed data before taking the requested action.

Derivatives and Peak Envelopes

A number of authors have proposed methods for identifying the underlying peaks responsible for a measured spectrum. Some of these ideas can be tried out on the known peak structures for PMMA, PVA, PVC and PIB.

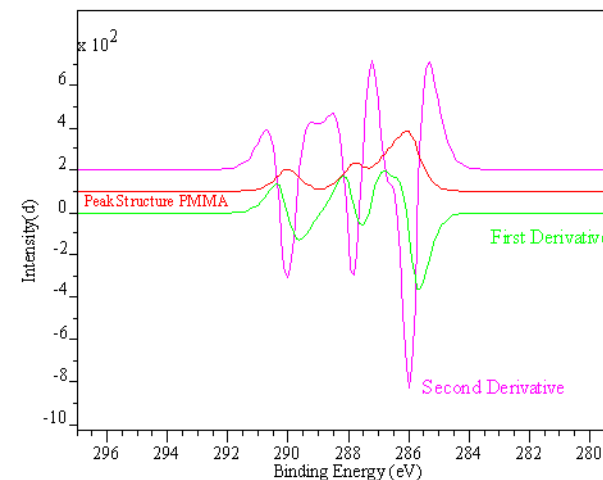


Figure 28: PMMA peak structure with first and second derivatives.

PMMA has four peaks, two well resolved and two that merge together to form a broad structure towards the lower binding energies. Spotting the shoulder associated with the third and fourth peaks and determining the locations for these poorly resolved peaks is critical for constructing a physically meaningful model. The presence of a background also interferes with identifying the true structure, although the data in Figure 27 has been prepared without this complication.

Differentiating the artificial data once and then twice can help to see the underlying peak structure (see Figure 28). The shoulder of the third peak in the PMMA envelope can be seen in the second derivative where a kink in both derivatives highlights the shoulder in the data.

Quantification by Example

The procedure for quantifying a set of spectra will be explained using an ISO 14976 file. This file includes measurements for four transitions taken after etch cycles have been used to change the state of the sample. The problem is to create suitable models for each of the transitions using a combination of regions and components, then to define the way these quantities should be combined in the form of a report. The multiple-document-interface (MDI) used by CasaXPS requires a new document before a file can be selected. When CasaXPS is first run, the initial state is an empty document and the "Open" menu option on the "File" menu is in an active state. If all the documents have been used or none are on offer then the "Open" option as well as the corresponding toolbar button will be inactive. To activate the "Open" option, press the "New" menu button; an empty ISO document frame will be created.



Figure 29: Main Toolbar

A file dialog window enables browsing of the file system and choice of an ISO file. Take care within the file dialog. It is the standard MFC dialog with drag and drop functionality. It is therefore possible to move files around simply by dragging the icon for the file over the icon for a folder.

Once read, the ISO file structure is displayed in the right-hand side of the Document view. Each transition appears as a column and each row is populated by spectra acquired with

the same etch time.

Table 3: Structure of the Browser view for a depth profile.

Etch time (s)	C 1s	O 1s	Al 2p	Mg 2p
0	C 1s/1	O 1s/2	Al 2p/3	Mg 2p/4
100	C 1s/6	O 1s/7	Al 2p/8	Mg 2p/9
200	C 1s/11	O 1s/12	Al 2p/13	Mg 2p/14
300
.....
2000	C 1s/101	O 1s/102	Al 2p/103	Mg 2p/104

The first step towards creating a model for each of these transitions is to identify a representative spectrum for each column of the browser. Click the column header for "C 1s". All the carbon spectra then become the browser selection. Pressing the toolbar button for displaying the blocks one per tile; the spectra are displayed in the left-hand scrolled view.



Figure 30: Spectrum Display Toolbar.

Choose a carbon spectrum that shows a well formed C 1s envelope. The limits for the energy range and hence the background to the spectrum needs to be well defined. A good idea of where the carbon peaks are located is essential for establishing a realistic background shape.

Before proceeding further ensure that your element library is loaded. (See "Element Library" on page 38).

Another important consideration is intensity calibration. Before quantifying the raw data it may be necessary to apply a transmission adjustment. See "Processing" on page 34 for a discussion of the issues involved with intensity calibration.

The next decision determines what values are used in the region. In the event that only the total counts for carbon is required, no synthetic peaks are necessary and the name plus the R.S.F. values will be the key information used by the quantification report. On the other hand, if the chemical states within the carbon envelope are to be profiled, then the R.S.F. within the region can be set to zero. The name for the region should be chosen to be different from any that will be used to identify the components.

Let us say the only chemical state information required is that related to the "Al 2p" transitions. For all other transitions, only the total intensity is needed to construct the profile. So for "O 1s" and "Mg 2p" repeat the following steps described for "C 1s":

1. Select all the spectra for "C 1s" by clicking the column header.
2. Enter the spectra into the scrolled list by pressing the button that displays one spectrum for each of the tiles (Figure 30).
3. Scroll to the spectrum that best illustrates the character of the data and click in the tile with the left-hand button of the mouse. This makes that tile the active tile.
4. Bring up the dialog window for quantifying the data (Figure 24).
5. Add a region via the "Region" property page and adjust

the parameters until the background and the integration limits look right (whatever that may mean).

6. Ensure that only the set of carbon transitions is still selected in the Browser view of the ISO file. Then right-click with the mouse cursor over the active display tile. The "Propagate Actions" dialog window will appear.
7. Select the check box for "Regions" and press "OK". The result is that each spectrum in the set of carbon spectra now has a region defined to be the same as the spectrum displayed in the active tile.
8. Scroll through the carbon spectra and check that sensible regions have been defined and that the backgrounds are indeed representative of the data. Repeat steps 6 through 8 until satisfactory results are obtained.

Once the above sequence of steps have been performed for "C s", "O 1s" and "Mg 2p" it is then time to generate synthetic components for the "Al 2p" data. Figure 25 shows a representative spectrum taken from a sample for which aluminium metal and aluminium oxides were present.

To create a model for an aluminium envelope, the first step is to create a region. In this case the region is best named "Re: Al 2p" and a value of zero should be entered for the R.S.F. value. By entering zero for the R.S.F. value, any quantification report will not include the intensity calculated from the region in the percentage concentration values.

Two pairs of components named "Al 2p Metal" and "Al 2p Oxide" describes the data seen in Figure 25 These correspond to two sets of doublets for aluminium. This knowledge can be introduced into the model via the constraints for the component parameters. The relative intensity and spacing of the components of a spin-orbit split (doublet) are usu-

ally well known, and inclusion of this information results in a chemically more meaningful model and also one in which the optimisation routines will work more effectively.

Armed with the component model for one spectrum from the "Al 2p" set, the other members of the set can be given similar models automatically. Ensure that the "Al 2p" set is selected in the Browser view then, in the tile view, right-click on the active "Al 2p" spectrum. Choose both the "Region" and "Component" check boxes, then press the "OK" button. Each spectrum within the set is fitted with the same region and components as the model which was prepared earlier.

When components are propagated the parameters are automatically fitted to the target data. If there is a trend in the data, for example, the oxide components dominate at the surface but smoothly diminish with depth, then the component parameters for the previous spectrum in the set of "Al 2p" narrow scans are a better starting point for the optimisation process. The "Propagate Action" dialog window includes a check box labelled "From previous block". If the check box is ticked, the parameters for successive blocks are taken from the preceding block. Otherwise the parameters derive from the block displayed in the active tile.

Again, it is advisable to view the results of these automatic fits by scrolling through the spectra in the tile view. Once satisfied, the ISO 14976 file is now ready to generate a quantification report.

The first step is to ensure that all the spectra to be included in the report should be selected within the Browser view.

A full report detailing parameters such as position, F.W.H.M., R.S.F., raw-area and percentage concentrations for all the quantification units is obtained by pressing the

"Combined" button on the "Report Spec." page. Regions and components can be reported in a similar format by pressing the buttons on the same property page but baring the appropriate names.

A custom report for the same data is usually in a format more suited to showing trends. To generate such a report it is necessary to complete a table of names and formulae.

The first column for a custom report is a list of the experimental variable values. In this example the column will contain etch time in seconds; the units are defined by the VAMAS file format. Subsequent columns are defined by the formulae, and are labelled by the corresponding names.

Pressing the "Reset" button initialises the list of names/formulae (Table 4). The entries are taken from the regions defined for the spectra currently selected within the Browser view. Typically this provides the right number of quantification names, but additional ones may be created or existing ones removed via the edit dialog window. Right-click the mouse button over a name field to display the edit dialog window.

Table 4: Name/Formula List.

Name	Formula
C 1s	C 1s
O 1s	O 1s
Re: Al 2p	Re: Al 2p
Mg 2p	Mg 2p

In the current example, three of the entries are appropriate for the custom report; only the "Re: Al 2p" region needs to be edited. Move the cursor over the name field and then press the right-hand mouse button. The dialog window for editing the name/formula field appears and the current values for these two items are entered into the text-edit fields. Change the name from "Re: Al 2p" to "Al 2p", then edit the formula to read "Al 2p Metal + Al 2p Oxide". Note that the list above the name/formula table contains all the names defined for the report. On the edit dialog window, press the button labelled "OK"; observe that the name/formula for aluminium has changed to the desired values (Table 5).

Table 5: Name/Formula List after editing.

Name	Formula
C 1s	C 1s
O 1s	O 1s
Al 2p	Al 2p Metal + Al 2p Oxide
Mg 2p	Mg 2p

The custom report is generated by pressing the "Apply" button on the "Report Spec" page. Nine columns of data are displayed in a scrolled list view. The first column is the experimental variable and is followed by eight columns corresponding to the names/formulae previously prepared. The first four of these eight columns list the raw areas that were specified via the formulae, and the second set of four col-

umns represents the percentage concentrations for the same items.

The values within this report may be written to file in an ASCII format. Each column of data is TAB separated. This allows the text file to be read into a spreadsheet program such as Excel, where the data can be formatted for printing either as a table or in graphical form.

Report Files and Excel

When a report is generated the data are displayed in a scrolled list. This represents an alternative view for a VAMAS file document. If a report view is the active MDI frame (indicated by the frame title bar colour), then the main menu only offers "File" and "Window" menus. The options on the "File" menu under these circumstances are restricted to "Save As ..."; this allows a File Dialog to be used to specify a text file to receive the report.

Once a report has been saved to disk, Excel can open the text file simply by selected the file name via the "File/Open" option on the Excel main menu. A Wizard for loading text-files into Excel guides the way through the available options. On completion the columns of data appear in a spreadsheet format.

Excel provides many tools for presenting spreadsheets in both tabulated forms as well as graphically. These far exceed the sophistication that could be implemented in CasaXPS and so the use of a spreadsheet for presenting the results seems most appropriate.

A range of physically possible line profiles in core-level XPS is possible, and simple Gaussian or Lorentzian functions are only rarely entirely adequate. In the case of metal samples it has been shown that asymmetric profiles should be expected on theoretical grounds [6], however recorded spectra exhibit deviations from idealized profiles due to a range of instrumental and physical effects:

- The response function of the electron analyzer (which may be asymmetric),
- The profile of the x-ray line shape (predicted to be asymmetric for non monochromatic lines from metal anodes [6]),
- Intrinsic life-time broadening of the core-level hole state (usually assumed to be Lorentzian in nature),
- Phonon broadening,
- Differential surface charging of the sample.

In addition to these instrumental considerations the shape of a synthetic peak is also influenced by the choice of background algorithm used to remove so called “extrinsic” elec-

trons from the data. Figure 31 and Figure 32 show examples

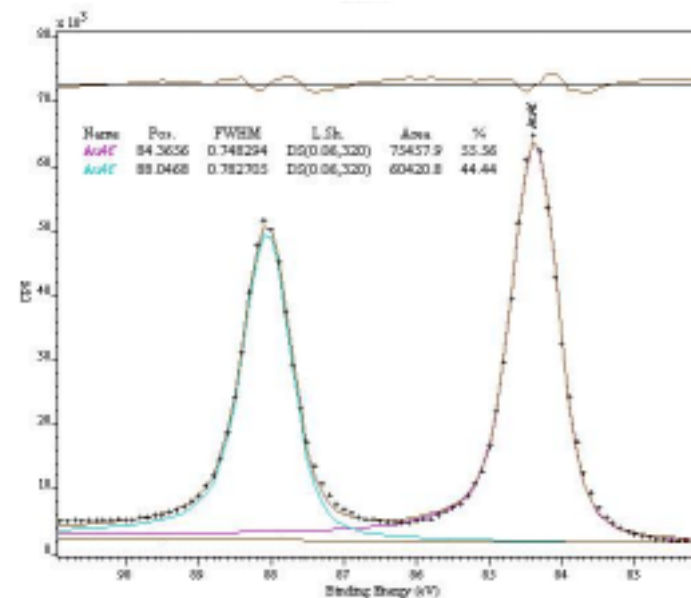


Figure 31: Au 4f doublet fitted using an offset linear background. The line-shape used in the fit is a Doniach Sunjic form convoluted with a Gaussian that results in an asymmetry index [6] of 0.49 for each peak

of peak fits where the difference between the two outcomes lies in the choice of background. Doniach Sunjic profiles have been used to model the recorded doublet, where it has become necessary to adjust the asymmetry parameter as well as the width of the Gaussian used to broaden the underlying Doniach Sunjic shape. The most obvious difference

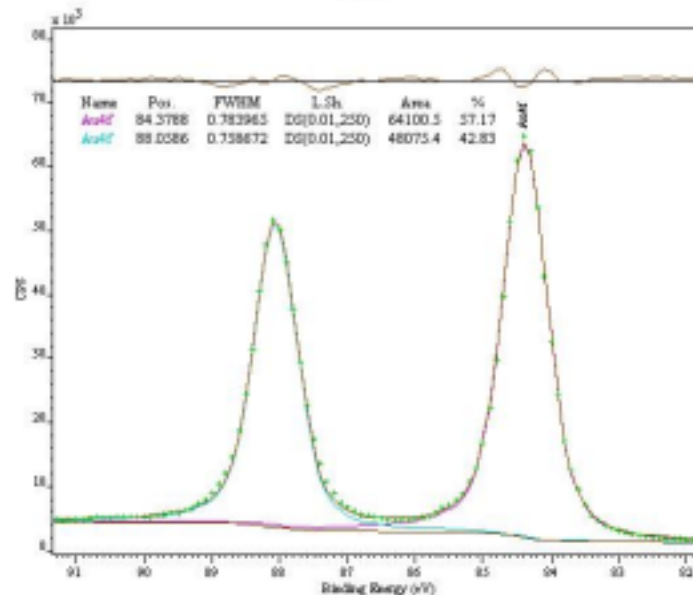


Figure 32: Au 4f doublet fitted using an iterated Shirley background. The line-shape used in the fit is a Doniach Sunjic form convoluted with a Gaussian that results in an asymmetry index of 0.008 for each peak.

between the two peak fits is the reduction in the amount of asymmetry required in the line-shape when a Shirley ^[5]

5. Shirley D. A., *Phys. Rev.*, **55**, 4709 (1972)

background is used. This is best measured by the asymmetry index ^[6], which changes from 0.49 in the case of a linear background to 0.008 when a Shirley background is employed.

Although theoretically based, the Doniach Sunjic profile suffers from ill-defined areas. The integral of the Doniach Sunjic function is infinite for non-zero values of the asymmetry parameter ^[7] and therefore any direct use of this line-shape requires a somewhat arbitrary use of cut-offs to allow finite peak areas to be reported. Changing the shape of the peaks through adjustments to the asymmetry parameter and the width of the broadening function has altered the relative proportion of the spin-orbit split peaks in the Au 4f doublet (Figure 31 and Figure 32). The background choice has altered the reported intensity as well as FWHM for the individual peaks. These factors emphasize the importance of comparing like-with-like when quantification results are used in practice.

The choice of line-shape is determined by the nature of the problem. The best fit is not always as important as producing quantification results that can be compared to historical results. On the other hand if the results represent a sequence of experiments for which a trend is more important than the absolute values then a good model may take precedence over the need to supply numerical values consistent with the past. Provided the models are self consistent the use of Doniach Sunjic line-shapes is usually acceptable.

Analysis driven by precise quantification lies at the root of

6. Doniach S. and Sunjic M., *J. Phys.* **4C31**, 285 (1970)

7. Evans S., *Surf. Interface Anal.*, **17**, 85 (1991)

many ad hoc procedures used in XPS. Shirley backgrounds were introduced precisely^[8] to remove as much asymmetry as possible from recorded data (Figure 32) in a well-prescribed fashion, so that near symmetric synthetic models can be used to characterize the intensity under a peak. A Gaussian/Lorentzian line-shape is finite and, with the appropriate relative sensitivity factors (RSF), can be used to compare intensities from fitted peaks to those calculated from integration regions. The key factor that has popularized Shirley backgrounds is the availability of well-characterized RSF values coupled with a specific background and line-shapes.

Some of the line-shapes introduced in CasaXPS have been constructed to allow Doniach Sunjic asymmetric behaviour to be associated with an underlying Gaussian/Lorentzian shape. The Voigt approximation is used to characterize the area, position and FWHM, while the asymmetric form approximates the rise in the signal much in the same way that the Shirley background is used to reduce the data to symmetric shape. The advantage of retaining a separate asymmetry parameter in the synthetic model is apparent when a Tougaard background is used to remove the extrinsic contribution to a metal spectrum.

To facilitate both trend analysis and basic quantification a wide range of line-shapes is required.

8. Castle J. E. *et al.*, *J. Electr. Spectr. Related Phenom*, **106**, 65 (2000)

A List of Line-Shapes

The line-shapes offered in CasaXPS are based around the following fundamental functional forms.

The **Voigt functional form** has been the basis for most quantitative analysis of XPS spectra and is simply a convolution of gaussian and lorentzian primitives.

Unfortunately this convolution proves to be mathematically intractable and an exactly accurate analytical form for a gaussian - lorentzian convolution is not available^[7], so practical systems have adopted two approximations to the true Voigt function, as described below.

Gaussian/Lorentzian Product Form

$$GL(x, F, E, m) = \frac{\exp\left(-4 \ln 2 (1-m) \frac{(x-E)^2}{F^2}\right)}{1 + 4m \frac{(x-E)^2}{F^2}}$$

Gaussian/Lorentzian Sum Form

$$SGL(x, F, E, m) = \exp\left(-4 \ln 2 \frac{(x-E)^2}{F^2}\right) (1-m) + \frac{m}{1 + 4 \frac{(x-E)^2}{F^2}}$$

Exponential Asymmetric Blend based upon Voigt-type line shapes

Given either of the above Gaussian/Lorentzian symmetric line-shapes, an asymmetric profile is obtained from a blend function as follows.

$$Y(x) = GL(x) + (1 - GL(x)) \times T(x)$$

Where

$$T(x, k, F, E) = \begin{cases} \exp\left(-k\frac{(x-E)}{F}\right) & x \leq E \\ 1 & \textit{otherwise} \end{cases}$$

Alternative Asymmetric Line-Shapes

An asymmetric line-shape due to Ulrik Gelius (Uppsala, Sweden) offers a class of profiles by modifying the Voigt function via an ad hoc adjustment. The profile is given by

$$A(x, a, b, F, E) =$$

$$\begin{cases} GL(x, F, E, m) + w(a, b)[AW(x, F, E, m) - G(F, E)] & x \leq E \\ GL(x, F, E, m) & x > E \end{cases}$$

Where

$$AW(x, a, F, E) = \exp\left[-\left(\frac{(x-E)}{F+a(x-E)}\right)^2\right]$$

and

$$w(a, b) = b\left[0.7 + \frac{0.3}{(a+0.01)}\right]$$

The parameters a and b determine the shape of the asymmetric portion of the curve.

Doniach Sunjic: a theoretically based asymmetric line shape

Doniach and Sunjic ^[6] performed an analysis for both photoemission and X-ray line-shapes, both of which result in an underlying profile given by the expression below. The formula includes an asymmetry parameter a that characterizes the asymmetry for a particular metal-like material. F is related to the FWHM and the position E is again related to but

not equal to the position of the maximum intensity for the line-shape. It is therefore difficult to relate optimization parameters determined from the Doniach Sunjic profile to similar quantities determined from Voigt-type line-shapes.

$$DS(x, \alpha, F, E) = \frac{\cos\left(\frac{\pi\alpha}{2} + (1 - \alpha)\tan^{-1}\left(\frac{x - E}{F}\right)\right)}{(F^2 + (x - E)^2)^{(1 - \alpha)/2}}$$

Nevertheless, the Doniach Sunjic profile offers an asymmetric shape that is particularly appropriate for non-monochromatic X-ray induced transitions; the profile is potentially present in both the photoemission process as well as the excitation source.

Line-Shapes Based upon Backgrounds

The above profiles are assumed to be entirely due to intrinsic electron energy variations, where a background subtraction algorithm is required before these line-shapes can be used to model the spectra. An alternative approach is to include the background shape as part of the model ^[8]. An analytical

form for the Shirley background can be determined for each of the line-shapes and the sum of these backgrounds plus line-shapes is used to approximate the variation in the spectrum. A simple constant background is all that is required for this procedure although other forms for the background are still an option.

Castle et al. ^[8] have developed a Shirley-type adjustment to a Voigt line-shape. The Shirley approximation is calculated from the current gaussian/lorentzian shape and a polynomial $b_0 + b_1(x - E)$ is used to scale the background in order to provide a fit to the observed spectra. The procedure yields a “Kappa” parameter (given by b_0) that characterizes the “intrinsic” step in the spectrum observed for a particular sample.

Line-Shapes Available in CasaXPS

Curve fitting in CasaXPS is performed via the Quantification Parameters dialog window where the synthetic line-shapes are defined from the scrolled-list on the Components

property page (Figure 33).

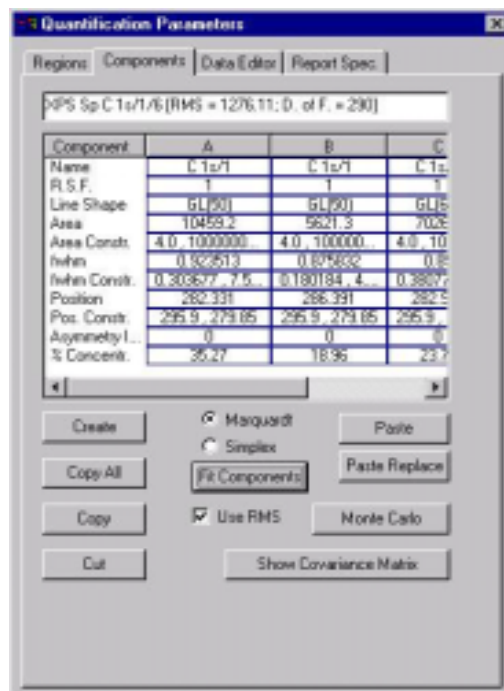


Figure 33: Property Page used to define synthetic line-shapes

The line-shape used to describe a photoelectric transition is entered in the row labelled “Line Shape” and takes the form of a text string.

GL(p): Gaussian/Lorentzian product formula where the mixing is determined by $m = p/100$, GL(100) is a pure Lorentzian while GL(0) is pure Gaussian.

SGL(p): Gaussian/Lorentzian sum formula where the mixing is determined by $m = p/100$, GL(100) is a pure Lorentzian while GL(0) is pure Gaussian.

GL(p)T(k): Gaussian/Lorentzian product formula modified by the exponential blend.

SGL(p)T(k): Gaussian/Lorentzian sum formula modified by the exponential blend

GL(p)K(b0,b1): Gaussian/Lorentzian product formula modified by a Shirley-type background prescribed by Castle *et al.* A linear polynomial determined from b0 and b1 adjusts the step in the Shirley background.

SGL(p)K(b0,b1): Gaussian/Lorentzian sum formula modified by a Shirley-type background prescribed by Castle *et al.* A linear polynomial determined from b0 and b1 adjusts the step in the Shirley background.

Further adjustments to the basic shapes

The basic shapes that result from the various functional forms can be further modified by numerically convoluting the profile with a Gaussian. For example, GL(100,50) will convolute a pure Lorentzian with a Gaussian characterized by a width of 50. The value for the width is the number of digital nodes used to describe the Gaussian and is therefore an arbitrary unit characteristic of the numerical form for the Gaussian.

N.B. Any line-shape that requires a digital convolution will cause slower performance than a shape that does not involve this procedure.

Asymmetric Line-Shapes

The exponential tail described by Sherwood in the book *Practical Surface Analysis* edited by Briggs and Seah attempts to approximate asymmetric line-shapes commonly found in photoelectric peaks. Several researchers have offered alternative deformations of the symmetric Voigt profile however none but Doniach Sunjic backs their approaches with a theoretical basis. Nevertheless, practical surface analysis requires practical solutions and the need for a rigid line-shape model that matches the observed profiles makes ad hoc forms acceptable tools. Philosophically, theoretically based solutions are preferred and hence, at least for metals, Doniach Sunjic profiles ought to offer “better” fits.

A practical solution offered by Ulrik Gelius (QSA 7)^[9] uses a Voigt-like function as the underlying shape and modifies the lower kinetic energy electrons using the equations given above. CasaXPS offers these line-shapes in the form:

A(a,b,n)GL(p): Gaussian/Lorentzian product formula modified by an asymmetric form define above. The parameters a and b allow the asymptotic form of the asymmetric tail to change whilst also altering the shape of the asymmetry prior to attaining the asymptotic shape. Figure 34 shows one such class of profiles where the parameter a has been held fixed at 0.35 while b varies between 0.2 and 1. The third parameter n defines the width of a Gaussian used to convo-

9. G. Wertheim, *J. Electron Spectrosc.* **6**, 239 (1975)

lute the basic shape of the profile.

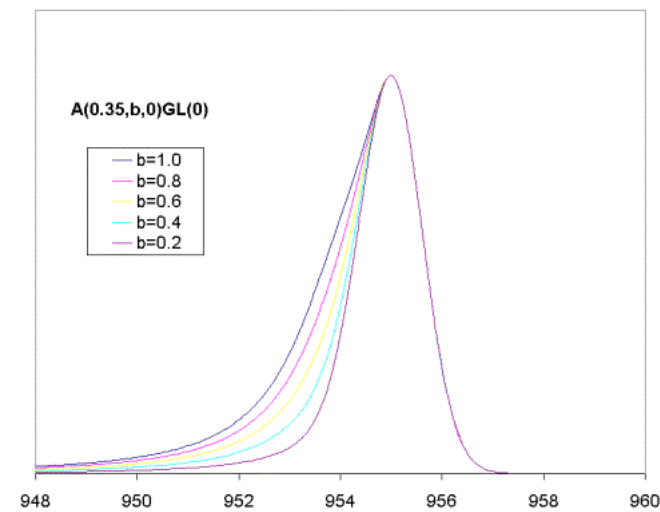


Figure 34: Asymmetric form due to Ulrik Gelius

The line-shapes defined by $A(a,b,n)GL(p)$ and $GL(p)K(b_0,b_1)$ are different from the Doniach Sunjic profile in that the Doniach-Sunjic profiles asymptotic limit is zero, while both of the ad hoc forms allow steps to be modelled within the line-shape itself. Another characteristic of these ad hoc line-shapes is that the peaks are defined in terms of intensity, position and FWHM taken from the underlying Voigt function. This is not unnatural in view of the fact that the area beneath all of these curves would be infinite without the introduction of some arbitrary cutoff limits.

DS(a,n): The basic Doniach Sunjic profile is defined in terms of the asymmetry parameter a and a convolution width

n. The profile defined above is numerically convoluted with a Gaussian whose width is determined from n to produce the final line-shape.

Although the Doniach Sunjic line-shape allows very good fits to experimental data, the infinite intensity defined by the curve makes this profile difficult to use under practical situations. Any scheme that limits the area through cutoff functions or ranges introduces inconsistencies into the peak parameters that make relationships such as stoichiometry invalid. The degree of asymmetry used to describe the data

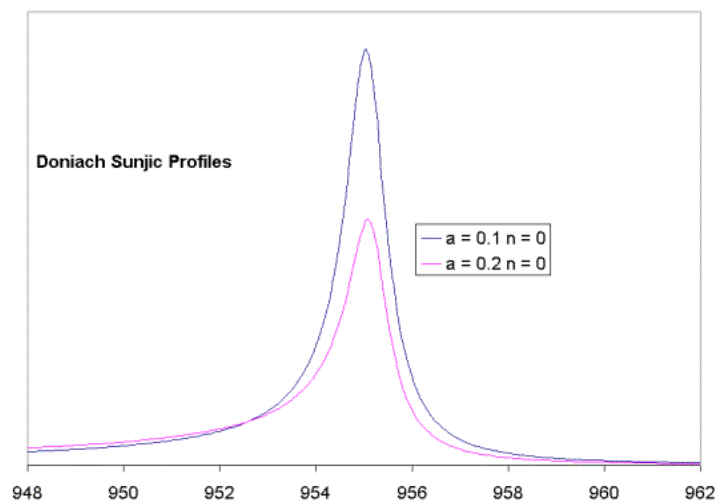


Figure 35: Doniach Sunjic profiles for asymmetry parameter equal to 0.1 and 0.2. Note how the asymptotic curves cross over illustrating how intensity is transferred away from the peak maximum

moves intensity modelled by the line-shape away from the primary peak position and towards the cutoff regions (Figure 35). For this reason it is essential to use the Doniach Sunjic line-shape in situations where the asymmetry parameters are close to the same value so that calculated areas can, in some sense, be compared.

Modifications to the Doniach Sunjic function

The advantages of the Doniach Sunjic profile are lost when synthetic models are required for quantification purposes. It is necessary to introduce ad hoc cutoff behaviour before finite intensities can be reported and, once obtained, the peak areas can only be used in a quantification reports when relative sensitivity factors are known for the specific cutoff criterion adopted.

Empirical observations suggest that pure Doniach Sunjic profiles do not always yield good fits for electrons with higher kinetic energy (KE) when monochromatic X-ray sources are employed. For example, Al 2p doublets recorded using a monochromatic X-ray source appears less Lorentzian in nature (to the higher KE side of the peak) than the Doniach Sunjic shape will allow (Figure 36), however the same line measured using an Aluminium anode does exhibit the predicted (but somewhat broader) shape. The latter is thought to be due to the double influence of the Doniach Sunjic shape from the photoemission process as well as the X-ray line profile, while monochromatic X-rays have lost the Doniach Sunjic distribution as a consequence of the fil-

tering process.

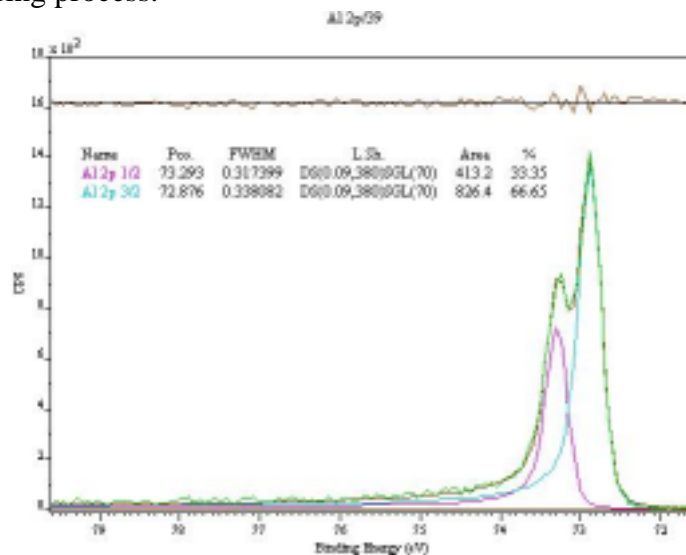


Figure 36: Al 2p Doublet from a monochromatic X-ray source.

In order to fit spectra such as that shown in Figure 36 the pure Doniach Sunjic shape has been modified. The lower KE asymmetry derives from the functional form for the Doniach Sunjic profile while the higher KE side of each component is a Voigt-type function. A numerical convolution is applied to the combined profiles to produce the line-shapes in Figure 36.

Given an approximation to the Doniach Sunjic profile of the form shown in Figure 36 then the possibility of characterizing the line-shape by the associated Voigt-type profile becomes feasible. The Voigt-type portion of the line-shape provides the position, width and intensity while an asymme-

try index ^[6] measures the departure of the line-shape from a symmetric form. Such a regime fits well with the philosophy adopted by Shirley, where the background shape was used to reduce an asymmetric step to a form that could be characterized by a Voigt function. The asymmetry index is given by

$$a = 1 - \frac{fwhm_{left}}{fwhm_{right}}$$

The asymmetry index is reported on the Components Property Page for each line-shape.

H(a,n)GL(p): Hybrid Doniach Sunjic/Gaussian-Lorentzian (product) line-shape.

H(a,n)SGL(p): Hybrid Doniach Sunjic/Gaussian-Lorentzian (sum) line-shape.

The “H” form of the DS/GL hybrid line-shape is identical to the DS(a,n)GL(p) form except the position, FWHM and area parameters are all determined from the Voigt-type shape. Compare the area values reported in Figure 37 to those from Figure 36. The area taken from the Voigt-type portion is significantly less than the intensity taken from the Doniach Sunjic curve where a cutoff has been used. Another difference between the two forms is that the area reported by the “H” is much less sensitive to adjustments in the asymmetry parameter (numerical convolution does move shape infor-

mation between the two halves of the line-shape).

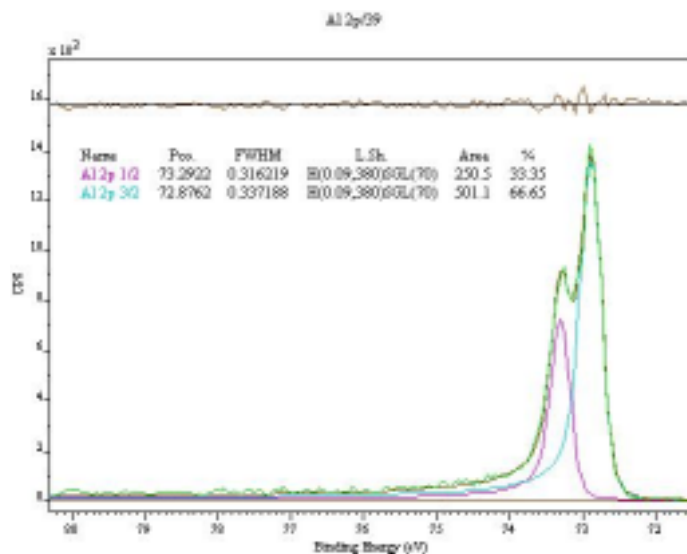


Figure 37: Hybrid DS/GL line-shape characterized by the Voigt-type portion.

A New Line-Shape

The hybrid form of a DS/GL line-shape is blended using a numerical convolution. The maximum for the two shapes diverge for larger asymmetry values, therefore characterizing the peak parameters by the Voigt-type portion is more difficult as the asymmetry becomes more pronounced. A new class of functions can be introduced to overcome this problem, namely, a set of line-shapes that are a blend of the asymptotic behaviour from the DS line-shape with a Lorentzian.

The cosine term in the Doniach Sunjic profile rapidly ap-

proximates a step function as the asymmetry parameter increases, while for small values of this parameter the cosine term is responsible for the movements in the DS maximum and for $a = 0$ results in a pure Lorentzian form. The formula for the “F” profile combines these characteristics by blending the asymptotic behaviour of the DS profile with one of the Voigt-type functions using a linear mapping.

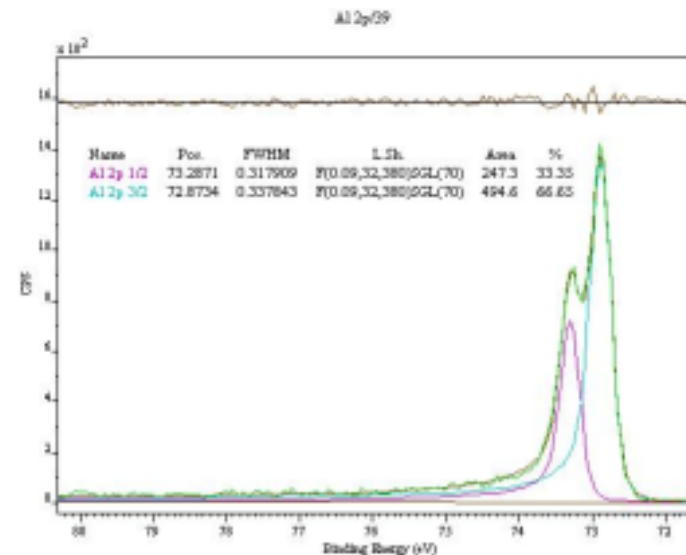


Figure 38: A new line-shape that blends the asymptotic behavior of a Doniach Sunjic profile with a Voigt-type function.

F(a,m,n)GL(p): Asymptotic Doniach Sunjic/Gaussian-Lorentzian (product) linear mapping line-shape.

F(a,m,n)SGL(p): Asymptotic Doniach Sunjic/Gaussian-Lorentzian (sum) linear mapping line-shape.

The linear mapping between the two functional forms is defined as a percentage through the m parameter. Figure 38 shows the same Al 2p data envelope as is seen in Figure 37, but the line-shapes are defined using the “F” functional form. In Figure 38, the same asymmetry value has been used as the one in Figure 37 however the linear mapping parameter is set to 32% DS asymptotic form 68% Voigt-type function. The resulting profile is convoluted with a relatively

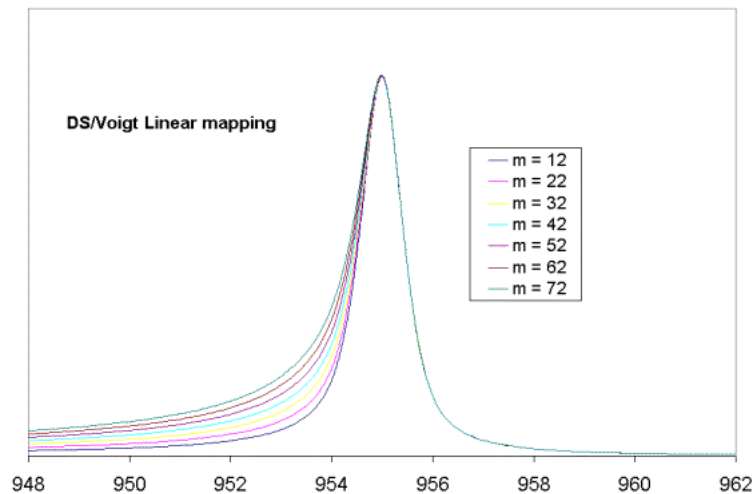


Figure 39: “F” function plotted over a range of values for the linear mapping parameter, a constant asymmetry parameter (0.09) and no numerical convolution applied.

wide Gaussian (characterized by the number 380) and accounts for the Lorentzian bias in the Voigt approximation.

Figure 39 shows how a pure Voigt-type function, SGL(70), is merged with the asymptotic form for a Doniach Sunjic

profile (asymmetry parameter equal to 0.09) for a range of the linear mapping parameter. The characteristic shape of the Doniach Sunjic profile can be retained to some degree but is supplemented by a class of shapes that go beyond what is possible with a DS profile. The principal difference between the DS profile and the “F” function lies in the static position of the maximum for the class of profiles. The moving maximum is a feature of the DS profile and can be achieved to some extent by use of the numerical convolution option. Increasing the width of the Gaussian caused the line-shape to lean in towards the asymmetric side of the peak (Figure 38).

A further variation on a theme is also possible where the “F” function can be replaced by:

$$E(x, \alpha, m, F, E) = m((Voigt(x))^{(1-\alpha)/2}) + (1-m)Voigt(x)$$

The “E” function is merely allows the shape to the right to be a part of the asymmetric shape to the left.

Background Subtraction

The line-shapes described above are very dependent on the availability of background subtraction algorithms that complement their use. If a background is incorrectly removed from a spectrum then the merit of attempting to model the result with theoretically correct synthetic line-shape is less clear.

Tougaard and co-workers [10] have devoted much time to understanding the backgrounds that are present in XPS spectra. The transport of electrons through a material after excitation by X-rays can be described by equations [11] that involve a one-sided convolution of the recorded data with an energy-loss probability distribution. This loss function of-

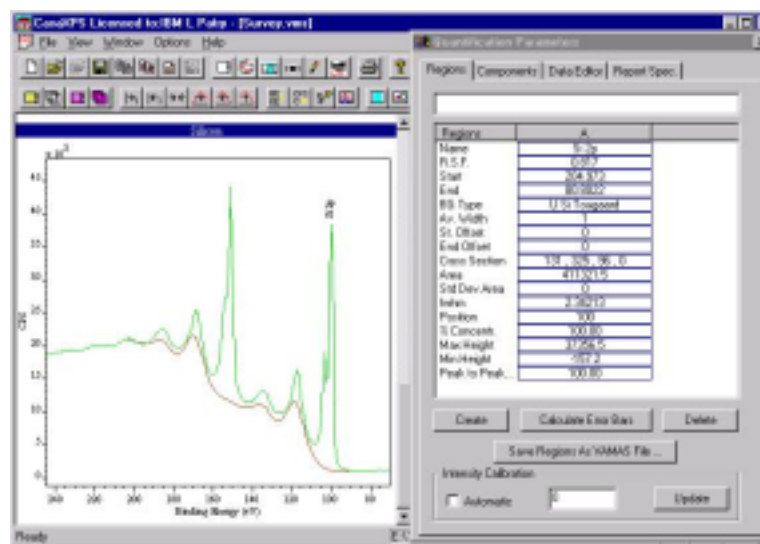


Figure 40: Three-parameter universal cross-section for Silicon

fers a prescription by which electrons leaving the sample with an initial energy may have their characteristic energy altered by the interactions with the surface. Each material

10. Tougaard S., *Surf. Interface Anal.*, **25** 137 (1997)

11. Tougaard S., *Surf. Interface Anal.*, **11** 453 (1988)

has a characteristic loss function of varying complexity and Tougaard has written many papers describing methods for establishing both generally applicable approximations (so called, universal cross-sections) and specific forms for individual materials.

The principal methods used by Tougaard for calculating the loss functions are theoretical dielectric response models and empirical estimates computed from spectra acquired using reflected energy-loss spectroscopy (REELS). Tougaard follows these calculations by fitting the results to a rational function, which models the resonance structure in the calculated loss-functions and allows the essential distribution to be described by a simple formula. These loss-functions are defined in terms of four parameters, three of which describe the shape of the rational function and one to allow for various band gaps in different materials. Tougaard refers to the distribution below as a “three-parameter universal cross-section” and has established values for a number of materials including aluminium, silicon, silicon dioxide and others [10]. The form of this universal cross-section is

$$P(T, B, C, D, T_0) = \begin{cases} \frac{BT}{(C-T)^2 + DT^2} & T > T_0 \\ 0 & T \leq T_0 \end{cases}$$

CasaXPS offers several ways of using the three-parameter universal cross-sections. The background type defined on the Regions Property Page of the Quantification Parameters dialog window (Figure 40) may be chosen from the following list:

U Si Tougaard (short form **U Si**) : Three-parameter cross-section determined for silicon.

U SiO2 Tougaard (short form **U SiO2**) : Three-parameter cross-section determined for silicon dioxide.

U Al Tougaard (short form **U Al**) : Three-parameter cross-section determined for aluminium.

U Ge Tougaard (short form **U Ge**) : Three-parameter cross-section determined for germanium.

U Poly Tougaard (short form **U Poly**) : Three-parameter cross-section determined for polymers.

U 4 Tougaard (short form **U 4**) : Custom three-parameter cross-section where the parameters are entered on the Regions Property Page (Figure 10) in the Cross Section text-field. The four numerical values entered in a comma separated list correspond to B,C, D and T0 in the energy-loss function above.

C 4 Tougaard (short form **C 4**) : Alternative custom three-parameter cross-section where the parameters are entered on the Regions Property Page (Figure 10) in the Cross Section text-field. The four numerical values have the same meaning as those found in the U 4 Tougaard form, however T0 is used to shift the resonance structure rather than simple acting as

a cutoff limit. It is therefore possible to retain the shape of the distribution but move the position of the maximum without changing the C and D parameters.

The custom Tougaard cross-sections offer a chance to change the background in response to the spectrum under investigation. Such an approach is not unprecedented ^[12]. “Direct, simultaneous determination of XPS background and inelastic differential cross sections using Tougaard’s algorithm” by Dr. M. Jo describes a method that extracts loss-functions from the XPS data using optimization procedures applied to a cross-section defined in terms of spline polynomials.

The rigid model offered by the three-parameter cross-section provides more guidance in the regions most influential in peak fitting, but will struggle to reproduce the asymptotic behavior for the cross-section as well as the adjustments needed to describe the near peak region of the background. Nevertheless, a practical approach to the use of Tougaard backgrounds in peak fitting is to choose a background that works in conjunction with the synthetic peaks. The custom form for the cross-section allows the background to develop at the same time as the peak parameters and choice of line-shapes. Figures 36 through 38 all use a custom three-parameter universal cross-section to describe the background under an Al 2p metal doublet; typically, researchers would choose Shirley background for such spectra.

12. Jo M., *Surface Science*, **320** 191 (1994)

Adjusting the universal cross-section

The rational function for the three-parameter cross-section is defined in terms of three constants B, C and D. This approximation is the quotient of a linear polynomial and a quadratic in terms of T^2 , but it is the denominator that determines the position and strength of the resonance in the cross-section. The parameters C and D both determine the position, width and height of the resonance, but to a first order (small D) the position is influenced by the square root of C and the sharpness of the cross-section is determined from the size of D.

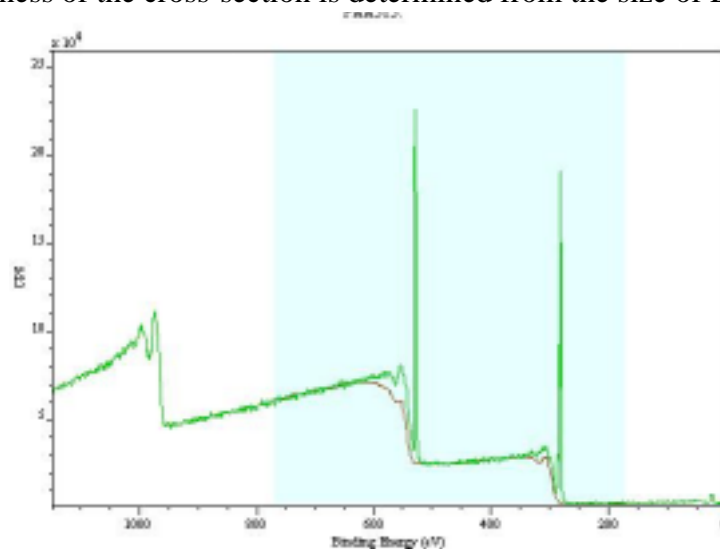


Figure 41: Wide scan spectrum from a PAA sample with a custom three-parameter cross-section applied.

The distance of a complex pole from the real-line in the domain of a complex function is instrumental in determining the width of a resonance when the function is plotted with

respect to a real variable. The D parameter principally moves the poles of the Tougaard cross-section away from the real-line, while C changes the projection of the pole onto the real-line. Small D corresponds to near singular behavior in the cross-section and therefore results in shape structures located, to some degree, by C while the magnitude of the function can be moderated by adjusting the B parameter. It should be noted, however, that B adjusts the behavior of the cross-section in the region close to zero energy-loss. This is precisely the region of interest for modeling a peak and so the values given to B and D can be influential in obtaining a reasonable synthetic model.

The three-parameter universal cross-sections for specific materials may be used as a basis for determining a background. A wide scan spectrum allows the essential form for a background to be determined in the case of materials for which cross-sections are not available. The resulting background is easily transferred to any high-resolution spectra from the same sample (Figure 41 and 42).

The meaning attributed to custom backgrounds should be viewed in the same way a synthetic model for a peak envelope is viewed. That is to say, only chemistry and knowledge of the sample allows a synthetic model to be constructed which has any meaning and this should also be the case for custom Tougaard backgrounds. The use of these backgrounds is entirely at the discretion of a researcher and bad backgrounds can be constructed in much the same way that synthetic models are open to abuse. However, a carefully constructed background can reveal features that may be overlooked or not accessible when a linear or Shirley form is applied. Figure 42 shows an O 1s structure from PAA (po-

ly(acrylic acid)) after a custom Tougaard background has been subtracted. The structures to the higher binding energy side of the peak may be real or an artifact of the background algorithm, but a linear background applied to the same data would fail to ask any questions at all regarding these features.

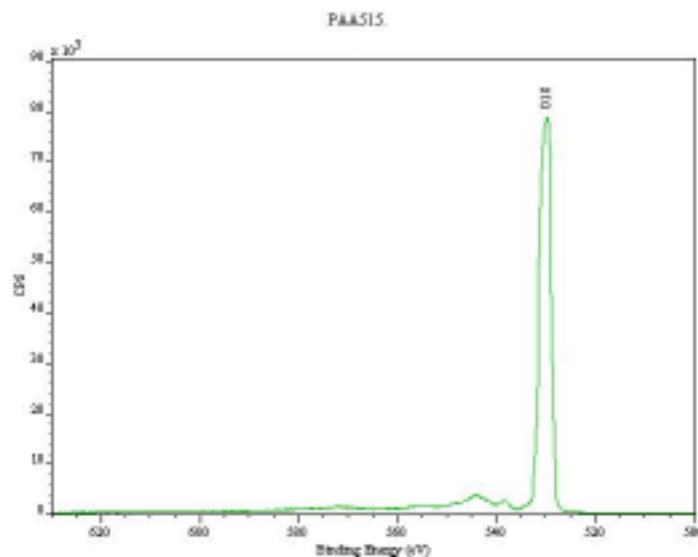


Figure 42: High-resolution spectrum from the same PAA sample shown in Figure 41. The custom Tougaard background has been subtracted from the data.

Simulating Spectra using Tougaard methods

Given a set of synthetic peaks and a three-parameter universal cross-section it is possible to construct a spectrum with

an associated background. The synthetic PAA spectrum in

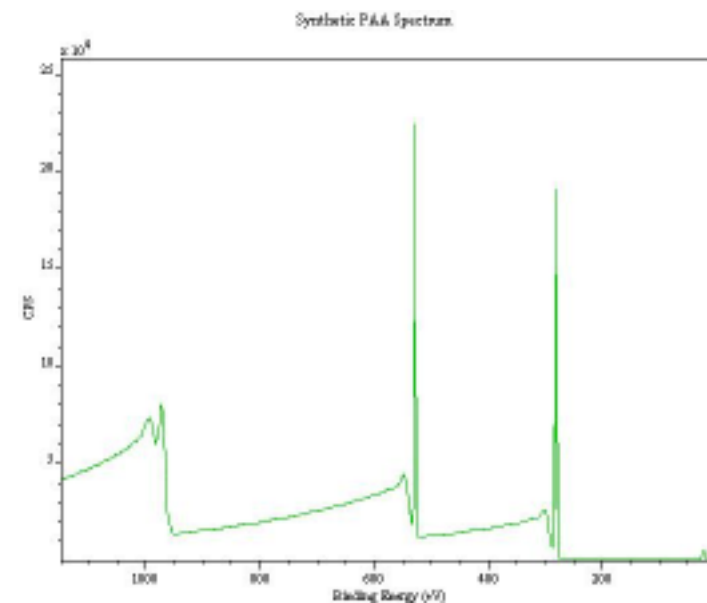


Figure 43: Synthetic PAA spectrum constructed from a set of Gaussian/Lorentzian line-shapes and a custom Tougaard cross-section.

Figure 43 has been constructed from a set of Gaussian/Lorentzian line-shapes approximately positioned at the experimental values for the valance band, two C 1s structures, two O AES peaks and one O 1s photoelectric peak. The widths and intensities are consistent with results for such a sample measured using a Scienta ESCA-300 at RUSTI, Daresbury Laboratory, UK, and the cross-section has been chosen similar to the one used in Figure 42. A sequence of iterations using these peaks and the three-parameter univer-

sal cross-section shown in Figure 44 converges to the synthetic spectrum in Figure 43.

This technique for generating theoretical spectra offers a way to understand the consequences of choosing a particular cross-section and if nothing else provides a basis from which real spectra can be viewed. Surface morphology and nanostructure causes most real spectra to exhibit background behavior that is not characteristic of a material but the environment in which the element is found. Attempting to construct a spectrum based purely on a resonant cross-section model clearly shows the limitations of such approximations when applied to real samples.

Although synthetic modeling has little direct use for real samples, it does offer an analyst the opportunity to investigate algorithms used to probe real spectra. For example, numerous techniques have been proposed from Multivariate Statistics for reducing XPS spectra to meaningful chemical information [13] [14]. Procedures for generating simulated spectra allow data sets to be constructed with similar characteristics to the experimental data but with well-defined information. These algorithms can be applied to the simulated spectra in order to establish what information is extracted for

13. Fiedor J.N., Proctor A., Houalla M. and Hercules D.M., *Surf. Interface Anal.*, **20** 1 (1993)
14. Do T., McIntyre N.S., Harshman R.A., Lundy M.E. and Splinter S.J., *Surf. Interface Anal.*, **27**, 618 (1999)

a perfectly understood data set.

Everyday procedures such as curve fitting can be evaluated using simulated spectra and therefore the ability to produce spectra with known shapes can enhance understanding for students and researchers alike.

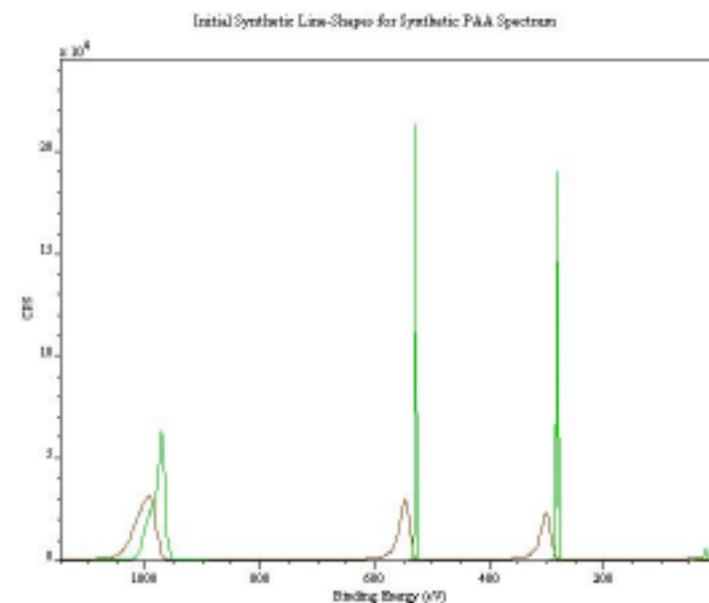


Figure 44: Initial synthetic components used to construct a PAA spectrum (green). The background curve (brown) represents the first iteration of the cross-section with the model data.

[

CasaXPS provides a “spectrum Calculator” which enables the normal arithmetic operations (addition, subtraction, multiplication and division) to take place between spectra or spectral regions. The Calculator is a means of managing and extracting information used to identify spectral features and then to take that information on to quantify regions and components. Calculator operations may be performed in a “one-to-one” or “one-to-many” fashion.

Creating a Comparison File

CasaXPS enables calculation operations on spectra **provided that the data blocks are located in a single ISO 14976 file**. It is clearly not always convenient for data acquired at different times to be stored in the same file, but spectra from one or more files may be easily copied into a single file in order to make such comparisons or calculations.

The procedure for moving the spectra is as follows:

1. Create an empty experiment frame window (use the “New” button or shortcut).

2. Load the separate files for which a comparison or calculator operation is required (each in its own experiment frame).
3. Select the ISO blocks of interest via the loaded experiment frame browser windows. To add to the selection the Control key must be pressed when new blocks within a file or from different files are selected.
4. Move the cursor over the right-hand-side of the empty experiment-frame window and click the right-hand mouse button. A dialog window will appear and offers the list of ISO blocks currently selected. Provided these data are compatible (in the ISO sense), the selected spectra will be copied into the previously empty window. If the tick box labeled “Data only” is ticked then the spectra will be copied using the results of any processing, otherwise an exact image of the original data will be transferred to the new file.

An empty frame represents an empty VAMAS file and by copying spectra into an empty frame one creates a new (uniform) experiment within which it is possible to compare the spectra, but in addition the transferred data may subsequent-

ly be saved to disk so that the results of any data manipulation performed on the spectra are preserved. For example, the difference of two spectra can be copied to a new file by performing the subtraction and then copying the VAMAS block altered by this operation using the “Data Only” check box option. Any charge correction or other processing can be preserved using this method.

The Calculator

The spectrum calculator is accessible from the Processing Dialog window.

An important use of the calculator is that of constructing difference spectra where the intention is to use the processed data as part of a quantification procedure. There is a fundamental requirement for these operations to maintain the integrity of the data and therefore a limited set of calculation options is available to support this end. The principal limitation is that the number of acquisition channels must be identical in each spectrum used in the calculation. The data may be charge corrected, but only those channels for which the charge correction lies within the original range of channels will the results be meaningful. Figure 45 illustrates how the data is adjusted to account for charge shift corrections prior to performing one of the calculator operations.

The Calculator is designed to allow a single spectrum to act upon a set of spectra. Any spectra selected in the right hand (browser) frame will be altered by the operation defined by both the Calculator property page and the current data in the active display tile.

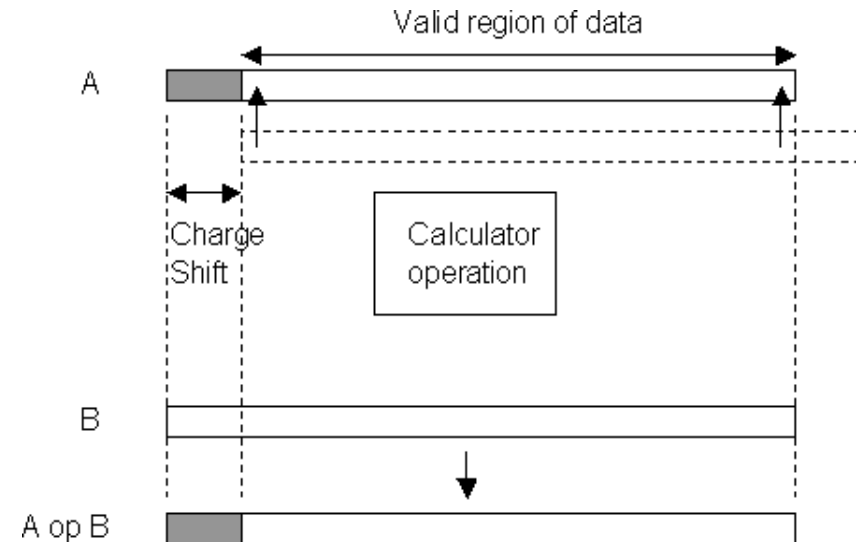


Figure 45: Calculator Operations and Spectrum Acquisition Channels after charge correction

Figure 46 shows a set of spectra in various states of oxidation. The spectrum in the left-side display tile (Figure 46(a)) represents a metallic surface and may be subtracted from the set of spectra selected in the right-hand side of the same experiment-frame to produce a set of difference spectra showing the trend in the Al 2p oxide peak. Figure 46(b) illustrates the result of the subtraction operation shown on the Calculator Property Page, where each of the spectra selected in the right-hand-side browser frame is displayed, showing the residual Al 2p oxide peak following the Calculator operation. Note the processing history indicates that the active tile on the left-hand-side shows data that has been modified by “vb1” or VAMAS BLOCK index 1 (where the indices start at vb0).

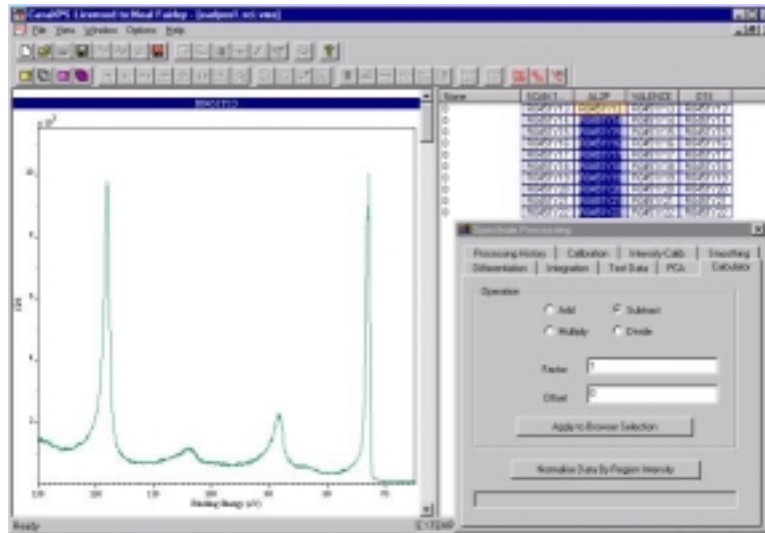


Figure 46: (a) The spectra selected in the browser will be modified by the spectrum displayed in the active display tile.

Normalizing the data (which are the result of a calculator operation) with respect to one another is sometimes necessary. One such situation arises when a peak lies on a plasmon loss structure. Chromium oxide supported on SiO_2 catalyst samples produces a spectrum where the O 1s line includes a pronounced loss structure associated with SiO_2 which impedes measurements for Cr 2p owing to the complex nature of the background, particularly when the amount of chromium is small. If a “standard” spectrum is available representing the SiO_2 catalyst material prior to modification by the Cr, subtracting the standard may result in data for which common

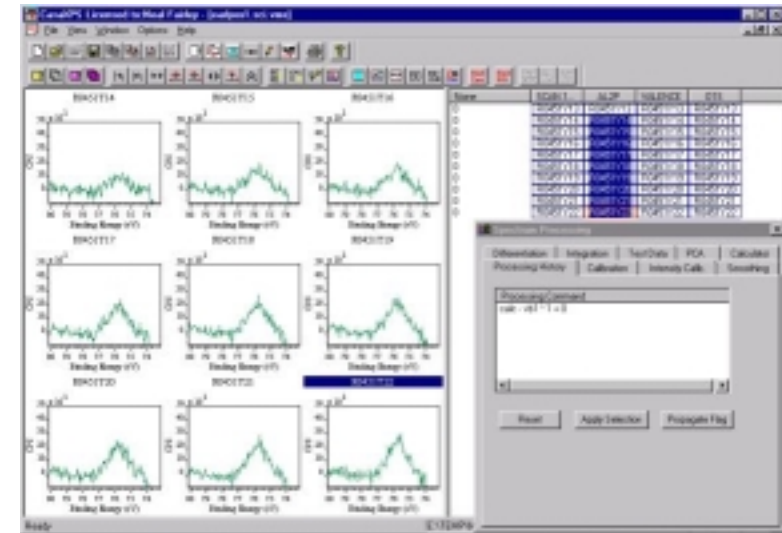


Figure 46: (b) Result of the operation shown alongside.

background types are appropriate. Figure 47 shows two spectra taken from a Cr 2p region. The O 1s plasmon structure is very evident in both spectra and clearly interferes with a consistent choice for the background to the Cr 2p lines. The first step in quantifying the Cr 2p lines is to charge correct one spectrum with respect to the other. Once performed the scale factor representing the difference in count rates may be estimated from the relative size of the oxygen peaks and is used as a factor that is applied to the standard spectrum before subtracting the two Cr 2p data intervals. First the data regions for the unknown and the standard must

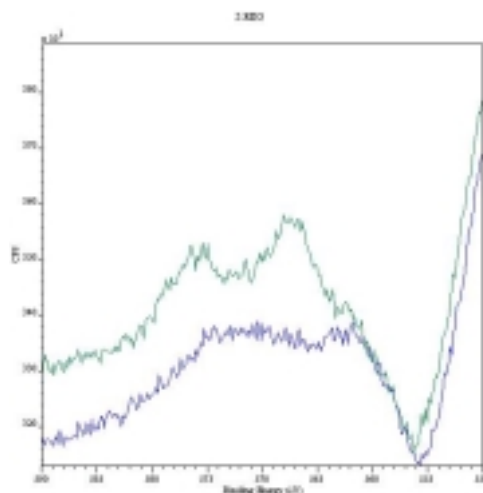


Figure 47: Cr 2p regions acquired from a catalyst before and after an uptake of Cr.

be copied into a new experiment frame. In Figure 48, six regions from each experiment have been copied into an empty file initially called CasaXP1. This file may be subsequently saved with an appropriately changed name.

The steps used to calibrate the energy scale are as follows (refer to Figure 49 and Figure 50):

1. Overlay the two O 1s peaks.
2. Select the Calibrate option on the processing property

page window. Use the mouse to drag out a box from the top of one peak to the top of the second. The limits of the box define the “Measured” and the “True” energies entered on the Calibrate property page, pressing Apply will calibrate the two spectra with respect to one another.

3. Once satisfied with the O 1s calibration, the values for the Measured and True energies may be applied to the corresponding Cr 2p region. Note that only one of the two overlaid O 1s spectra will have an entry in their processing history following the calibration step. The Cr 2p spectrum from the same file as the O 1s with the calibration entry is the Cr 2p spectrum for which these calibration values are appropriate.

Figure 51 and Figure 52 show how to define the calculation for removing the underlying trend from the Cr 2p peaks.

The quantification report printed over the wide spectrum (Figure 48) for the unknown sample shows the results compiled using a Shirley background applied to the difference spectrum for the Cr 2p region. Note that the quantification report is compiled using the results from the high-resolution spectra even though it is displayed over the survey spectrum. Without subtracting the reference spectrum the validity of the Shirley procedure would be in question for the Cr 2p data.

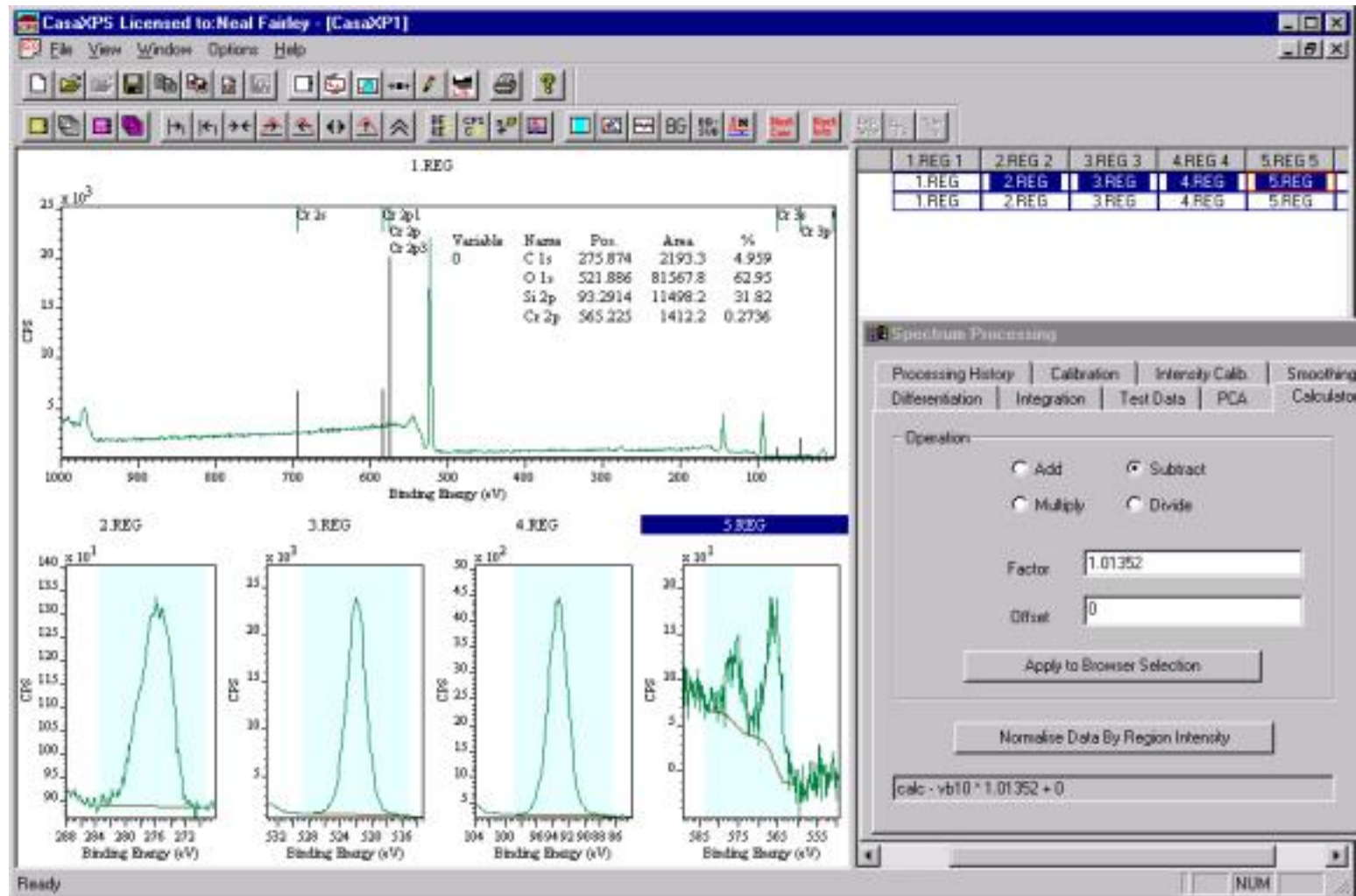


Figure 48: Quantification based upon the Cr 2p data after subtracting the corresponding spectrum from a standard sample

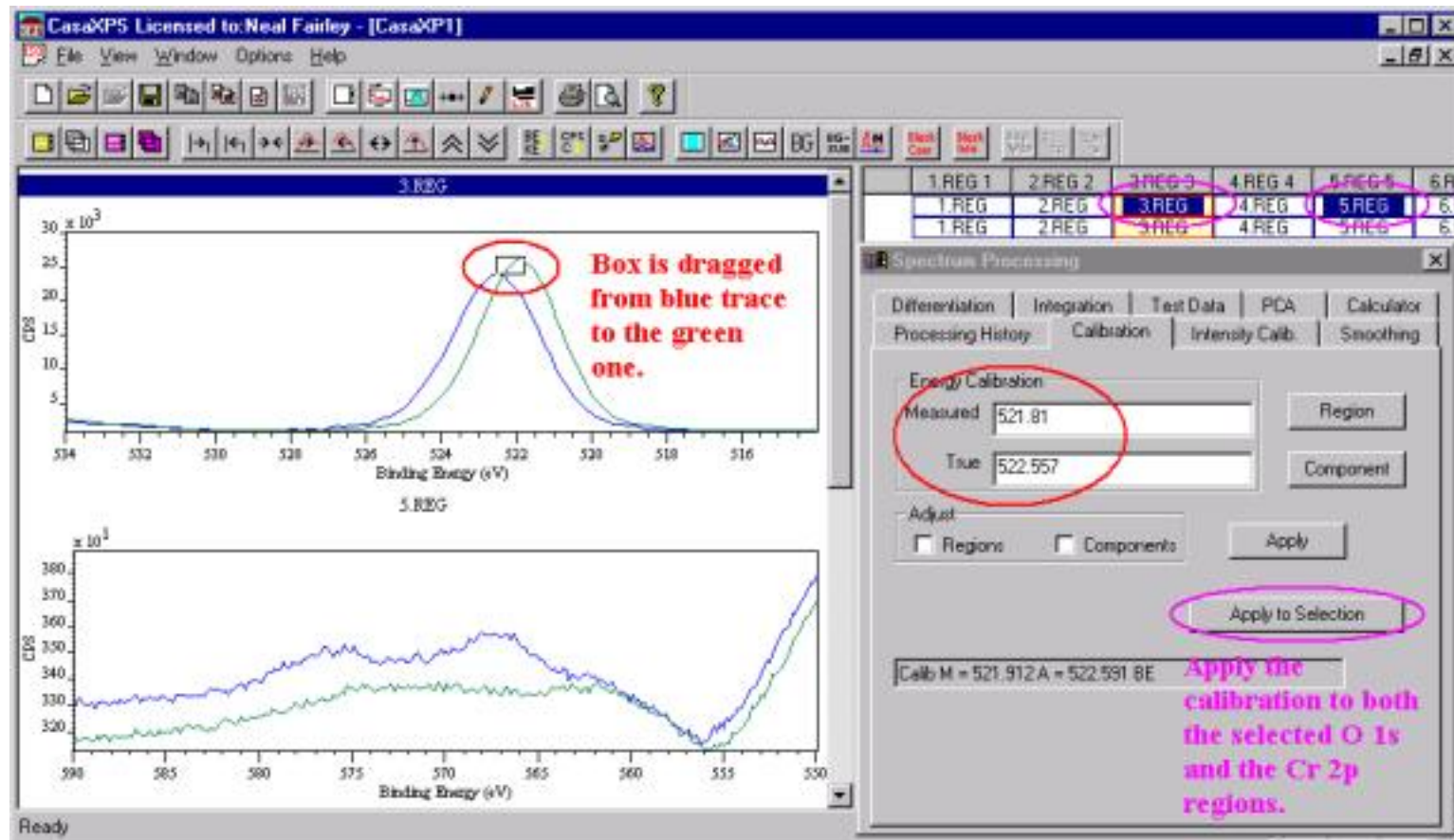


Figure 49: Calibrate the spectra with respect to one another. The calibration point is determined from the overlaid O 1s peaks

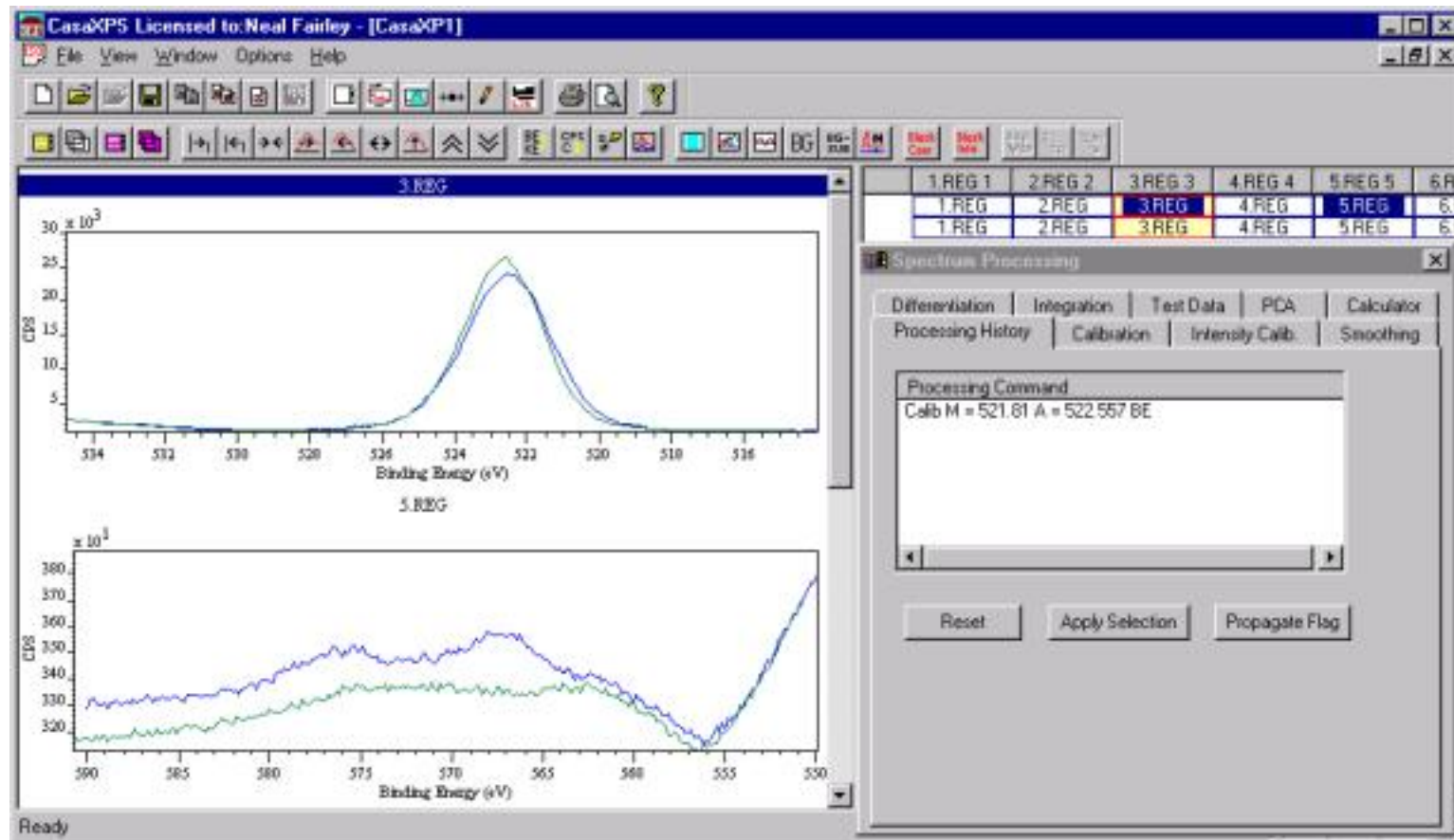


Figure 50: Regions after the calibration has been applied to both the O 1s and the Cr 2p regions using the “Apply to Selection” button on the Calibration Property Page, Figure 46

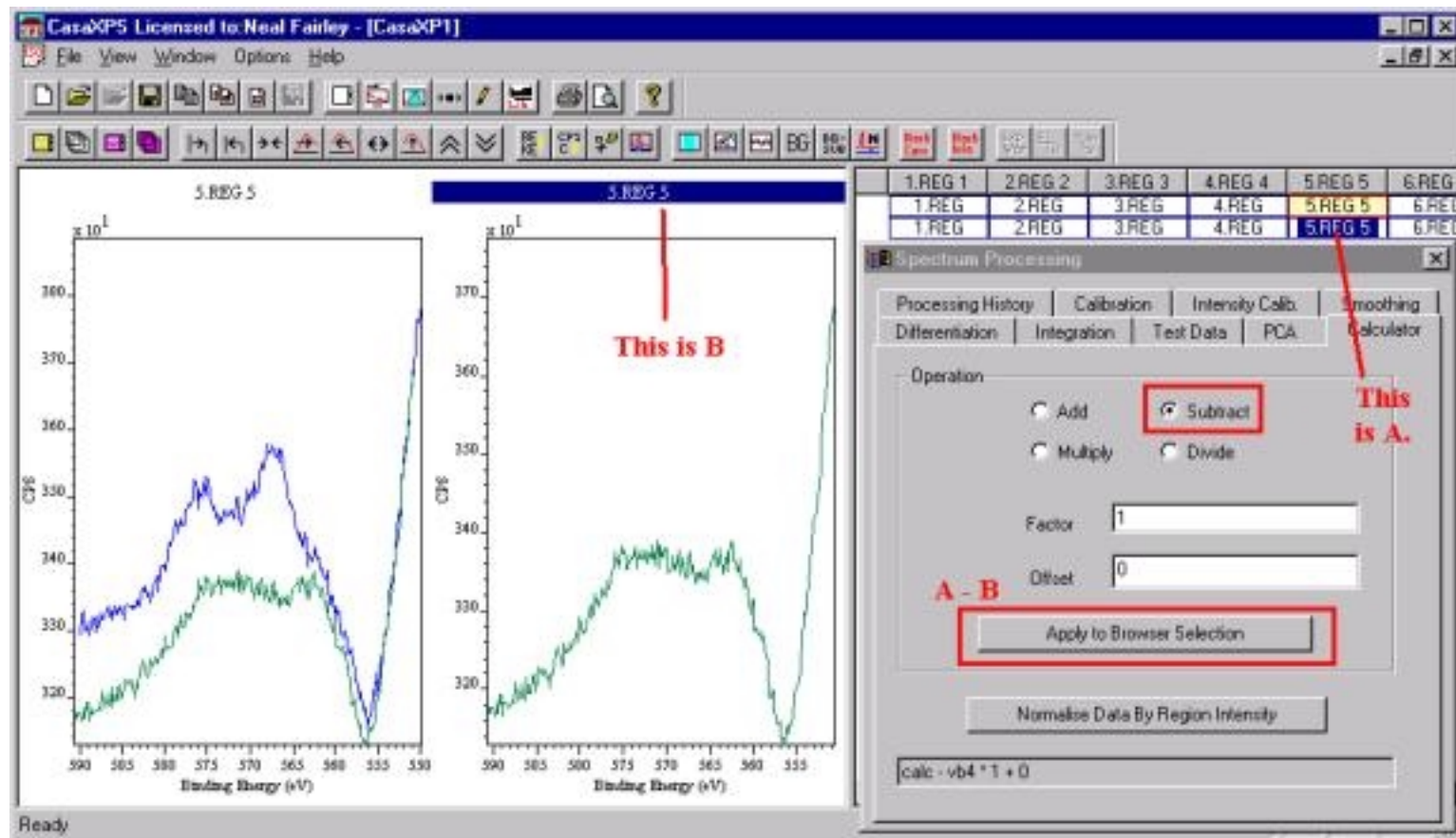


Figure 51: Calculator page setup up to subtract the standard data away from the Cr modified spectrum. The standard is the spectrum labeled B and must be the spectrum in the active display tile. The modified data (A) must be selected via the right-hand-side

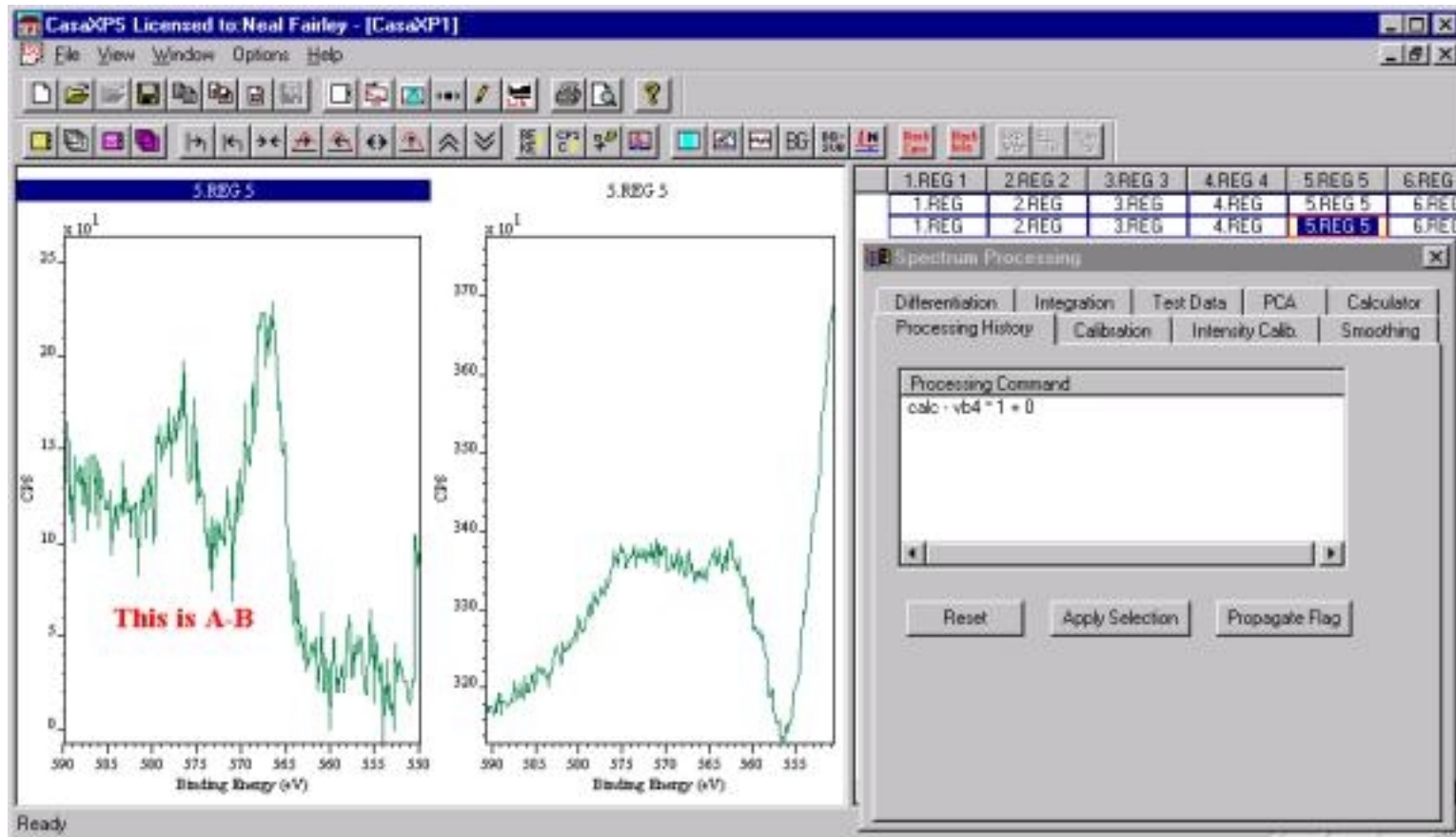


Figure 52: Cr 2p region after the equivalent data from the standard has been subtracted.

Introduction

XPS is a technique that provides chemical information about a sample that sets it apart from other analytical tools. However, the key information sought by the analyst is locked into a data envelope and as a consequence the need for powerful algorithms is paramount when reducing the data to chemically meaningful quantities. Two approaches have been employed on XPS data:

- Curve synthesis and fitting (see “Quantification” on page 42).
- Techniques from multivariate statistical analysis of which Principal Component Analysis (PCA) is the most common form.

Curve synthesis is probably the most widely used method for data analysis practised by XPS researchers. Unfortunately, statistically good curve fits are not always physically meaningful and, in many cases, great care must be exercised when choosing the model to describe the data. Any assist-

ance in understanding the model is therefore of great value and it is with this end that Principal Component Analysis is offered as a supplementary tool.

Factor analysis is a field that is as broad as it is deep. It is a mathematically challenging tool that requires knowledge of matrix algebra coupled with a feel for a statistical approach to data interpretation. A true understanding for the subject can only be obtained by studying the literature and through practical experience. Therefore the material presented here is only an introduction rather than a complete set of works.

Theory of Principal Component Analysis

Factor analysis is a multivariate technique for reducing matrices of data to their lowest dimensionality by use of orthogonal factor space. The challenge is to identify the number of significant factors (principal components) and use this information to model the data using techniques such as Target Transformations or curve fitting.

In XPS the data matrix is composed of spectra where each

acquisition channel is viewed as a co-ordinate in an r -dimensional space; r is equal to the number of acquisition channels per spectrum. The problem addressed by PCA is that of determining the number of distinct spectroscopic features present in a particular set of c spectra.

The following example tries to illustrate the nature of the problem.

Consider a set of three spectra; each spectrum has three acquisition channels:

$$\mathbf{s1} = (4, 3, 6), \mathbf{s2} = (2, 3, 2), \mathbf{s3} = (2, 0, 4)$$

The data matrix is given by

$$D = \begin{bmatrix} 4 & 2 & 2 \\ 3 & 3 & 0 \\ 6 & 2 & 4 \end{bmatrix}$$

These three vectors belong to a 3-dimensional space, however they do not span 3-dimensional space for the following reason. If a linear combination of the vectors s_1 , s_2 and s_3 is used to construct a new vector v , then v always lies in a plane (a 2-dimensional sub-space of 3-dimensional space). The fact that v lies in a plane is a consequence of the following relationships between the three spectra.

$$\mathbf{s3} = \mathbf{s1} - \mathbf{s2},$$

$$\begin{aligned} \text{so } \mathbf{v} &= a \mathbf{s1} + b \mathbf{s2} + c \mathbf{s3} \\ &= a \mathbf{s1} + b \mathbf{s2} + c (\mathbf{s1} - \mathbf{s2}) \\ &= (a + c) \mathbf{s1} + (b - c) \mathbf{s2}. \end{aligned}$$

Thus, two principal components exist for the set of three spectra.

The analysis of the data matrix in the above simple example

has been performed by observation. Unfortunately real spectra are not so simple and spotting the linear relationships between the columns of the data matrix requires a more sophisticated approach.

PCA, also known as Eigenanalysis, provides a method for identifying the underlying spectra that form the building blocks for the entire set of spectra. The data matrix is transformed into a new set of r -dimensional vectors. These new vectors span the same subspace as the original columns of the data matrix, however they are now characterised by a set of eigenvalues and eigenvectors. The eigenvalues provide a measure for the significance of the abstract factors with respect to the original data. Various statistics can be computed from these values that aid in identifying the dimensionality of the subspace spanned by the spectra.

The procedure for calculating the abstract factors has its roots in linear least square theory. In fact the preferred method is to form a Singular Value Decomposition (SVD) for the data matrix.

$$D = USV'$$

Where D is the data matrix formed from c spectra, each containing r channels. U is the same dimension as D , while S and V are c by c matrices. S is a diagonal matrix; the diagonal elements are the square root of the eigenvalues of the correlation matrix

$$Z = D'D$$

The abstract factors are computed from US . The rows of V are the corresponding eigenvectors of Z ; the co-ordinates of the eigenvectors represent the loading for the abstract factors and specify how linear combinations of these factors can be used to reproduce the original data. Including all of

the abstract factors with the appropriate loading enables the data to be reproduced to an accuracy only limited by the precision characteristic of the Eigenanalysis procedure.

The essential feature of the SVD procedure is to compute the abstract factors so that the factor corresponding to the largest eigenvalue accounts for a maximum of the variation in the data. Subsequent abstract factors are generated such that 1) as much variance as possible is accounted for by each new factor and 2) the newest axis is mutually orthogonal to the set of axes already located. The procedure therefore computes an orthogonal basis set for the subspace spanned by the original data matrix that is oriented with respect to the data in a linear least square sense.

In principle, the number of non-zero eigenvalues is equal to the number of linearly independent vectors in the original data matrix. This is true for well posed problems, but even the presence of errors due to numerical operations will result in small eigenvalues that theoretically should be zero. Numerical errors are an insignificant problem compared to the one presented by the inclusion of experimental error in the calculation. Noise in the data changes the underlying vectors so that almost every data matrix of c spectra with r acquisition channels, where $c \leq r$, will span a c -dimensional subspace. This is true even though the underlying vectors should only span fewer than c dimensions.

Various statistics are available for identifying the most likely dimensionality of a data matrix. These statistics are designed to aid partitioning the abstract factors into primary and secondary factors. The primary factors are those corresponding to the largest n eigenvalues and represent the set of abstract factors that span the true subspace for the data. The

secondary factors are those factors that can be associated with the noise and, in principle, can be omitted from subsequent calculations. It is not possible to completely disassociate the true data from the error within the measured data, however the statistics guide the analyst in choosing the most appropriate number of abstract factors that describe the data and therefore the "best guess" dimensionality for the data matrix.

In the case of XPS spectra the experimental error is known to be the square root of the number of counts in an acquisition channel. Under these circumstances where the experimental error is known, a number of statistics have been proposed for determining the size of the true factor space.

Residual Standard Deviation ("Real Error")

An alternative name for the residual standard deviation (RSD) used by Malinowski is "real error" (RE).

The RSD is defined to be:

$$RSD_n = \sqrt{\frac{\sum_{j=n+1}^c E_j}{r(c-n)}}$$

where E_j is the j th largest eigenvalue, n is the number of abstract factors used to reproduce the data; c spectra each with r channels are used to construct the data matrix.

RSD_n must be compared against the estimated experimental error. If the value computed for RSD_n is approximately equal to the estimated error then the first n abstract factors span the factor space. The dimensionality of the original

data matrix is therefore n .

Two further statistics may be derived from RSD_n , namely, IE_n (Imbedded Error) and IND_n (Indicator Function) given by:

$$IE_n = RSD_n \left(\frac{n}{c} \right)^{\frac{1}{2}}$$

And

$$IND_n = \frac{RSD_n}{(n-c)^2}$$

IE_n and IND_n are statistics that should decrease as the number of primary abstract factors is increased. Once all the primary factors have been included, these statistics should begin to increase since at this point factors from the noise subspace start to interfere with the accuracy of the data description. This minimum is therefore an indicator of the dimensionality of the data subspace.

Chi square

Bartlett proposed using the chi square criterion for situations similar to XPS data, where the standard deviation varies from one data channel to the next.

The procedure involves reproducing the data matrix using the abstract factors. Each abstract factor is progressively included in a linear combination in the order defined by the

size of the eigenvalues and weighted by the co-ordinates of the corresponding eigenvectors. The chi square value for a set of n abstract factors is computed using:

$$\chi_n^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(d_{ij} - \underline{d}_{ij})^2}{\sigma_{ij}^2}$$

where d_{ij} is an element of the data matrix, \underline{d}_{ij} is the corresponding approximation to the data point constructed from the first n abstract factors with the largest eigenvalues. The standard deviation for XPS data σ_{ij} is the square root of d_{ij} . The expected value for each n is given by:-

$$\chi_n^2 (\text{expected}) = (r-n)(c-n).$$

A comparison between the expected value and the computed value is the basis for determining the number of principal components. Both χ_n^2 and its expected value decrease as n increases. χ_n^2 initially is larger than χ_n^2 (expected) but as n increases a crossover occurs. The true dimensionality of the data matrix is chosen to be the value of n for which χ_n^2 is closest to its expected value.

Note that smoothing the data will alter the characteristics of the noise. Performing such pre-processing therefore invalidates the χ_n^2 statistic.

Target Factor Analysis

Principal Component Analysis provides a set of basis vectors that describe the original set of spectra. Although useful as a means of characterising the data, these abstract factors

are in general not physically meaningful. Target Factor Analysis is concerned with identifying vectors that can also describe the data, but with the additional property that they are recognisable as spectra rather than simply abstract vectors in an r -dimensional space.

There are numerous methods for transforming the PCA abstract factors to provide vectors that are more open to chemical interpretation. These involve constructing abstract rotation transformations that map the abstract factors into one of the infinite number of alternative basis sets for the factor space. Fortunately there is a technique which when coupled with curve synthesis, lends itself to the analysis of XPS data, namely, Target Testing.

Target Testing

Once a Principal Component Analysis has been performed, the mathematical bridge between abstract and real solutions is Target Testing. Individual spectra can be evaluated to assess whether the corresponding vector lies in the subspace spanned by the chosen primary abstract factors. The essential feature of Target Testing is to form the projection of the target vector onto the subspace spanned by the primary factors, then compute the predicted target vector using this projection. Statistical tests applied to the predicted and test vectors determine whether these two vectors are one and the same. These tests serve as a mean of accepting or rejecting possible fundamental components of the sample.

Ultimately, the goal of target testing is to identify a set of spectra that span the same subspace as the primary abstract factors. Complete models of real factors are tested in the target-combination step. In the combination step the data ma-

trix is reproduced from the real factors (spectra) rather than from abstract factors and by comparing the results for different sets of real factors, the best TFA solution to a problem can be determined.

Testing a target vector \mathbf{x} with respect to the chosen set of primary abstract factors involves forming the projection \mathbf{t} onto the subspace spanned by the PCA primary abstract factors. The predicted vector $\hat{\mathbf{x}}$, calculated using the co-ordinate values of \mathbf{t} to load the corresponding abstract factors, is compared to the original target vector. A target vector that belongs to the subspace spanned by the primary abstract factors should result in a predicted vector that is identical to the initial target vector. Errors in the original data matrix and similar errors in the measured target vector mean that the predicted and target vector differ from each other as well as from the pure target vector \mathbf{x}^* (\mathbf{x} but without error). Estimates for these differences allow a comparison to be made between the predicted and target vector and a decision as to which targets to include in the target combination step.

The apparent error in the test vector (AET) measures the difference between the test and predicted vectors in a root mean square (RMS) sense. Similarly two other RMS quantities estimate the real error in the target vector (RET) and the real error in the predicted vector (REP). These error estimates form the basis for the SPOIL function defined to be approximately equal to the ratio RET/REP.

Principal Component Analysis by Example

The first example illustrating the characteristics of PCA uses a set of artificial data.

Three sets of spectra prepared from synthetic components are used in the PCA. The structure of the artificial data derives from Carbon 1s states within three compounds, namely, PMMA, PVA and PVC (Figure 53). The proportion of each compound varies throughout each set of ten VAMAS blocks. The data is located in the files c1stest1.vms, c1stest2.vms and c1stest3.vms. The underlying trends introduced into each file are as follows: peaks corresponding to PMMA and PVC obey quadratic adjustments in intensity over the set of ten spectra (PMMA decreases while PVC increases). The difference between the three files is the proportion of PVA in each data envelope. The first file (c1stest1.vms) has a constant level of PVA (Figure 54); the second file (c1stest2.vms) varies linearly, first increasing then decreasing; the third file (c1stest3.vms) includes a linear increase in the level of PVA.

The objective is to show how the statistics used in PCA behave for a known problem. Data matrices constructed from the three sets of spectra should have a dimensionality of three.

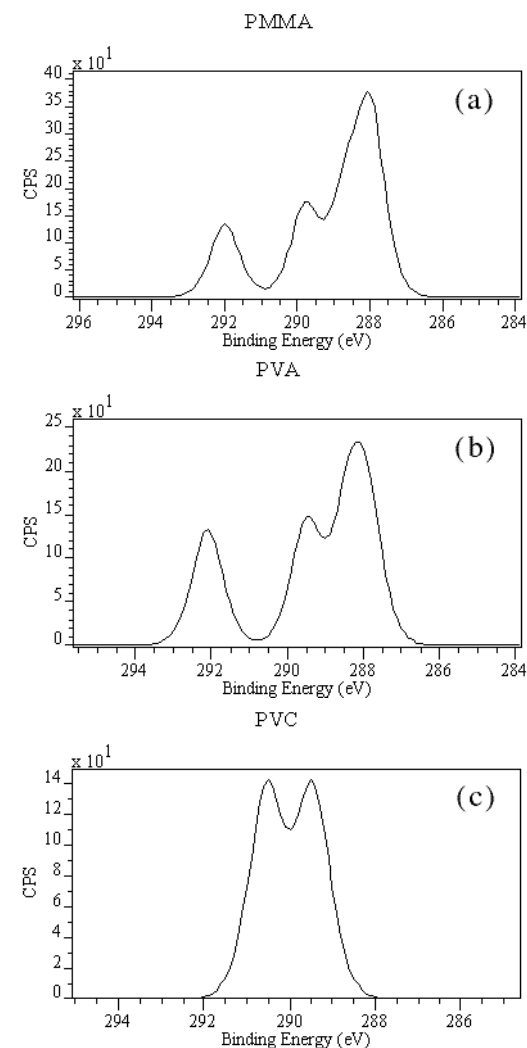


Figure 53: (a - c) Artificial C 1s Data

Note that, although each compound is constructed from a number of C 1s peaks (PMMA 4, PVA 4 and PVC 2), the stoichiometry of these compounds masks the true number of synthetic components actually present in the data. Hence the dimensionality of the data should be three not ten (4+4+2). An additional twist to this example is that two of the underlying envelopes are similar in shape to each other, though not identical (see Figure 53 (a) and (b)).

The trend throughout the first data set may be seen in Figure 54.

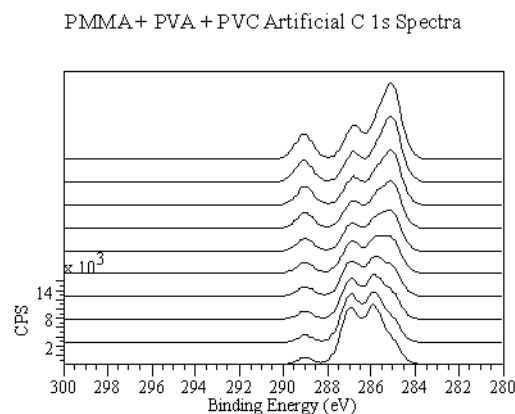


Figure 54: c1s_test1.vms C 1s Spectra

No noise is present in the data; therefore eigenvalues belonging to the primary set of three abstract factors should be non-zero, while the remaining seven eigenvalues should be zero.

The results of applying PCA to these data sets (Table 6) illustrate the uncertainty associated in estimating the dimensionality of the data matrix from the statistics. The fourth largest eigenvalue in each case is small but non-zero. Also the statistics for IE and IND indicate a minimum at eigenvalues other than the expected result. Chi square is not a valid statistic since no noise is present in the data, however it does show that three abstract factors are sufficient to reproduce the data to within reason.

Table 6: PCA report for file c1stest.vms

Factor	Eigenvalue	RMS	RE (RSD)	IE	IND * 1000	Chi-sq Calc.	Chi-sq Expected
C 1s/1	10861690000	266.7944	615.8053	194.7347	7602.534	29055.89	1800
C 1s/2	684568100	13.0963	29.86145	13.35445	466.5852	172.5176	1592
C 1s/3	1433862	0.01064209	0.02964045	0.01623474	0.6049071	0.000190564	1386
C 1s/4	1.230771	0.000391784	0.002107629	0.001332982	0.05854525	9.86356E-08	1182
C 1s/5	0.003712633	0.000385433	0.001279202	0.000904532	0.05116807	1.20367E-07	980
C 1s/6	0.000545838	0.000230365	0.001168993	0.000905498	0.07306205	4.86336E-08	780
C 1s/7	0.000473059	0.000240069	0.001018602	0.000852223	0.113178	5.18512E-08	582
C 1s/8	0.000306354	0.000155331	0.000891207	0.000797119	0.2228016	2.20338E-08	386
C 1s/9	0.000200465	0.00012725	0.00076887	0.000729414	0.7688698	1.91008E-08	192
C 1s/10	0.000118823	4.68061E-13	0	0	0	6.13259E-23	0

It is interesting to see how the eigenvalues change with respect to the three data sets (Figure 55 and Figure 56). The same spectra varied in different ways results in slightly different orientations for the principal component axes and hence different eigenvalues.

The PCA statistics IE and IND have implied a dimensionality other than three (Table 6). The clue to the correct dimensionality of the data lies in the relative size of the eigenvalues. The fourth eigenvalue is in two cases better than five orders of magnitude smaller than the third eigen-

value. This statement has been made with the benefit of a good understanding of what is present in the data. In real situations such statements are themselves suspect and so require support from other data reduction techniques. For example curve fitting using three sets of synthetic peaks all linked with the appropriate stoichiometric relationships would lend support to the hypothesis. Curve fitting such structures is not an exact science and such fits themselves should be supported by statistics gathered from the fitting parameters.

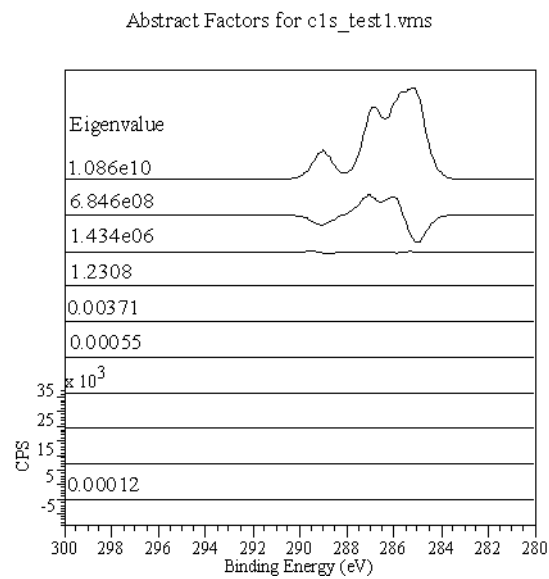


Figure 55: Abstract Factors and Eigenvalues.

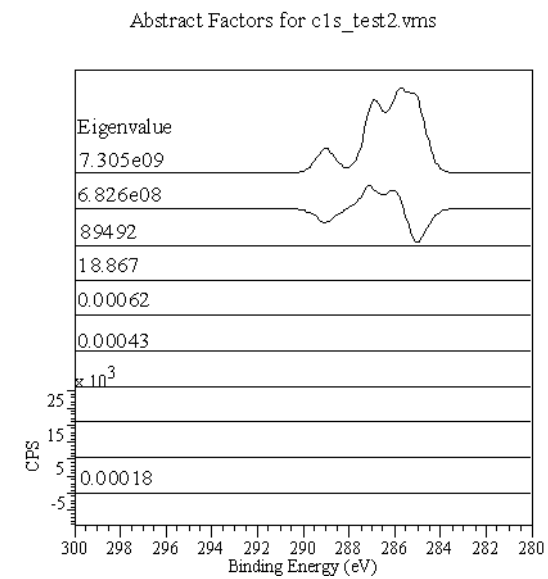


Figure 56: Eigenvalues for c1s_test2.vms Abstract Factors

The second example illustrates the effects of experimental error on a PCA calculation.

Table 7: PCA Applied to Artificial Data with Simulated Noise

Factor	Eigenvalue	RMS	RE (RSD)	IE	IND * 1000	Chi-sq Calc.	Chi-sq Expected
C 1s/1	10883620000	267.626	617.9728	195.4202	7629.294	50489.63	1800
C 1s/2	687172800	18.93302	47.75373	21.35612	746.152	3572.687	1592
C 1s/3	2714766	6.90718	26.01386	14.24838	530.8951	483.7729	1386
C 1s/4	946794.7	0.1649431	2.106561	1.332306	58.51557	0.5688374	1182
C 1s/5	5095.574	0.0622169	0.5048702	0.3569971	20.19481	1.247441	980
C 1s/6	226.9936	0.01656361	0.1904914	0.147554	11.90571	0.002773735	780
C 1s/7	29.13143	0.002688086	0.00847462	0.007090376	0.9416245	5.0833E-05	582
C 1s/8	0.03943089	0.000339289	0.003105158	0.002777337	0.7762894	3.05649E-07	386
C 1s/9	0.003409306	0.000206314	0.001523904	0.001445703	1.523904	9.22657E-08	192
C 1s/10	0.000466779	1.26006E-12	0	0	0	7.4912E-22	0

Real data includes noise. The effect of noise on a PCA calculation can be seen from Figure 57 together with the report in Table 7. The data in the file c1stest1.vms has been used together with a pseudorandom number generator to simulate noise that would typically be found in XPS data. The consequence of including a random element in the data is that the

eigenvalues increase in size and lead to further uncertainty with regard to which eigenvalues belong to the set of primary abstract factors. Note that the abstract factors in Figure 57 are plotted in the reverse order to the ones in Figure 55 and Figure 56

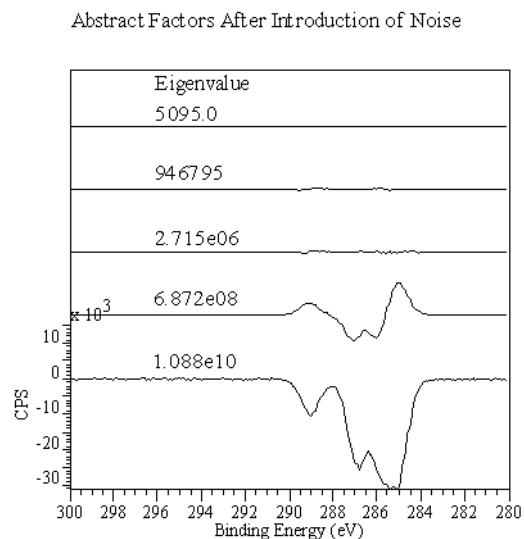


Figure 57: c1stest1.vms data plus artificial noise.

Fortunately, the chi square statistic becomes more meaningful when noise is introduced into the problem. A comparison between the computed chi square and its expected values do seem to point to a 3-dimensional subspace. The crossover between the two quantities suggests the need for three abstract factors when approximating the data matrix using the results of PCA.

Principal Component Analysis and Real Data

XPS depth profiles generate sets of spectra that are ideal for examination via PCA. The spectra are produced by repeatedly performing etch cycles followed by measuring the count rate over an identical energy region. The resulting data set therefore varies in chemical composition with respect to

etch time and the common acquisition conditions provide data in a form that is well suited to PCA.

An aluminium foil, when profiled using a Kratos Analytical Axis ULTRA, provides a good example of a data set that can be analysed using some of the features on offer in CasaXPS. These data are not chemically or structurally interesting, but do show how trends can be identified and anomalies isolated.

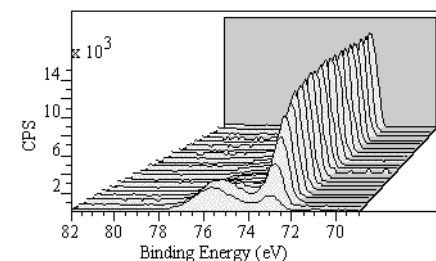


Figure 58: Al 2p Depth profile

Figure 58 shows a plot of the Al 2p energy-region profiled against etch time. The data envelopes change in shape as the surface oxide layer is removed by the etch cycles to reveal the homogeneous bulk aluminium metal.

It should also be noted from Figure 59 that the data contains an imperfection. One of the energy scans includes data acquired during an instrumental event. Noise spikes are superimposed on the true data and these should be removed before techniques such as curve synthesis are applied. In this case the spikes appear in the background and are therefore obvious to the eye, however similar non-physical structure that occurs on the side of a peak is less obvious and could be missed.

The first step in performing the Principal Component Analysis is to define a quantification region for each of the spectra to be used in the analysis. These regions specify the acquisition channels that will be used in the data matrix. Also any shifts in the data due to charging can be removed from the calculation by using an offset in the energy region for those spectra affected.

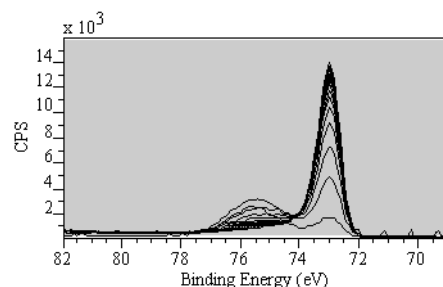


Figure 59: Al 2p Spectra

Next, select the set of spectra in the Browser View and display the data in the active tile. The processing property page labelled "PCA" offers a button labelled "PCA Apply". On pressing this button, those spectra displayed in the active tile are transformed into abstract factors. Figure 59 displays the spectra before the PCA transformation while Figure 60 shows the abstract factors generated from the eigenanalysis.

Note that the abstract factors are truly abstract. The first factor (Figure 60) looks like an Al 2p metal doublet, however this is because the Al 2p metal envelope dominates the data set and therefore a vector having a similar shape accounts

for most of the variation in the overall set of spectra. A more even weighting between the underlying line-shapes would produce abstract factors that are less physically meaningful in appearance.

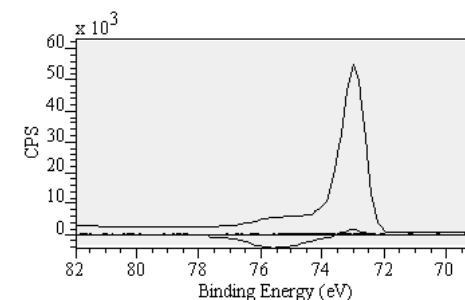


Figure 60: Al 2p Abstract Factors

The only real use for the abstract factors is judging their significance with respect to the original data. Abstract vectors that derive from noise look like noise, factors that contribute to the description of the data contain structure. The dividing line between the primary and secondary abstract factors can sometimes be assessed based on the appearance of the abstract factors.

Analysing the Al 2p spectra generates abstract factors and eigenvalues that represent the PCA fingerprint for the data. Table 8 is a report of the Al 2p data set generated by CasaXPS and formatted using a spreadsheet program. Each row of the report is labelled by the ISO block name that contains the abstract factor corresponding to the listed eigenvalue.

Table 8: Report generated by a PCA for AI 2p Profile.

Factors	Eigenvalue	RMS	RE (RSD)	IE	IND * 1000	Chi-sq(Calc.)	Chi-sq(Expected)
AI 2p/3	4727872000	82.5236	159.3106	34.76442	398.2765	26032.99	2600
AI 2p/8	65086870	3.784555	23.78739	7.340948	65.89305	136.4193	2451
AI 2p/13	393178.7	5.875731	20.74929	7.842494	64.04102	133.7949	2304
AI 2p/18	305001.5	3.48073	17.85783	7.793798	61.79182	78.0949	2159
AI 2p/23	156694.8	3.249367	16.25038	7.929371	63.47803	59.6866	2016
AI 2p/28	98359.06	2.757161	15.2192	8.135007	67.64091	51.1035	1875
AI 2p/33	86168.29	2.48836	14.18397	8.189118	72.36718	41.82519	1736
AI 2p/38	65267.54	2.22333	13.35424	8.242415	79.01916	37.98182	1599
AI 2p/43	53613.14	2.247765	12.61316	8.257255	87.59142	42.46731	1464
AI 2p/48	43569.08	0.1744253	11.97161	8.261198	98.93895	0.1869893	1331
AI 2p/53	32387.23	1.710532	11.52946	8.344409	115.2946	29.01139	1200
AI 2p/58	28174.98	2.021671	11.12658	8.410907	137.3652	33.42906	1071
AI 2p/63	24742	1.261896	10.75487	8.461885	168.0448	15.17705	944
AI 2p/68	23980.27	0.40768	10.29759	8.407944	210.1548	1.649807	819
AI 2p/73	20710.42	1.113217	9.867347	8.33943	274.093	10.69842	696
AI 2p/78	18345.12	1.155456	9.424948	8.226769	376.9979	15.75993	575
AI 2p/83	16109.71	0.7358818	8.960655	8.062218	560.0409	5.605129	456
AI 2p/88	13003.76	0.7303461	8.600543	7.962556	955.6159	5.882052	339
AI 2p/93	12307.15	0.7049177	7.99876	7.608339	1999.69	6.193881	224
AI 2p/98	9285.948	0.000443667	7.554815	7.372745	7554.815	2.12747E-06	111

Table 8: Report generated by a PCA for Al 2p Profile.

Factors	Eigenvalue	RMS	RE (RSD)	IE	IND * 1000	Chi-sq(Calc.)	Chi-sq(Expected)
Al 2p/103	7476.855	1.90305E-13	0	0	0	2.03675E-25	0

The chi square indicates that the data matrix can be reproduced to within experimental error using two abstract factors. This is a result that is consistent with the physical nature of the sample. It is also interesting (from a mathematical standpoint) to note that using all the abstract factors to reproduce the data matrix returns a chi square of zero (allowing for round-off errors in the computation). This should always be the case and provides an easy check to see that the calculation has been performed correctly.

All the statistics except the Indicator Function point to two abstract factors being sufficient to span the factor space for the data matrix.

It is worth examining the data set using a subset of the spectra and Target Testing the spectra not used in the PCA. This allows anomalies to be identified such as spikes in the data. Selecting a representative subset of spectra for the PCA then target testing the remainder is particularly useful for large sets of data.

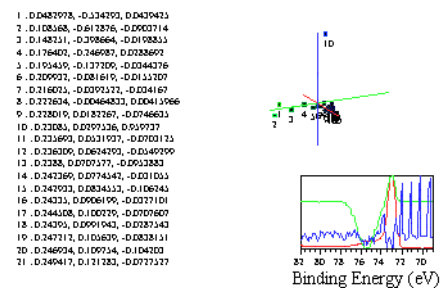
Table 9: Target Test Report for a Subset of Al 2p Data Set

Target	AET	REP	RET	SPOIL	Al 2p/3	Al 2p/8
Al 2p/23	20.93032	10.38374	18.17295	1.750135	0.39033	-0.04947
Al 2p/28	23.83028	11.05117	21.11288	1.910467	0.41839	0.017257
Al 2p/33	19.83736	11.47927	16.17861	1.409376	0.42997	0.065749
Al 2p/38	19.8507	12.01348	15.80274	1.315418	0.44268	0.10609
Al 2p/43	19.9069	12.46508	15.52116	1.245171	0.4531	0.133366
Al 2p/48	57.16561	12.70691	55.73546	4.386233	0.45854	0.14688
Al 2p/53	15.37333	13.18052	7.912861	0.600345	0.46791	0.174614
Al 2p/58	21.39836	13.30379	16.76004	1.259795	0.46901	0.184805
Al 2p/63	19.92528	13.5238	14.63296	1.082016	0.47386	0.195062
Al 2p/68	27.73522	13.78354	24.06775	1.746122	0.48087	0.203826

Table 9: Target Test Report for a Subset of Al 2p Data Set

Target	AET	REP	RET	SPOIL	Al 2p/3	Al 2p/8
Al 2p/73	19.10189	13.88023	13.12332	0.945469	0.48192	0.210646
Al 2p/78	20.9575	13.98145	15.61204	1.116625	0.48264	0.218455
Al 2p/83	19.03813	14.15492	12.7314	0.899433	0.48483	0.229382
Al 2p/88	18.38591	14.11378	11.78317	0.83487	0.48374	0.228046

The SPOIL function and AET statistics (Table 9) show that Al 2p/48 differs in some respect from the other spectra in the list tested. The spectrum in question corresponds to the trace displaying the spikes seen in Figure 58. Also, another spectrum that could be looked at is Al 2p/68. The AET value is high compared to the other spectra. Such spectra may highlight interfaces where either new chemical states appear (either directly from features in the data or indirectly through changes in the background due to features outside the acquisition region) or energy shifts due to sample charging have altered the characteristics of the data.

**Figure 61:** 3-D factor space

The PCA report in Table 3 includes the spectrum labelled Al 2p/48 in the data matrix. The consequence of not removing the spikes is apparent in the 3-D factor space shown in Figure 61, where the abstract factor with third largest eigenvalue clearly contains spikes and the projection point number 10 derived from the Al 2p/48 spectrum is obviously a statistical outlier.

PCA and CasaXPS

Principal Component Analysis is offered on the "processing" window. The options on the property page labelled "PCA" allow spectra to be transformed into abstract factors according to a number of regimes. These include covariance about the origin and correlation about the origin. Each of these pre-processing methods may be applied with and without background subtraction.

Quantification regions must be defined for each spectrum included in the factor analysis. In addition, each spectrum must have the same number of acquisition channels as the others in the set of spectra to be analysed. The first step in the calculation replaces the values in each spectrum by the result of interpolating the data within the defined quantifica-

tion region for the spectrum. This is designed to allow energy shifts to be removed from the data used in the factor analysis.

The quantification region also provides the type of background to the spectrum. Performing the analysis on background subtracted data attempts to remove artifacts in the spectrum that derive from other peaks within the vicinity of the energy region. Background contributions can be significant in PCA. Additional primary abstract factors are often introduced as a consequence of changes in the background rather than the underlying peaks within the region of interest. The presence of such abstract factors can be viewed as information extracted from the data, although in many circumstances they can lead to incorrect synthetic models if background contributions are misunderstood.

A factor analysis is performed on the set of spectra displayed in the active tile. Although PCA is offered as a processing option, it is the only processing option that acts on a collection of spectra. Any other option from the processing window would only act upon the first VAMAS block in a selection when that selection is displayed in a single tile.

The principal component analysis is performed when the "Apply" button is pressed. Each spectrum displayed in the active tile is replaced by the computed abstract factors. The order of the VAMAS blocks containing the spectra is used as the order for the abstract factors. The factor corresponding to the largest eigenvalue is entered first. Subsequent blocks receive the abstract factors in descending order defined by the size of the corresponding eigenvalues. A report showing the statistics for understanding the dimensionality of the factor space appears in a dialog window.

A button labelled "PCA Report" allows the current PCA report to be re-displayed. Care should be exercised since the values are subject to any additional processing (including PCA) that may subsequently be applied to any of the spectra included in the original analysis.

The PCA property page includes a button to reset the processing operations for every spectrum displayed in the active tile. This allows a PCA calculation to be undone in one stroke. It will also undo any processing previously performed on the data. PCA is aimed at the raw data; the chi square statistic is referenced to the raw data and has an undefined meaning when the data have been processed prior to performing factor analysis.

Target Factor Analysis in the form of target testing is also available on the PCA property page. Following a PCA, candidates for the physically meaningful components may be assessed individually or collectively. Choose an abstract factor from the PCA and entering this factor into the active tile. Then select the number of primary abstract factors for use in the target test procedure. A text field is offered on the PCA property page for this purpose and is found in the section headed "Target FA". Next, select the target test spectra in the Browser view and press the button labelled "TFA Apply". A report detailing the statistics calculated from the TFA procedure will appear in a dialog window.

The TFA report may be written to file in an ASCII format with TAB separated columns. When pressed, any of the buttons above the columns on the report will display a file dialog window from which the output text-file can be specified. This method for saving a report to file is used by the PCA report (above) and the Linear Regression Report described

below.

Once a set of target spectra has been identified, these spectra can be used to reproduce the original set of spectra through a linear regression step. Enter the set of target spectra into the active tile; then select the original spectra in the Browser view. Press the button labelled "Linear Regression". A report shows the RMS differences between each of the original spectra and the predicted spectra calculated from a linear combination of the set of target spectra displayed in the active tile. The loading used to compute the predicted spectra are listed in the report. The report may be written to file using a similar procedure to the TFA report described above.

Viewing the Data in Factor Space

CasaXPS offers an option on the "Geometry" property page on the "Tile Display" dialog window labelled "Factor Space". If selected, the VAMAS blocks displayed in a tile are used to define the axes for a subspace and the original data are plotted, if possible, as a set of co-ordinates with respect to these axes. The plot represents a projection of the data space onto the subspace defined by a set of two or three abstract factors.

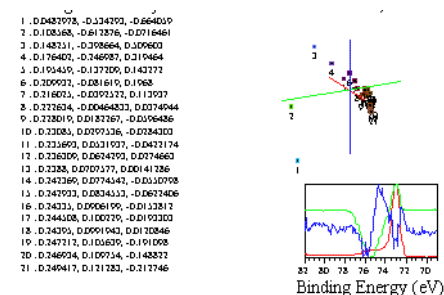


Figure 62: Projection onto Abstract factors 1, 2, and 4

The abstract factors defining the axes are graphed together with a list of the co-ordinate values for each of the spectra projected onto the subspace spanned by the chosen abstract factors (Figure 62). A 3-dimensional plot provides a visual interpretation for the spectra. Patterns formed by the spectra highlight trends within the data set and the relative importance of the abstract factors can be examined. A plot in which the axes are defined by unimportant factors generally appear random, while factors that are significant when describing the data typically produce plots containing recognisable structure.

Monte Carlo Methods

- Uncertainties in Intensity Calculations

CHAPTER 11

Optimization routines are completely deterministic, in that the same set of parameters, used in the same functional forms, applied to the same set of data with the same initial and final conditions will result in the same set of parameters on termination. Vary any of the above conditions and the result from the optimization routine will change in some respect. One method for assessing the uncertainty in the parameters for a peak model is to vary these optimization conditions by repeating an experiment on, what are hoped to be, identical samples. Then for each set of data apply the same optimization routine to the same synthetic model and so determine a distribution for the set of parameters used to quantify a sample. Such a procedure will vary almost every aspect of the measurement process and so result in a distribution for the parameters that truly represent the nature of the experiment.

The basis for such an approach as described above lies in the assumption that there exists a set of parameters (only known to nature) that does not depend on any optimization routine nor any other errors introduced into the measurement proc-

ess, and these true values will lie inside the region of the N-dimensional parameter space defined by the set of outcomes to this sequence of experiments. Obviously, if the synthetic model does not describe a set of parameters in tune with nature, the results may be in more doubt than the measured distribution might suggest. However, given that all is well then the task is to offer a means of understanding the uncertainties in the peak parameters within the context of these parameter distributions.

Peak identification in XPS spectra represents a challenge since synthetic models more often than not involve overlapping line-shapes (Figure 63), the consequence of which is correlated optimization parameters. That is to say, if a single data envelope results from two overlapping peaks and if one of these underlying peaks is reduced in intensity then in order to describe the same data envelope the other must increase in intensity. Adjustments to peak parameters of this nature are inherent to any optimization procedure and the choice between the possible combinations of these peak intensities is made based upon a chi square or root mean

square metric. The problem is therefore to identify the point at which these goodness-of-fit metrics fail to produce results that can be believed, and provide some means of illustrating the degree of uncertainty.

Given a distribution for each of the parameters, the best way to describe the nature of the uncertainties is to offer an error matrix as well as a tabulated form for the distributions. The error matrix provides numerical values from which the degree of correlation can be assessed while scatter plots taken from some subset of these distributions allows visual inspection for the same information. Ideally a single number for each parameter would provide the simplest means of stating the uncertainties, but as the old adage goes “To every complex problem there is a simple solution that is wrong”, and so it is for correlated parameters. The unfortunate fact is that if the peaks weren’t correlated then synthetic models would be unnecessary.

Monte Carlo Data Sets

Repeating an experiment many times is not always practical and for most XPS spectra peak models are developed based upon a single acquisition sequence. Estimates for uncertainties in the peak parameters must be made using assumptions about the noise typical of XPS spectra. The essence of Monte Carlo methods for estimating uncertainties is to take a data set, remove the noise from the data, then repeatedly add noise to the synthesized data to generate a set of simulated experimental results and apply the set of operations to that

artificial data. The results represent a distribution for each of the parameters from which an error matrix and tabulated parameter distributions can be extracted.

Error Matrix

The result of a Monte Carlo simulation for a peak fit takes the form of m optimization parameters from each of the n simulated data sets. An error matrix derived from these parameter distributions is defined to be an $m \times m$ matrix as follows:

$$e_{ij} = \frac{1}{n-m} \sum_{k=1}^n (x_{sik} - \hat{x}_i)(x_{sjk} - \hat{x}_j)$$

Where the x_{sik} are parameter values calculated for each of the simulation steps and each distribution is centered with respect to the mean rather than the initial parameter value.

The standard error in each parameter is given by $\sqrt{e_{ii}}$ and the correlation between parameters i and j is given by $\frac{e_{ij}}{\sqrt{e_{ii}e_{jj}}}$

An alternative method for estimating uncertainties in the peak parameters is to quote the inverse of the Hessian matrix used in the Marquardt Levenberg optimization routine. So the question is, why bother with Monte Carlo simulation when this very information desired is offered by the optimization procedure itself? The answer comes in two forms. Firstly, the error matrix derived from Monte Carlo is not limited to the problem of just fitting the peaks. The simulation procedure can also take account of the background as

well as the peaks, and so determine a global error matrix for the calculation rather than one that focuses on the stability of the peak parameters alone. Secondly, the Hessian matrix has a tendency to be near singular when the Marquardt algorithm terminates. Indeed the optimization routine follows a sequence of steps that switch between a direct solution of the Hessian matrix and a solution of a regularized form of the matrix. The regularization step occurs when the direct Hessian inversion fails to produce an improvement in the fitting parameters and a common reason for this to occur is that the Hessian is ill conditioned. Under these circumstances the covariance matrix derived from the Hessian should not be trusted.

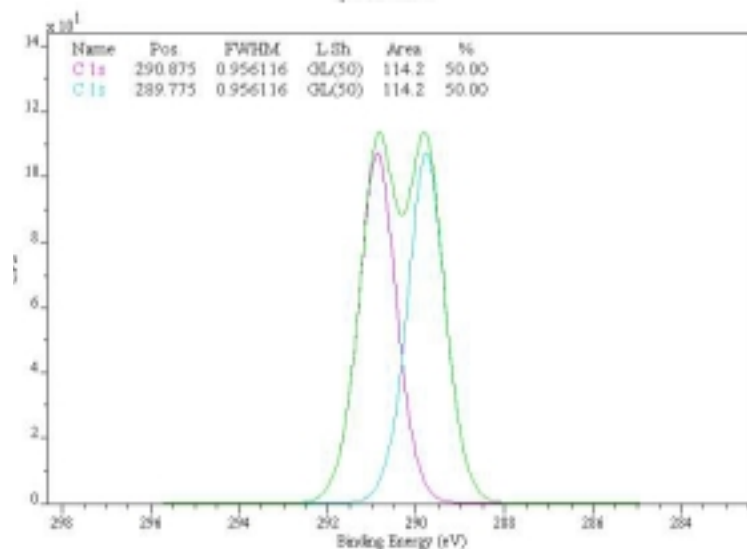


Figure 63: Simulated PVC C 1s Data Envelope.

Monte Carlo: A Simple Example

The best way to introduce the ideas behind Monte Carlo methods and how the results can be interpreted is to look at a simple example. (Note that the “Monte Carlo” button may be found on the Components tab of the Quantification module)

Consider the data envelope in Figure 63. The spectrum is a synthetic envelope created from two GL(50) line-shapes without any background and where the peaks are separated by an energy gap consistent with C 1s lines in a PVC spectrum (this is actually the PVC data available from the “Test” tab of the Processing module). In the absence of noise and experimental error the optimization routine always returns the peak parameters tabulated on the spectrum in Figure 63. The next step in the simulation is to introduce noise onto the data envelope that is consistent with noise found in experimental XPS spectra, i.e. variations about the channel intensity of magnitude related to the square root of the counts. Figure 64 shows the data envelope from Figure 63 after noise has been added and the peak parameters refitted.

This procedure yields the first set of simulation results. If repeated many times, the output is six distributions, one for each of the peak parameters involved and these can be used to assess the influence of experimental noise on these quantities. Note that this differs in some respects from adopting a purely experimentally determined parameter distribution. The initial starting point for the peak parameters will not be identical for an experimental data set since the experimental data may be subject to sample charging, and any errors in the

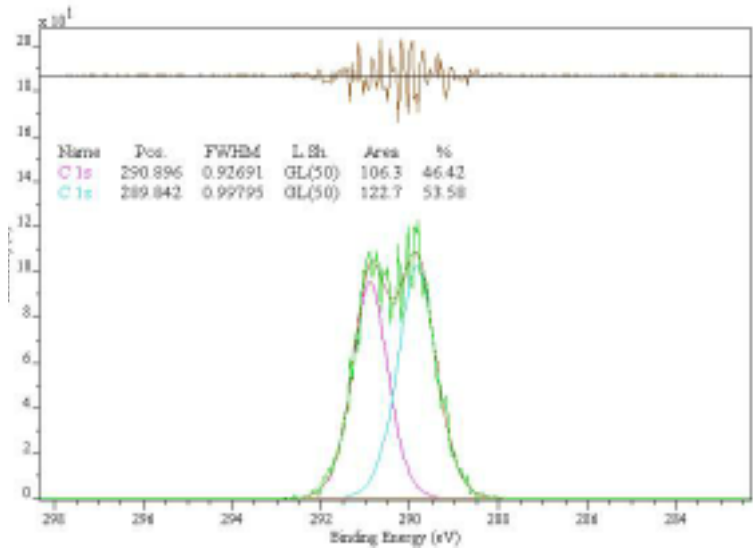


Figure 64: Simulated C 1s data where noise has been introduced to the system.

measurement procedure that can not simply be described by the random nature of the counting system will be omitted. A Monte Carlo simulated data set only tests the stability of a model with respect random noise and therefore may neglect other factors present in an experimental data set.

Table 10 is the error matrix that results from a Monte Carlo simulation for the spectrum and synthetic model in Figure 63.

If viewed as uncorrelated error distributions (which is certainly not true) then one standard deviation in each of the parameters is given by the square root of the diagonal elements in the matrix.

Table 10: Error Matrix for the C 1s spectrum in Figure 63.

	1:Area	1:Pos'n	1:FWHM	2:Area	2:Pos'n	2:FWHM
1:Area	39.230	0.122	0.196	-31.512	0.128	-0.186
1:Pos'n	0.122	0.001	0.001	-0.120	0.000	-0.001
1:FWHM	0.196	0.001	0.002	-0.185	0.001	-0.001
2:Area	-31.512	-0.120	-0.185	38.330	-0.125	0.202
2:Pos'n	0.128	0.000	0.001	-0.125	0.001	-0.001
2:FWHM	-0.186	-0.001	-0.001	0.202	-0.001	0.002

It may be seen that the areas of the two peaks are anti-correlated, where the quantity $-31.512 / \sqrt{(39.230 \times 38.330)}$ characterizes the degree of interaction between the two parameters. The minus sign indicates the parameters are anti-correlated. A scatter plot constructed from the two distributions for the peak areas is shown in Figure 65, where the anti-correlation is obvious from the bias in the scatter along a direction 135° to the positive x axis. An alternative scatter plot (Figure 66) for the area and FWHM parameters taken from peak C 1s 1 (Figure 63) shows that these two parameters exhibit positive correlation.

The uncertainty for correlated parameters can be estimated from plots such as those shown in Figure 65 and Figure 66. By projecting horizontal or vertical lines onto the parameter axes in such a way that the appropriate number of points lie between the projection lines for a given confidence limit, the uncertainty for a parameter can be assessed in the context of others. Again, the uncertainty for a parameter should be viewed in the context of all other correlated distributions, and yet the ellipsoid in 6-dimensional parameter space is difficult to quantify; the procedure based on scatter plots should be seen as merely a step in the right direction rather than arriving at the ultimate destination. In the case of the peak area shown in both Figure 65 and Figure 66, the estimate taken from the diagonal element of the error matrix would seem to be reasonable. Both scatter plots show that about 65% of the points lie inside projection lines positioned at about 0.95 and 1.05. This interval represents about $\pm 5\%$ of 114.2 CPS/eV (see Figure 63) and is not too different from the uncertainty taken from the error matrix ± 6.2 CPS/eV.

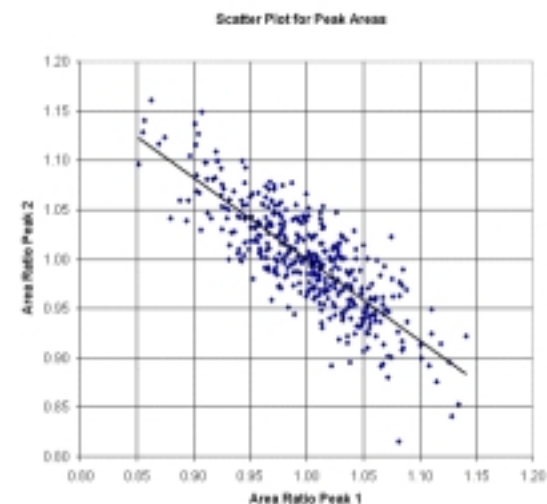


Figure 65: Scatter plot showing the anti-correlation between the peak area parameter distributions.

All the parameter distributions are reported relative to the initial parameters used to define the data envelope. As a result the relative parameter distributions may not be symmetrical about the initial value and asymmetric confidence intervals are possible. Note that the error matrix is calculated from the distributions centered on the mean parameter value for the distribution, not the initial values.

One of the real advantages of using Monte Carlo error analysis is that it highlights when a quantification parameter is poorly determined by the combination of model and optimization procedure. It also allows the influence of constraints within a model to be evaluated. Adding information about chemical shifts, relative peaks widths and/or peak areas can

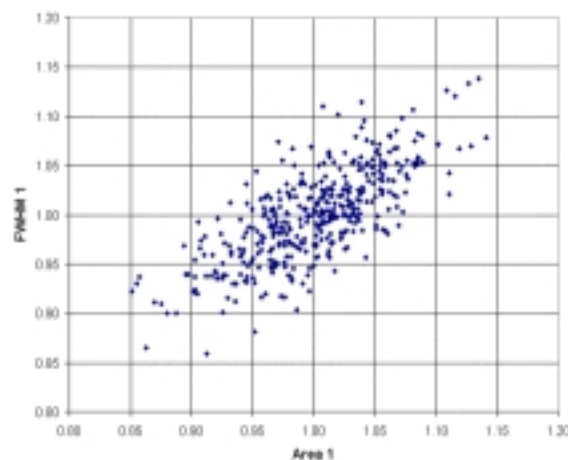


Figure 66: Scatter plot between Area and FWHM parameters for Peak C 1s 1 in Figure 65

alter the manner in which noise adjusts the parameters from their initial value. Monte Carlo derived scatter plots can often help to understand how rigid models based upon chemical knowledge can reduce the range of outcomes for a given set of parameters.

Quantification

Quantification of AES/XPS spectra is routinely performed

using combinations of intensities from integration regions and synthetic components. The purpose of this section is to point out some of the less obvious consequences that result from the specification of the background associated with the recorded peaks.

Monte Carlo End Point Determination

Consider an all too common situation where a pair of unresolved peaks is the subject of interest and the total intensity is the value that is calculated from the data.

The data in this example (Figure 67) are purely synthetic; two Gaussian peaks of equal intensity and equal FWHM of 2eV are offset by 1eV on a linear background. This provides an envelope open to analysis by Monte Carlo simulation that will provide visual insight into the errors associated with noise in the data.

Integration Region Limits

The purpose of this analysis is to illustrate the consequences of choosing inappropriate end points for the integration region. Note that the peak-fitting procedure is not at issue since only the data are used to calculate the intensities for the combined peaks.

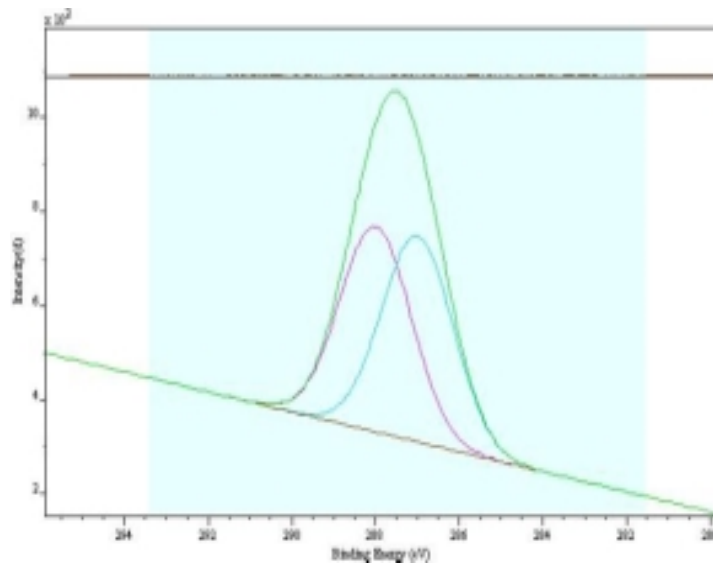


Figure 67: Two “unresolved” peaks

Figure 66 shows a scatter plot where each point is a normalized area plotted against simulation index. These areas are calculated from the original data envelope after the introduction of pseudorandom data and the computed area is referenced to the initial area before noise was superimposed on the spectrum

. Each calculation includes determining a new linear back-

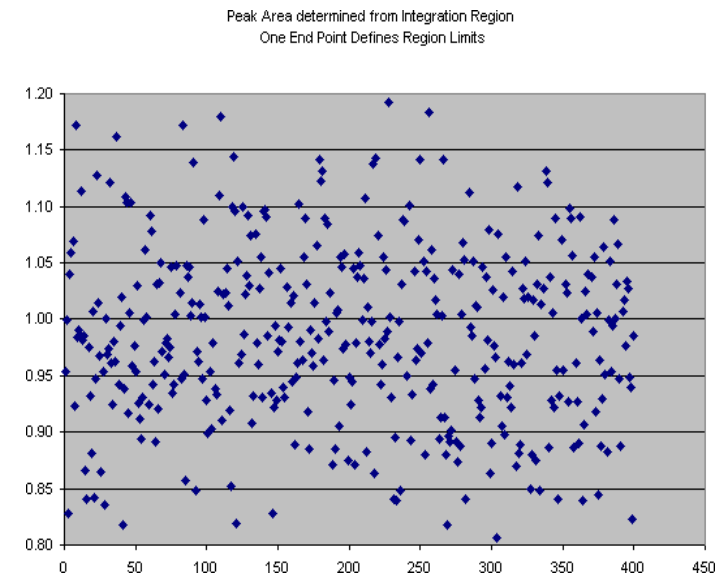


Figure 68: Peak areas from one end point

ground from the information in the simulated noise. In Figure 68 the background is computed using only one point at each end of the integration region while Figure 68 is the equivalent experiment but in this case an average of 21 points defines the background limits.

It is clear from the two scatter plots that the intensity calculation using just one data point at each end of the integration region is much more sensitive to noise. The introduction of more data channels when determining the background improves the situation.

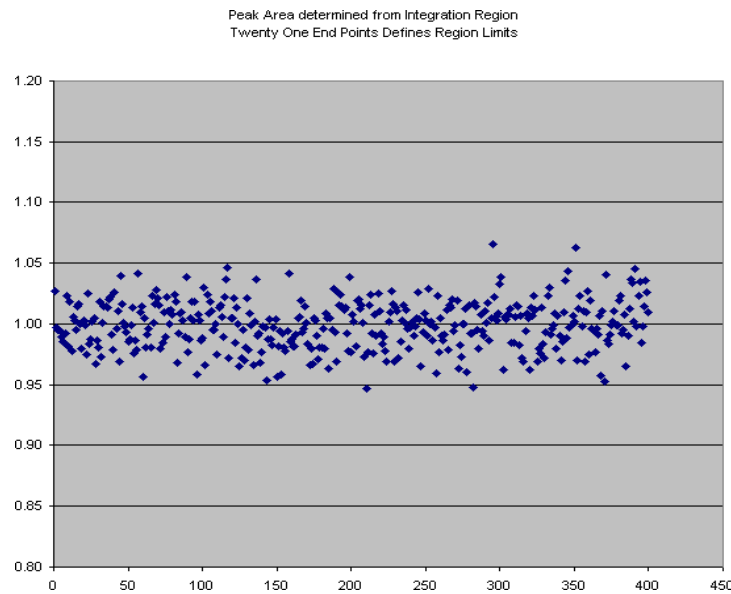


Figure 69: Peak areas - end points averaged

Not all peaks are isolated and so under these situations it falls to the acquisition phase to improve the statistics for the data channels that determine the background. An acquisition system that accurately determines the points used to define the background will yield intensities with significantly more precision than those that treat the end point as equal to any other in the acquisition region¹⁵.

15. K. Harrison and L.B. Hazell, *Surf. Interface Anal.* **18**, 368 (1992)

Intensities Determined by Peak-Fitting

Two peaks of equal width and area but separated by 1eV represent an intractable problem for a peak-fitting algorithm if the intensity of the individual peaks is required and no further information is available that can be used to provide constraints for the peak fitting parameters.

Figure 70 shows a Monte Carlo simulation for the peaks shown in Figure 67. Each Monte Carlo simulation step takes the envelope for the two peaks plus the background then superimposes a normal deviate of mean zero and standard deviation one scaled by the square root of the counts in each data channel. The envelope constructed in this way simu-

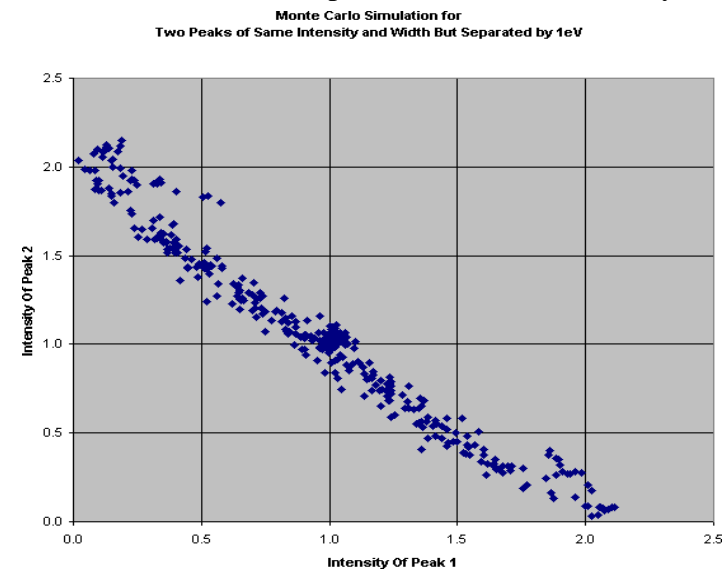


Figure 70: Linear background using single end points..

lates a sequence of experiments performed on the same sample, so that the only variation in the data is the random noise inherent in the acquisition process. For each Monte Carlo envelope the peak parameters are refitted. The points used in the scatter plot are determined from the intensities for the two peaks after automatically fitting the peak parameters (Area, FWHM and Position) to the new data envelope; these intensities are normalized with respect to the initial values for the peak areas. It is apparent from Figure 70 that the uncertainty in the individual peak intensities may be in error by as much as 100% although the total area is determined much more accurately (Figure 71). The scatter in the computed intensities in Figure 71 should be compared with that for the intensities shown in Figure 68.

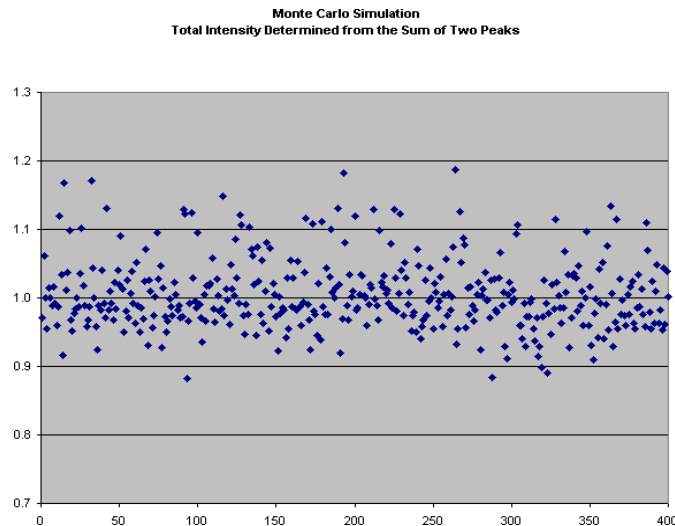


Figure 71: Total intensity, linear background from one end point.

In both cases the background used for the Monte Carlo simulations shown in Figure 68 and Figure 71 are determined using one end point for the integration region. The asymmetry in the distribution derives from the errors introduced by the background calculation that occurs prior to performing the peak fitting procedure. A comparison of the distributions suggests that the intensity determined from the peak fitting procedure will be more accurate than the integration region approach although there is a tendency to over estimate the total area when the sum of two fitted peak intensities is used and the background calculation is sensitive to noise. Again, the use of more acquisition channels to define the integration region end points improves the precision in the results for intensities calculated from the sum of the fitted peak areas (Figure 72).

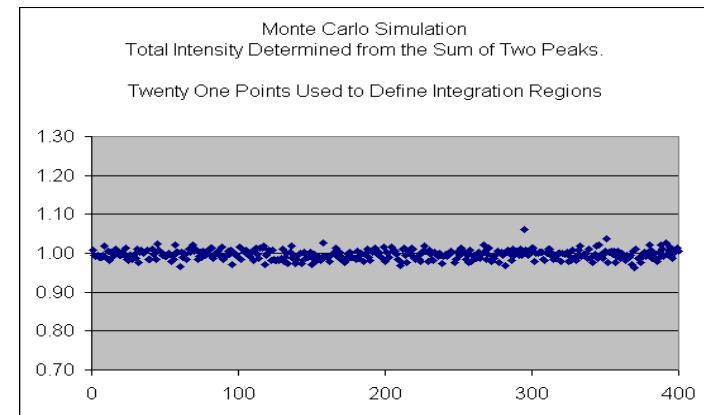


Figure 72: Linear background with twenty-one end points.

In addition, when the uncertainty in the background is reduced, asymmetry in the distribution of area ratios is im-

proved dramatically (compare Figure 71 and Figure 72).

Introducing additional information into the background calculation does not improve the uncertainty in the intensities for the individual peaks determined by a peak fitting procedure, but it does remove some of the correlations in the peak intensities due to the background. These correlations are visible in Figure 70 at the extremes of the distribution but are not present when the same experiment is performed but with an improved background specification (Figure 73).

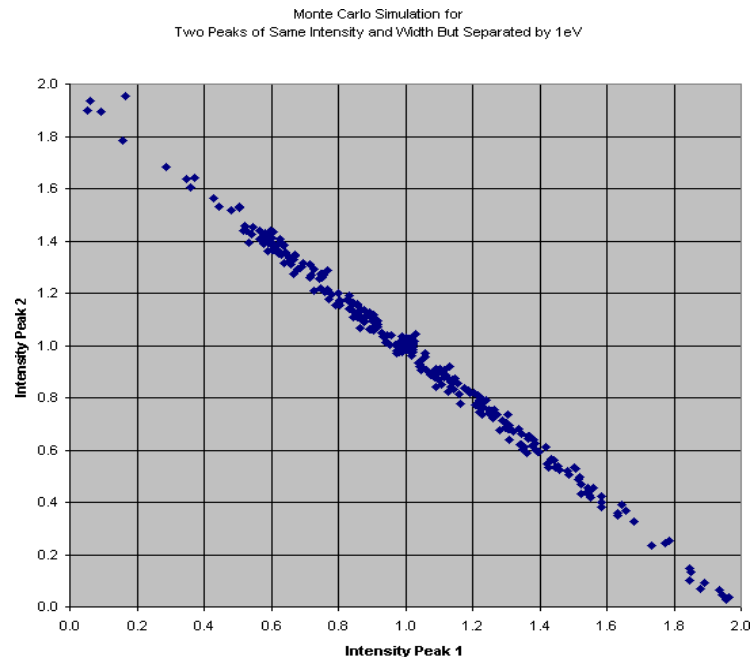


Figure 73: Individual intensities - Linear background with twenty-one end points.

Summary

Monte Carlo simulation provides a valuable tool for understanding the uncertainties associated with data reduction for AES/XPS spectra. The results presented here are available in other forms ^[15] ^[16] ^[17] however the visual feedback offered by Monte Carlo simulations ^[17] provides an insight for analysts who feel less comfortable with a mathematical description of the same concepts. By using Monte Carlo simulation, an understanding of the errors involved in extracting information from AES/XPS spectra is hopefully opened to a wider audience.

16. P.J. Cumpson and M.P. Seah, "Random Uncertainties in AES and XPS: Peak Energies, Areas and Quantification." NPL Report DMM(A) 26, May 1991; *Surf. Interface Anal.* **18**, 345 (1992) and 18, 361 (1992).

17. S. Evans, *Surf. Interface Anal.* **18**, 323 (1992).

The following sections provide examples of the treatment of particular analytical problems in greater depth, emphasising the special tools and procedures available within CasaXPS as appropriate. It is intended that these examples should provide a greater insight into some of the powerful features of the system and an indication of a possible analytical approach which may be applied effectively to certain problems. It is not suggested that these are the only approaches possible or available - even within CasaXPS - but merely that these indicate known, good methodology.

Organic Polymers and Curve Fitting

The availability of data from well-characterized samples such as those offered by Beamson and Briggs¹⁸ owe much to the popularity of XPS as a tool for understanding the chemistry of polymers. A typical C 1s envelope¹⁹ (Figure

18. Beamson G. and Briggs D., "The XPS of Polymers Database - CD-ROM", Surface Spectra Ltd (2000)

74) includes structure that offers chemical information about a sample, but without some initial starting point it is difficult to construct an appropriate model for the data envelope.

The high-resolution spectrum in Figure 74 derives from poly (acrylic acid) PAA reacted with inorganic material (or partially reacted) in an acid-base reaction. If Gaussian/Lorentzian (GL) line-shapes are added in an arbitrary way, the curve fit yields little information about the sample other than to say that it deviates from the published data for PAA and therefore demonstrates the presence of additional chemistry at the surface. Figure 75 shows a synthetic model for this data envelope where three GL peaks have been added, then fitted using a Marquardt-Levenberg²⁰ optimization algorithm. No constraints have been applied and the result is a reasonable fit to the experimental data but the fitting param-

19. Jones F. et al, "Fluoride uptake by glass ionomer cements: a surface analysis approach", *to be published*

20. Press W.H. et al, "Numerical Recipes in C", Cambridge University Press (1988)

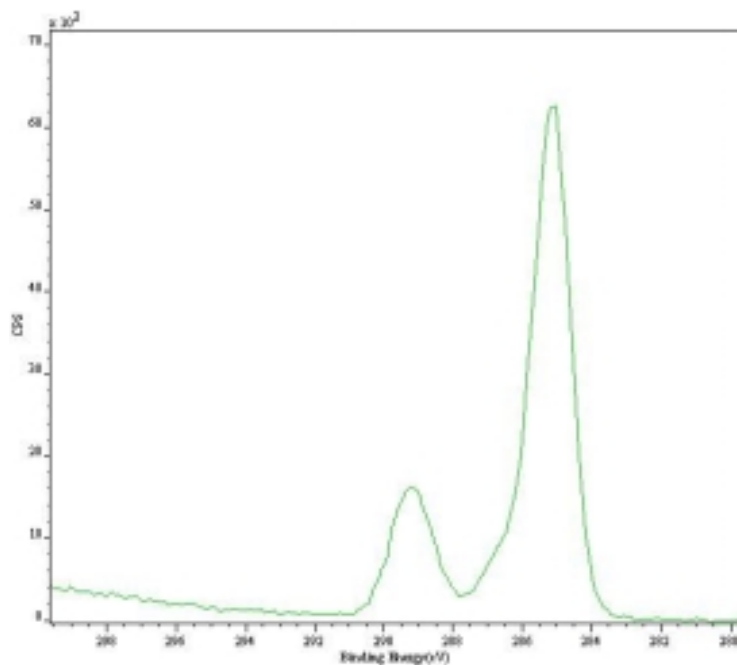


Figure 74: C1s high-resolution spectrum taken from a polymer-based sample using an ESCALAB 220i at University College, London.

eters are not readily open to chemical interpretation. For example, the FWHM are much bigger than would typically be expected from C 1s profiles the given instrumental resolution. The three-peak model suggests further deconvolution is required before the sample can be fully understood.

If it is assumed that the FWHM for a C 1s photoelectron peak is 1.1 eV (only a guess), then applying peaks with said constraint results in a peak fit shown in Figure 76. Two additional peaks appear in the model and further more three of the peaks look like they may have something to do with pure

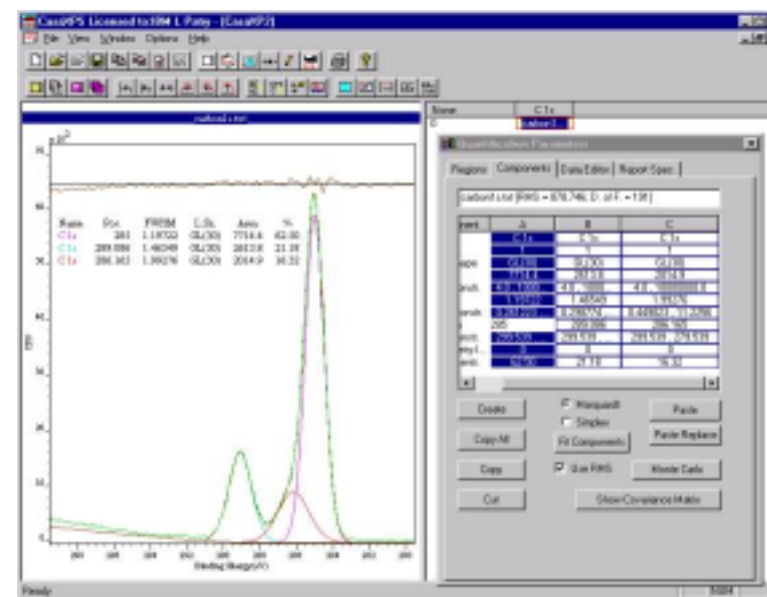


Figure 75: An initial fit to the C 1s spectra shown in Figure 74

PAA. The PAA stoichiometry is still doubtful but the essential positions for a pure PAA envelope (as indicated by Beamson and Briggs) have more or less appeared. Further input is required to make sense of the new synthetic model.

To constrain a parameter so that it does not adjust during an optimization step, set the constraint interval to have the same value as the fixed value: e.g., if it is required to set the FWHM value to 1.1, the constraint interval should be entered with the value "1.1,1.1". Actually, it is sufficient to set the parameter value outside the constraint range currently defined for the parameter.

Spectra contained in the Beamson and Briggs polymer database offer the opportunity to examine more complex polymer data in the context of known synthetic models. Figure 77 is an example of such data where a set of three line-shapes has been used to model the clean PAA C 1s data envelope. The important feature is that the stoichiometry and chemical shifts for the C 1s lines are incorporated into the model and this information is then transferable to other polymer spectra. The peaks in Figure 77 are linked in area, but only the position of peak “C 1s b” and “C 1s a” are constrained by an offset.

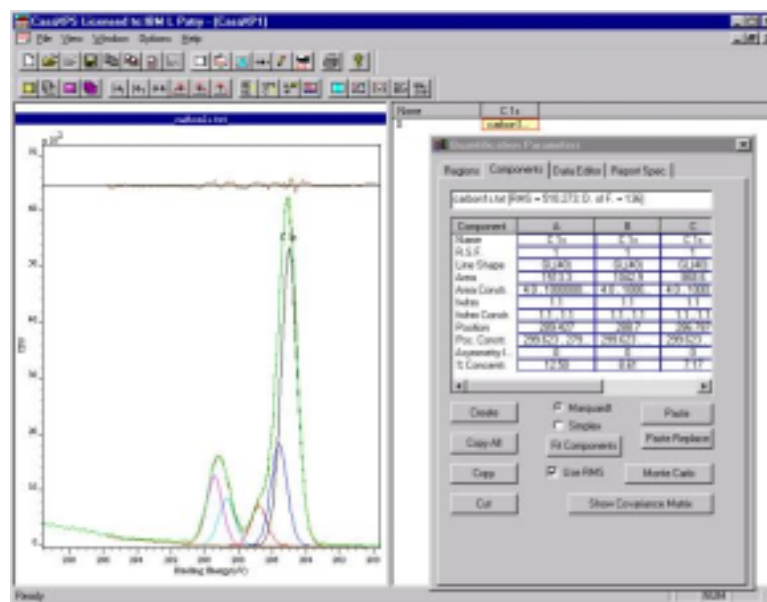


Figure 76: Same C 1s envelope as Figure 75 but the synthetic peaks are all constrained to have FWHM equal to 1.1 eV.

To link a component parameter the constraints must be adjusted as follows. Each synthetic component defined on the Quantification Parameter dialog, Components Property Page, appears as a column of parameters in the scrolled list shown in Figure 76. These columns are headed “A”, “B”, “C” and so on. To constrain the area of the component in column B to be half of the area of the component in column A, the area constraint in column B should be set to “A * 0.5”. Similarly, to offset a component in column C by 0.2 from the component in column B, enter “B+0.2” in the position constraint field in column C.

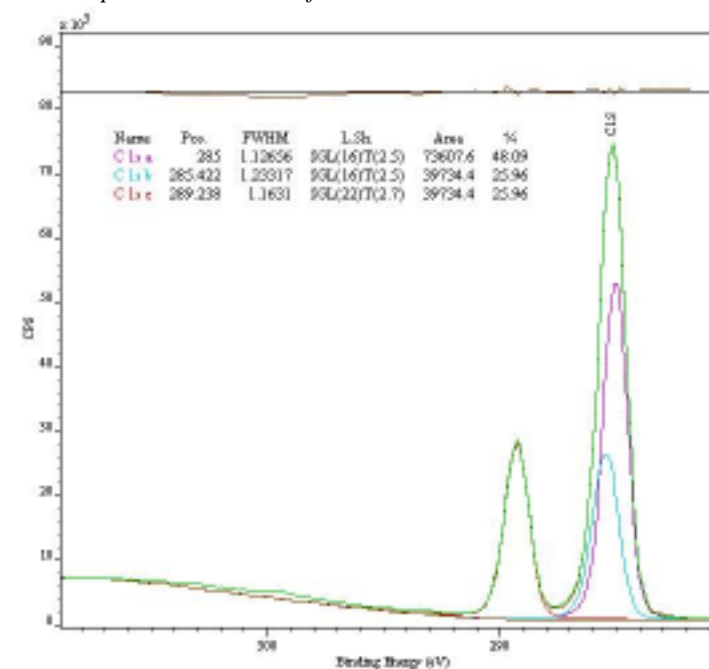


Figure 77: C 1s envelope from clean PAA acquired on a Scienta-300, RUSTI, Daresbury Laboratory UK.

The peak shapes from a Scienta ESCA-300 may differ from a VG ESCALAB 220i (the source of the real data in Figure 74) or any other manufacturer's instrument, but the essential structure should be suitable as a basis for the new model. Copying the Beamson and Briggs pure PAA model into the data in Figure 76 leaves a residual that requires an adjustment for the two non-PAA peaks together with the introduction of a third peak. The new peak in Figure 78 is constrained to be the same width and position as the saturated PAA C 1s peak located at 285 eV (BE). The area of this new peak is allowed to adjust at will and accounts for carbon with the same characteristics as the PAA peak at that position. The consequence of introducing the new peak is that the PAA synthetic model can adjust without breaking the stoichiometric relationships for pure PAA, while differences in the intensity of the saturated peak from the PAA structure are allowed for by this additional constrained component (Figure 78).

Synthetic models such as the one in Figure 78 can be tested using Monte Carlo simulation techniques²¹ to assess the stability of the peak parameters with respect to noise in the data. Once a set of peaks and constraints has been developed, the Monte Carlo procedure simulates repeated identical experiments on the same sample, and for each simulation a new fit is determined for the peak parameters. The result of this procedure is a set of distributions for the individual parameters from which scatter plots may be constructed that

21. Cumpson P. J. and Seah M. P., "Random Uncertainties in AES and XPS", *Surface and Interface Analysis*, **18** 361 (1992)

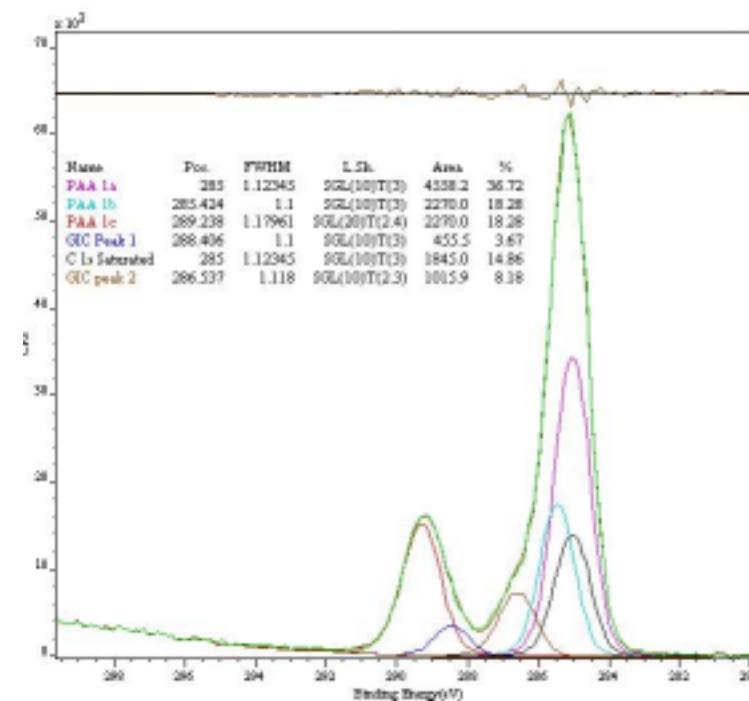


Figure 78: Final form for the synthetic model. The Glass Ionomer Cement (GIC) [2] C 1s envelope containing three peaks from PAA plus three additional peaks not seen in a clean PAA spectrum.

highlight the way noise influences the data model. Figure 79 is a scatter plot for the normalized peak areas of the non-PAA peaks against the saturated C 1s peak from the pure PAA model. The peaks at 288.4eV and 285eV have areas that are anti-correlated with the pure PAA model (all the peak areas within the PAA model are constrained to one another,) but the synthetic component at 286.54eV shows a degree of correlation with the intensity of the pure-PAA area.

At first glance the behavior of the GIC 2 peak at 286.54eV is unexpected since one might think that two peaks next to one another should produce anti-correlated area distributions. The constraints have altered the concept of “next to” since the PAA sub-model spreads across the entire envelope and it becomes difficult to judge by eye what the influence of noise might be on the final result. This type of insight can only help to understand what constraints do to a fitting procedure as well as provide a rule-of-thumb estimate for error bars (multi-dimensional error distributions can seldom be described by a single number.)

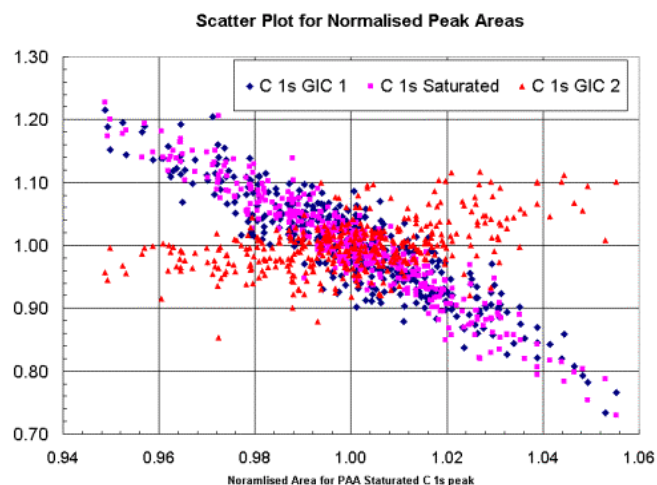


Figure 79: Monte Carlo simulation results for normalized peak areas. The three peaks associated with GIC are plotted against the saturated C 1s peak area from the pure PAA model

The exact meaning for the model in Figure 78 is left to the

experimentalist and may need changing in view of other input, however such a model is only possible when peak fitting routines offer mechanisms for fixing parameters with respect to one another. The role played by the pure PAA synthetic model is that of a foundation shape from which differences in the unknown polymer can be assessed. These additional peaks may still require further interpretation, but with the aid of chemical knowledge and supporting evidence a meaningful model can emerge from seemingly intractable data.

Quantification using “Tagged” Regions

Introduction

Routine measurements using XPS often involve acquiring a wide scan spectrum to determine the general composition of a sample. The survey spectrum is then used to select narrow energy regions where detailed structures are present. These narrow scan spectra are typically needed when overlapping peaks are responsible for the data envelope such as is often seen in the case of C 1s spectra. Synthetic line-shapes must be used to extract the chemical-state information in the data, but intensities determined from these models are not always comparable with intensities determined from integration (wide scan) regions. The most notable case is when asymmetric line-shapes are used to model the data. Moreover, the transmission characteristics of an instrument require that each operating mode must be characterized by a transmis-

sion function and although it is possible to correct the data for such differences, problems may arise due to changes in the system from aging or tweaks to lens functions by the operator. Both require recalibration of the instrument, without which peak areas from a wide scan acquired using one pass energy cannot be used with intensities measured via other modes.

If a spectrometer is characterized by a set of relative sensitivity factors (RSF's) applicable only to the survey mode of an instrument, but the resolution of the survey mode is insufficient to provide the desired chemical state information, then one way to provide a more detailed quantification is to reference the intensities from the to high-resolution spectra to this survey mode. Peak areas determined from the high-resolution data provide the relative proportions for the chemical-state intensities and therefore the corresponding concentration from the survey data can be subdivided using these proportions, thus allowing detailed quantifications without the need for transmission correction. Such a procedure effectively takes out the transmission function from the quantification step and lessens the need for the time consuming calibration of each operating mode. Indeed, for many instruments transmission function correction is not available and so this procedure represents the only way to combine data from different operating modes, which would otherwise require the maintaining of multiple sets of RSF values.

“Tagging” regions in a survey scan

Figure 80 shows a typical situation from an XPS measurement where a survey spectrum provides the overall elemental quantification but a narrow C 1s scan offers a wealth of

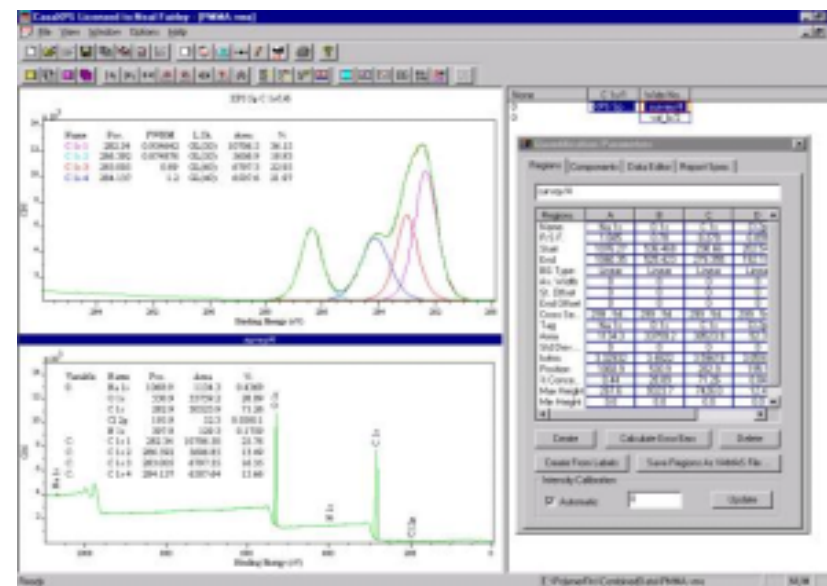


Figure 80: C 1s high-resolution data used to proportion the quantification for C 1s in the survey spectrum.

information lost by the limited resolution used for the survey spectrum. In the case of relatively pure samples, such as the one in Figure 80, it might be possible to run a set of high-resolution spectra for each of the peaks identified from the survey data, but many samples include so many elements that full quantification via high-resolution spectra would be both time consuming and costly. Degradation of a sample from exposure to X-rays and financial considerations can make the use of high-resolution spectra unattractive as a routine analysis regime.

The data in Figure 80 are quantified via a table that shows the elemental concentrations plus a further breakdown of the

C 1s elemental composition into chemical state concentrations. The proportions determined from the peak model for the high-resolution envelope are used to show how the elemental concentration is subdivided into chemical-state intensities. This is achieved by assigning a “tag” to each of the quantification items used in the analysis. That is to say, in addition to the name field in each integration region or synthetic component there is a Tag entry (in the form of a string), which is used to link intensities from synthetic peaks to a specific integration region.

If the information is available, CasaXPS will enter into the Tag field the element and transition associated with the acquisition, and only those items with tags are included in the quantification table. To remove an item from the quantification step delete the string from the Tag field and press return. The system will enter the key word NoTag when return is pressed to indicate that the item is not to be used in the quantification. Note that the NoTag entry is necessary to exclude from the elemental concentration table the integration region defined for the high-resolution scan seen in Figure 80. The background to the peak fit for the C 1s spectrum requires a region to be defined, but the intensity from this region must not be included in the results from the survey spectrum. In this example the tag field in the region used to define the background to the peak fit must be set to NoTag, while each of the synthetic components are tagged with the same name as the corresponding region in the survey quantification, namely, “C 1s”.

Not all forms of quantification tables use tags. For example, the Component annotation option used to display the component table in the upper spectrum does not discriminate be-

tween the components based upon tags, nor do any of the annotation options involving quantification tables unless the “Use Tag Field” checkbox is ticked. The tag mechanism is, however, always used when the Combined button on the Report Spec property page from the Quantification Parameters dialog window is selected. No other quantification reporting options use tags.

“Tagging” for different line shapes

Although the primary reason for using tagged reports is to reference results from peak fits to survey spectra, it is possible to use the same mechanism to remove some of the ambiguities from comparisons between intensities derived from peak fits that employ different line-shapes. Any line-shape involves a functional form that may extend beyond the acquisition regions to which it is applied, especially when asymmetric peaks are involved. The tag mechanism allows the peak fitted results to be referenced to the underlying integration region and so intensities for the chemical states are calculated based upon the data rather than the implementation of the synthetic components.

Figure 81 shows a set of quantification tables. These tables derive from components and integration regions taken from the two high-resolution scans below the survey spectrum. Note that the Al 2p doublet is fitted using a Doniach-Sunjic asymmetric line-shape, while the O 1s spectrum is modeled using symmetric Voigt approximations. The table headed “Components Only” shows the atomic concentrations where the intensities are calculated from the line-shapes. These concentrations do not agree with the elemental results calculated from the integration regions alone, where the Al 2p

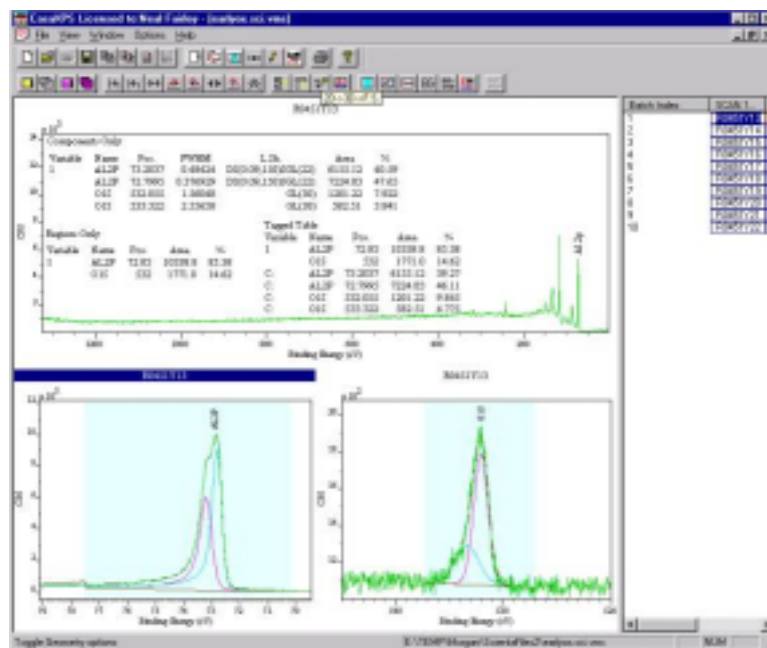


Figure 81: Tag mechanism applied to differing line-shapes. The Al 2p has been modeled using Doniach-Sunjic profiles, while the O 1s spectrum is fitted with Gaussian-Lorentzian line-shapes.

concentration is computed to be 85.38% while components only reports 88.24% for the same quantity. The differences are due to the cutoff criterion used to limit the infinite area under a Doniach-Sunjic profile. Using the tag mechanism the peak fit can be used to subdivide the concentrations from the integration regions using the individual components from the peak models. The concentrations for the synthetic components are now consistent with the results of the integration regions but are in the proportions determined by the

models.

Trend Analysis for Metal Oxidation

Introduction

Many XPS/AES experiments are performed to monitor how the chemical composition of a sample varies with one or more variables. These variables may be time, tilt angle of a sample (depth), ion gun etch-time or any other quantity that characterises the state of an experiment. The challenge is to devise methods for identifying changes in the spectra and therefore provide a way of following these changes as a function of the experimental variable.

The usual tools available for reducing XPS/AES spectra are simple quantification regions and/or synthetic spectra determined by optimisation procedures with various forms of parameter constraints. Implicit in both of these data reduction methods is the introduction of a background approximation and in the latter case rigid models for the synthetic line-shapes. The consequence of the assumptions involved in these data reduction techniques is hard to assess with respect to the results for individual spectra. Techniques for estimating the errors due to noise in the data are available for both peak-fitted parameters and quantification region values, but there is still the question regarding what influence these assumptions have on the trend itself.

An alternative means of following a trend through a data set is to perform a Principal Component Analysis (PCA). This technique is a linear least square procedure that transforms a

set of spectra (when viewed as vectors) into a set of orthogonal basis vectors (abstract factors) that span the same subspace as the original data set. Trends within the original spectra can be assessed by examining the so-called abstract factors and plotting the co-ordinates of the spectra with respect to the principal abstract factors as a function of the experimental variable involved. Here the term principal abstract factor means those abstract factors that contain significant structure and are necessary for providing an adequate description for the entire set of spectra. Plotting the co-ordinates for the spectra with respect to the subspace spanned by the principal components shows how each of these designated abstract factors contributes to the description of the spectra and depending on the nature of the abstract factor trends can be identified. No a priori model is required by PCA although pre-processing the spectra can be used to assess the consequences of including, for example, a particular background approximation.

Principal Component Analysis is not a substitute for the conventional quantification techniques and requires care before the results can be used to assess trends. There have been attempts to place physical meaning on the abstract factors but it should always be remembered that the abstract factors are truly abstract and owe their nature and size to a mathematical procedure for generating an orthogonal basis set for an n-dimensional vector space. The first abstract factor is chosen to be in the direction of the maximum variation in the entire set of spectra in the linear least square sense. Subsequent factors are generated with the same criterion as the first but subject to the constraint that each new vector is orthogonal to all those that have already been constructed. Sometimes a

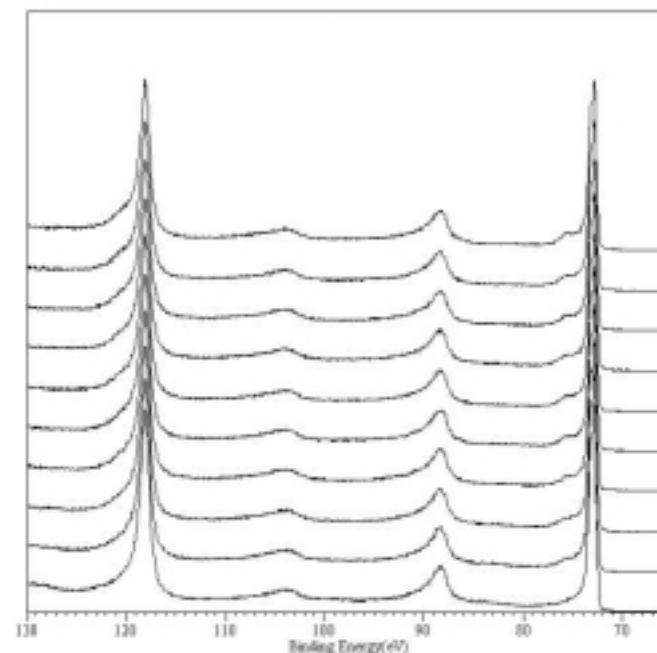


Figure 82: Oxidation Sequence

data set may have a dominant trend and therefore the abstract factors apparently have chemical meaning in their own right. The temptation to interpret these abstract factors as chemically meaningful quantities should be avoided nor should the number of significant factors be slavishly used to imply the number of synthetic components that are required to model a data envelope. However PCA can produce some interesting insights into the data and with the appropriate care provides supporting evidence for a synthetic model.

A trend analysis for the oxidation of a piece of aluminium

provides a good example of how PCA can assist in the interpretation of a data set. The essential problem is to construct a synthetic model for the Al 2p or Al 2s photoelectric lines measured periodically as an Aluminium sample oxidises in a vacuum chamber. The relative sizes for the individual synthetic lines provide a quantitative measure for the trends within the data. The problem lies in knowing what chemical states are present and therefore what features the model must accommodate.

PCA for Aluminium Oxidation Sequence

The spectra shown in Figure 82 were recorded on a Scienta ESCA300 XPS spectrometer at RUSTI and represent a trend observed as a clean Aluminium sample, created by argon-ion etching, oxidised in a vacuum chamber as a result of combination with low level residual water and oxygen (1×10^{-9} mbar). These are particularly good data for a PCA since the spectra have been recorded over the Al 2s as well as the Al 2p photoelectric lines and both structures should vary together as a unit over the course of the experiment.

A PCA for the data set shown in Figure 82 using correlation about the origin produces the abstract factors and associated eigenvalues displayed in Figure 83. The first abstract factor is clearly a vector in the direction of the maximum variation in the overall data set. There are shapes within this first abstract factor that belong to what can be identified as aluminium bonded with oxygen. The second abstract factor exhibits structure associated with each of the aluminium peaks and what is more the structure is of similar form.

The third abstract factor contains almost no significant structure and all the PCA statistics point to the need for two

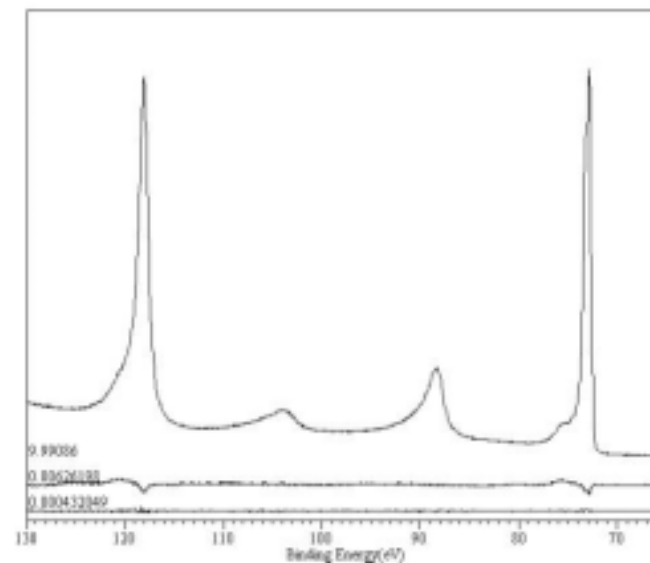


Figure 83: First three abstract factors generated from the spectra shown in Figure 82.

principal components when reproducing the original spectra. The trends within the data set can therefore be assessed by plotting the co-ordinates for the spectra when projected onto the two dimensional subspace spanned by the first two abstract factors.

At this point it would be easy to conclude that any synthetic model for the data envelopes requires two fundamental line shapes. The structure associated with the aluminium peaks is positive for the region where the oxide peak should occur and is negative in the region of the main metal line shapes. The second abstract factor represents an adjustment to the

average shape manifested in the first abstract factor that is required to describe each spectrum in the data set to within experimental error. The loading for each of the abstract factors demonstrates that the average form is almost constant across the data set but that the adjustment smoothly moves from a negative contribution for the second abstract factor to a positive loading as the oxidation proceeds. An interpretation for this trend might be the consistent conversion of one form of Aluminium to another.

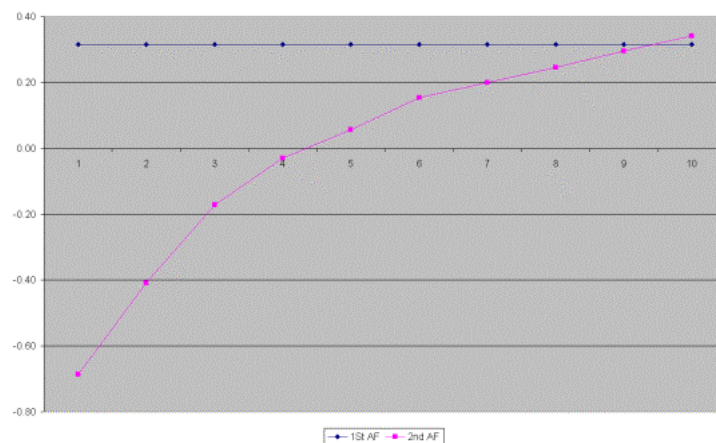


Figure 84: PCA loading for first two abstract factors

The graph in Figure 84. shows a plot for the loading associated with each of the first two abstract factors. The loading factor for the overall shape seen in the first abstract factor is virtually constant across the set of spectra. All the variation within the data set is accounted for by the second abstract factor. This is an interesting result in the context of the structure seen in the second abstract factor. The initial negative

loading represents a removal of the oxide shape from the average description represented by the first abstract factor, while the metal line shape is enhanced by virtue of the negative structure beneath that part of the spectrum. The adjustment to the average shape reverses as the oxidation reaction progresses. The PCA has managed to describe the trend using essentially a single loading and a careful choice of the basis vectors. An interpretation for this trend could be a migration of the metal form of aluminium to aluminium bonded with oxygen and/or the correlated attenuation of the metal signal due to the build up of an oxide over-layer. A strong conclusion from the PCA is that, whatever processes are involved, these processes occur in a linearly related fashion.

This statement does not exclude a hydride moving to an oxide state with the same speed as the metal making the same adjustment. It does, however, show that metal and oxide line-shapes are sufficient to describe the oxidation process measured by the XPS spectra.

The next step is to construct a synthetic model that produces the same trend as identified by the PCA.

Synthetic Model for an Oxidation Sequence

Photoelectric peaks for metals require synthetic models capable of describing asymmetry in the observed line shapes. Various approximations have been introduced to account for the asymmetry seen in the aluminium lines, ranging from ad hoc functions to theoretically based shapes. As always, regardless of the asymmetry function a background approximation must be introduced to facilitate the optimisation procedures that ultimately determine the peak parameters.

The Al 2s line is an interesting one to study since the data envelopes should result from only two photoelectric lines namely the metal and the oxide Al 2s resonance features (although a small plasmon structure from the Al 2p line should also appear at the same BE as the Al 2s lines.) This is in contrast to the Al 2p data envelope since in this case the data envelope is constructed from doublet structures that therefore complicate the nature of the model. The data set under study offers a near pure Al 2s Metal spectrum from which the asymmetry can be assessed and then applied to the subsequent spectra in the oxidation sequence.

The theoretical line shape for a metal such as the aluminium has been shown to have an asymmetric line shape given by Doniach and Sunjic²², however from a practical perspective the Doniach Sunjic line shape presents a number of problems. The principal problem is the question of area under the curve and how it relates to the intensity of the photoelectric line. The Doniach Sunjic line provides a very good fit for the observed data provided the background does not attempt to model the background contribution from the Al 2s line itself. The functional form for the Doniach Sunjic line typically extends beyond the data region and therefore integrates to a larger intensity than is calculated from the data region in question. Truncating the function to the integration region destroys stoichiometric relationships between peaks or relative intensities for multiple-peak models. The underlying line shapes must extend far enough to reduce the influence of offset peaks within the same region but whatever choice is made for the extent of the Doniach Sunjic line shape the

22.

intensity will not be consistent with the direct integration over the data region. If the data region is used to determine the limits for the theoretical line shape then the results for the intensity will vary with the acquisition region. Alternatively, a more extensive but arbitrary limit will result in more repeatable results for intensity values however these intensities will not compare to quantification values derived from different synthetic models.

The Voigt type synthetic line shapes are more ad hoc in nature than the Doniach Sunjic approximation but have the merit that they make no attempt to model the background resulting from the primary peak. The form for the background must derive from one of the standard approximations normally applied (Linear, Shirley, Tougaard etc). Modifying the Voigt function by an exponential tail revisits the problem associated with the meaning of the intensity when compared to the actual acquisition region (i.e infinite extent vs defined region). The grey area introduced by modelling data with functional forms that blur the boundaries between peak and background are also present when line shapes are extracted from experimental data as well as issues associated with arbitrarily truncating the model data. Good statistics for a model do not necessarily translate into good quantification results for practical reasons.

Peak Fit Using Doniach Sunjic Line Shapes

The preceding discussion attempts to highlight the dangers of modelling spectroscopic data using synthetic line shapes. Nevertheless trends within data such as the set shown in Figure 85 must be assessed by some means and an approximation chosen. In the case of the Al 2s peaks the Doniach

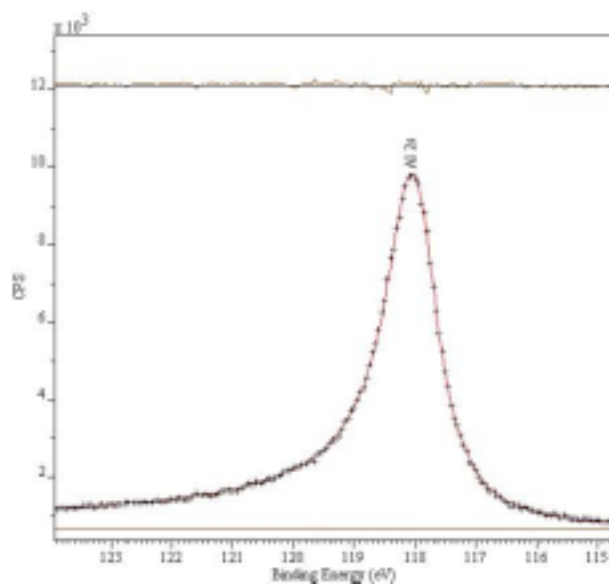


Figure 85: Al 2s metal data fitting using a Doniach Sunjic line shape.

Sunjic line shape provides the best form for modelling the observed data. Given these doubts it is therefore wise to test the results against other techniques in order to build confidence in the conclusions. In this case PCA is used to “cross check”

Figure 85 shows the initial spectrum in the data set fitted using a single Doniach Sunjic line shape. Note how the background approximation is a simple linear form offset from the spectrum.

The model from the first spectrum in the sequence is used as the basis for modelling the remaining spectra. An additional

component is introduced to account for the oxide structure seen on the other spectra. The position and width for the oxide component is determined from the last spectrum in the data set (Figure 86) and then fitted across all the spectra in the experiment.

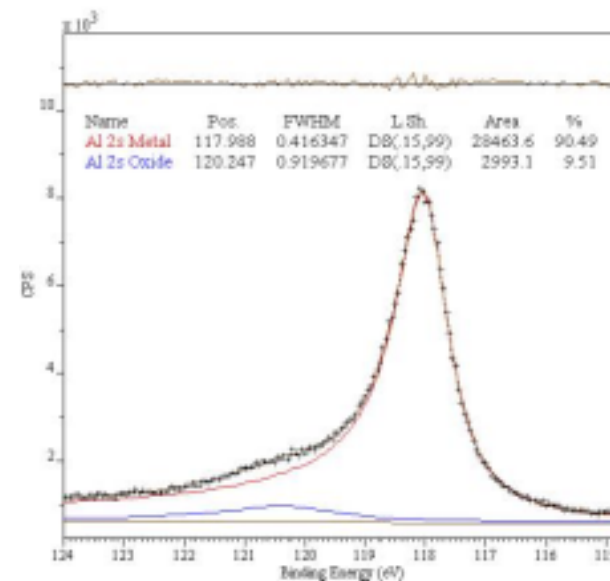


Figure 86: Oxidised Aluminium data envelope. Note that the oxide line-shape has been chosen to be the same as the metal even though the asymmetry of the metal should not exist in the oxide. This has been done to so that the intensity measured by the oxide synthetic line-shape is, in some sense, comparable to the metal.

A trend for the oxide peak intensity can be identified but the true intensities involved are difficult to assess for the reasons given above. Figure 87 shows a comparison between the results from modelling the spectra using two synthetic compo-

nents and the principal component analysis. The two PCA trends correspond to covariance about the origin with (PCA BG) and without (PCA) background subtraction. Note how the trend PCA BG contains elements from both the PCA and the normalised peak-fitted (Normalised PF) results. Clearly the background approximation influences the nature of the trend computed using PCA and therefore similar influences must also be present when peak fitting is used to extract the trend.

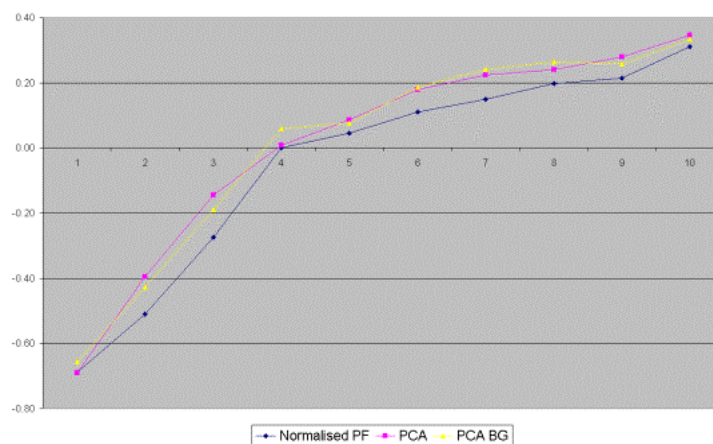


Figure 87: Comparison of trends identified using Curve-fitting and PCA.

The PCA without background subtraction shows the smoothest trend in Figure 87 and the limited input assumptions suggests this curve is the best description of the trend within the data set, although not necessarily the best descrip-

tion of the chemical processes involved. The trend analysis based on PCA is modelling changes in both the background and the intensity of the photoelectric lines, but does show that these adjustments throughout the data set are well modelled by two underlying shapes in the form of abstract factors.

Adjusting a Custom Quantification Report

The standard quantification reports treat each quantification item as a separate entity and therefore a percentage concentration is reported for each item used to quantify a sample. Custom reports, on the other-hand, sum the intensities for any quantification items defined with the same name. It is therefore important when using the Custom report option to label synthetic peaks with a different name from the region used to define the background for those peaks.

Consider the following example:

Figure 88 shows three high-resolution spectra used to quantify a sample. The O 1s region is fitted using two synthetic components, whilst the C 1s and the Al 2p spectra are quantified using integration regions only.

Each spectrum has an integration region defined and the names for these regions are O 1s, Al 2p and C 1s. In addition to the integration regions, the O 1s spectrum is fitted using two synthetic components and both of these components are named using the same name as the integration region for the O 1s spectrum. A Standard report generated from these inte-

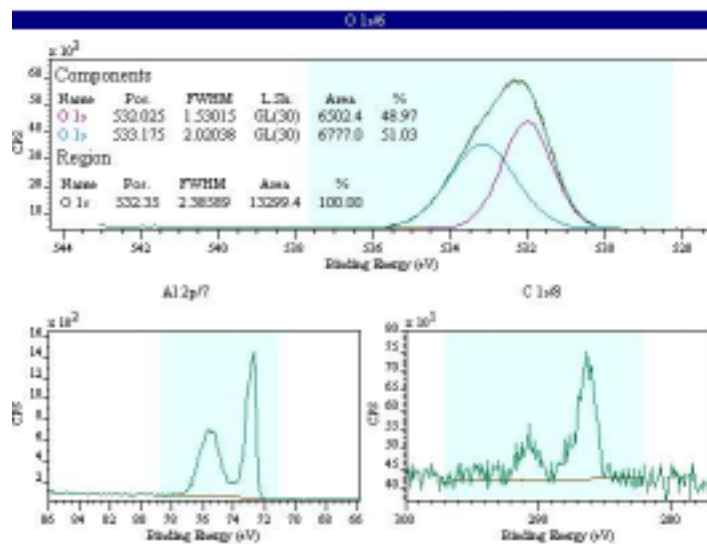


Figure 88: Three High-resolution Spectral Regions

gration regions is as follows:.

Table 11: Standard report generated from the quantification regions shown in Figure 88

Name	Energy	FWHM	R.S.F.	Area	% Conc
O 1s	532.35	2.38589	0.78	13299.4	54.52
Al 2p	72.7	0.740902	0.193	2192.6	36.33
C 1s	286.3	1.39428	0.278	795.3	9.148

A custom report using the same names as those shown in Table 11 is as follows:

Table 12: Custom report using the same names as those used in Table 11.

O 1s CPSeV	Al 2p CPSeV	C 1s CPSeV	O 1s %	Al 2p %	C 1s %
34075.5	11360.8	2860.9	70.5538	23.5227	5.92353

Note that the O 1s percentage concentration is larger than the equivalent value reported by the Standard report. The

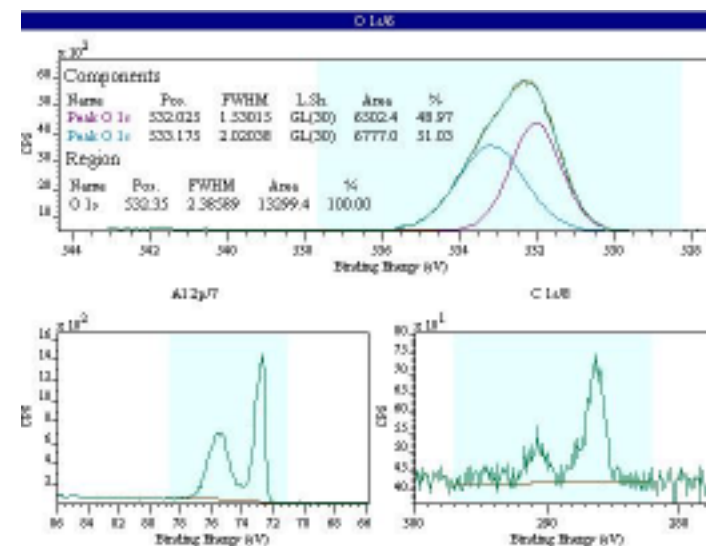


Figure 89: Synthetic peaks for the O 1s spectrum are now named differently from the integration region.

problem lies with the same name being used for more than one quantification item. The integration region for the O 1s spectrum has the same name as both the synthetic components and therefore the three quantities associated with the O 1s spectrum are summed to provide the intensity used in the custom report.

This was not the intention. To prevent this type of ambiguity it is best to use a different name for the integration region from that used for the synthetic components. Figure 89 shows the same data where the synthetic peaks have been assigned a common name, which is different from the integration region.

A Custom report generated from this newly named set of quantification items is shown in Table 13 and can be seen to be identical to the report in Table 11

Table 13: Custom report after the synthetic components have been renamed

O 1s CPSeV	Al 2p CPSeV	C 1s CPSeV	O 1s %	Al 2p %	C 1s %
17050.5	11360.8	2860.9	54.523	36.3287	9.14836

The value of the Custom report lies in the ability to combine intensities, such as those from the synthetic components used to model the O 1s region, with the intensities from the integration regions for the C 1s and the Al 2p regions. Table 14 has been calculated from a combination of the “Peak O s” intensities and the previously used Al 2p and C 1s regions. The real value comes when the O 1s region, say, has an interference peak which requires excluding from the final results by fitting multiple peaks to the data but only selecting a subset of these peaks for the purposes of the quantification report

Table 14: Custom report using the Synthetic Peaks for the O 1s intensity but the integration regions for the C 1s and Al 2p intensities

Peak O 1s CPSeV	Al 2p CPSeV	C 1s CPSeV	Peak O 1s %	Al 2p %	C 1s %
17024.9	11360.8	2860.9	54.4856	36.3585	9.15586

In order to further enhance the “cross platform” approach inherent in ISO 14976, CasaXPS provides several file conversion routines for the “native” file format of systems which either do not have any “VAMAS” output capability or which do not implement the specification fully. brief descriptions of procedures for these conversions are given below.

Kratos DS800 Binary Files

Kratos DS800 binary files can be directly converted through CasaXPS.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty sub-frames are amongst the list of windows then the New toolbar button must be pressed to create an empty sub-frame.

Figure 90 shows the “Convert to VAMAS file” dialog win-

dow where the “Files of type” filter allows only “.mpa” files to appear in the listing. Simply select a DS800 file and press Open or double-click the desired file. CasaXPS will convert the binary file into the ISO 14976 file format, where the new file will have the same name as the original .mpa file but with a .vms extension added.



Figure 90: Convert to VAMAS file showing DS800 files.

VG Eclipse Files

VG Eclipse files are binary files, one file per spectral region, organised in directories where files with common acquisition characteristics appear in the same directory or inside sub-directories within that directory. CasaXPS will convert these directory structures into a single ISO 14976 file where experimental information and relationships are maintained. Quantification in CasaXPS can be performed in an identical way to the Eclipse data system; transmission correction is accounted for by CasaXPS whenever the information is included in the original files.

Individual spectral regions are stored in files with a .dts extension. Sets of regions are stored in individual files all located in the same sub-directory. Older versions of Eclipse may have .col extensions associated with these data directories, but later versions of Eclipse generate such directories without this naming convention. Similarly, depth profiles or angle resolved experiments appear as sets of directories, one per acquisition region, where each sub-directory contains the .dts files for one of the acquisition regions involved in the experiment. Again, older versions of Eclipse use a naming convention where each sub-directory was given a .mle extension.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled un-

less the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is amongst the list of windows then the New toolbar button must be pressed to create an empty Frame.

There are three ways to convert VG Eclipse data.

- If an individual file is required then selecting the .dts file via the Covert to VAMAS file dialog window will result in the creation of a VAMAS file that contains the spectrum in that .dts and no others.
- If, on the other hand, all the .dts files are require in the same VAMAS file, the Covert to VAMAS file dialog window should be used to find the directory containing the spectral regions. The name of the new file entered in the dialog window must be specified in the text-field, but with a .col extension. It is the extension given to the file name that flags the type of conversion required, however the new file created by CasaXPS will replace the entered extension by .vms.
- The third method converts depth profile type file structures. Using the Covert to VAMAS file dialog window, move to the directory that contains the sub-directories containing the spectral regions. Whilst at the level of the sub-directories enter the name of the new file name and add a .mle extension. The .mle flags that CasaXPS should read each sub-directory within the current directory and convert all the .dts files found within the sub-directories. The new file name will replace the .mle

extension with .vms. Figure 91 shows the dialog window



Figure 91: Convert to VAMAS File Dialog Window

where the visible directories are about to be converted using the .mle flag. The file name profile .mle indicates that each of the directories form part of a depth profile and should be interpreted as the names for the regions found inside these directories. The result of converting

these directories is shown in Figure 92.

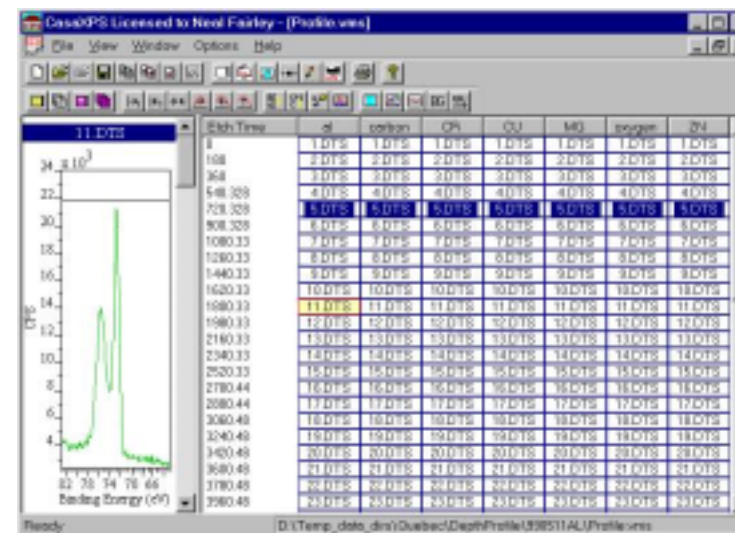


Figure 92: VAMAS File Block Structure from a VG Eclipse Depth Profile

VGX900 (Ron Unwin)

Ron Unwin developed the VGX900 system and so all the various forms of Ron Unwin's files can be converted in the same way.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is

amongst the list of windows then the New toolbar button must be pressed to create an empty Frame.

Files generated by the VGX900 system are ACSII files and each file may contain one or more spectral regions. The method for loading these files into CasaXPS is to collect a set of the files in a sub-directory and then enter a new file name in the “Convert to VAMAS file” dialog window, but add an extension of .unw to the specified name (Figure 93).

All files in the directory will be read and appear in a single VAMAS file. Since the files have no characteristic file type it is essential that only files generated from the VGX900 system are present in the sub-directory and no others.

Some formats of the VGX900 system do not include information necessary for a proper description of the data (in the VAMAS sense), but the system does allow the writing of VAMAS files based on a user specified template. This allows data to be written in a form directly readable by CasaXPS, but the user must ensure that the correct fields have been used to describe the spectra.

Alternatively, the conversion option in CasaXPS can specify some of the missing information using flags. The X-ray anode defaults to Aluminium and the energy-scale defaults to Binding Energy. Appending the following strings to the file

name at conversion time can alter these two options:

- anode Mg specifies Magnesium anode.
- energy KE specifies scans with increasing KE sweep.

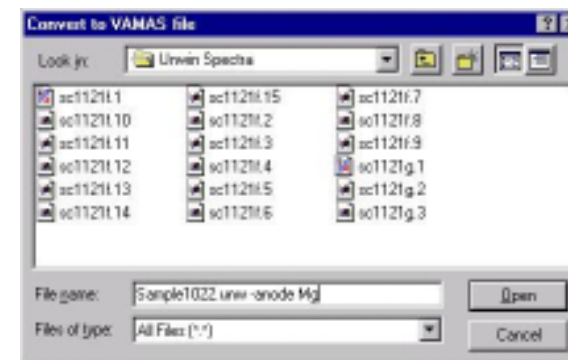


Figure 93: Convert to VAMAS file dialog window VGX-900.

Note the .unw extension used to indicate that the set of files contain spectra acquired using Ron Unwin's data system. The -anode Mg is only used to force the X-ray source to be Magnesium instead of the default setting. The latest release of Ron Unwin's system includes this information.

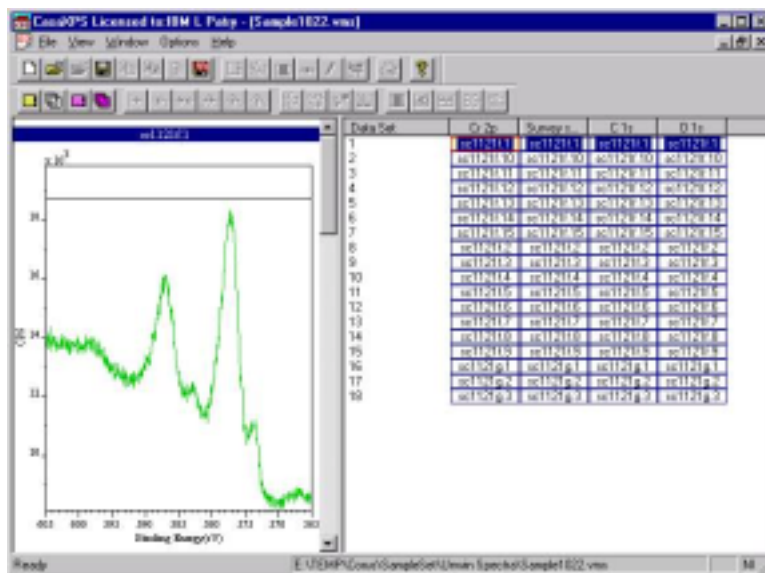


Figure 94: The set of spectra contained in the sub-directory shown in Figure 93 are offered in CasaXPS using the browser and spectrum display frames.

“Dayta” System Files (Bristol IAC system)

Spectra acquired by the Bristol “Dayta” system are stored as sets of ASCII files with .SP_ file extensions. These files may be connected to a particular experiment through an associated .seq file. Sets of .SP_ spectrum files are listed in the .seq file together with the any experimental information (e.g. etch times).

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is amongst the list of windows then the New toolbar button must be pressed to create an empty Frame.

There are two methods for loading Bristol files.

- The simplest method is to convert files that are listed in a .seq file. All that is required is for the .seq file and all the .SP_ files listed within that file are located in the same sub-directory, then the .seq file is selected through the “Convert to VAMAS file” dialog window. If for any reason the set of .SP_ files does not match the set listed in the .seq file, then the .seq will need to be edited to remove or add entries as necessary. The VAMAS file that results will include an experimental variable for each spectrum loaded.
- The alternative method for loading the .SP_ files is to collect a set of these files in a sub-directory and then enter a new file name in the “Convert to VAMAS file” dialog window, but add an extension of .bri to the specified name (Figure 95). All .SP_ files in the directory will be read and appear in a single VAMAS file. The default action is that no experimental variable is defined for each spectrum. As a consequence spectra read will all appear in the same row in CasaXPS (i.e. in a format suitable for quantification of high resolution spectra). If on the other hand it would be more convenient to separate spectra by an experimental variable then the new file name entered in the dialog window should be followed by a –col string e.g. “newfile.bri –col” (omit the quotation marks). The

experimental variable is the index number determined from the order in which the files are read.

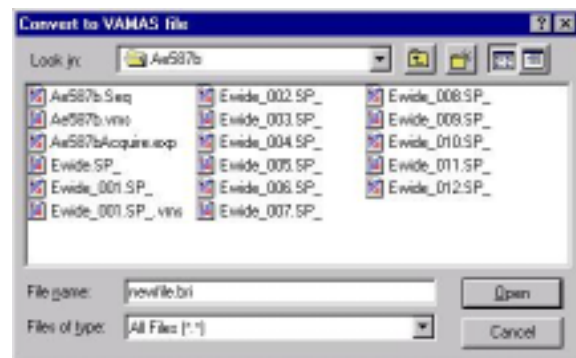


Figure 95: Convert to VAMAS file dialog window for Dayta A.bri file extension has been specified. The result is a file called newfile.vms containing all the .SP_ data files seen in the window

Kratos Vision 1.x/2.x ASCII files

The Kratos Vision data sets are probably the closest in logical design to the ISO 14976 file format. A single binary file maintains the context for an experiment, where an experiment may include multiple acquisition regions characterized by, so-called, state changes. These state-changes are equiv-

alent to the experimental variable fields used in the ISO 14976 standard and the Vision Objects are equivalent to the VAMAS blocks.

The principal problem with converting the Vision 1.x/2.x data sets is the internal data-structures used to maintain the experimental information. The file structure is a link-list of file records that hold a hierarchical description of the data. Fortunately there is a utility called `dump_dataset` (available from Kratos), which converts the binary format to an ASCII version and it is the ASCII version that CasaXPS converts to the ISO 14976 standard.

To create an ASCII version of a Kratos Vision data set follow this procedure:

at the command line level on the Unix workstation, move to the data directory and type

```
$dump_dataset filename all > new_filename.kal
```

(where represents space and `$` the command line prompt)

The extension `.kal` is recognized by CasaXPS and tells the conversion routine to parse the data using the Kratos Vision option.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is among the list of windows then the “New” toolbar button must be pressed to create an empty Frame.

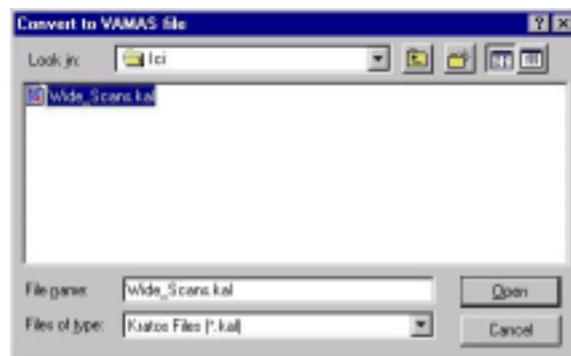


Figure 96: Convert to VAMAS file Dialog Window. The file display filter is set to show only .kal files

Selecting a .kal file via the “Convert to VAMAS file” dialog window causes an ISO 14976 file to be generated with the same name, but with a .vms extension added. The spectra will be converted and other experimental information extracted so that the ISO 14976 file may be quantified to produce identical results to those from the Vision systems. That is, transmission correction is included as part of the ISO 14976 file.

Element library files from the Vision 1.x system can also be converted to CasaXPS format and Vision 2.x element libraries can be converted on request.

SSI M-Probe Files

SSI M-Probe files may exist in an ASCII format where multiple regions appear in the same file or alternatively, the data is described by a control file with the spectra stored in separate .reg files in a sub-directory where the sub-directory has the same name as the base-name of the control file.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is amongst the list of windows then the New toolbar button must be pressed to create an empty Frame.

There are two methods for loading M-probe files:

- ASCII format files must have an .mrs extension and may be opened directly through the Convert Dialog window.
- The binary format consists of an .mrs file together with a directory of the same name as the base-name of the .mrs file. To load the information held in this binary format the base-name of the .mrs file must be specified but with an .ssi extension replacing the .mrs extension (Figure 97.)

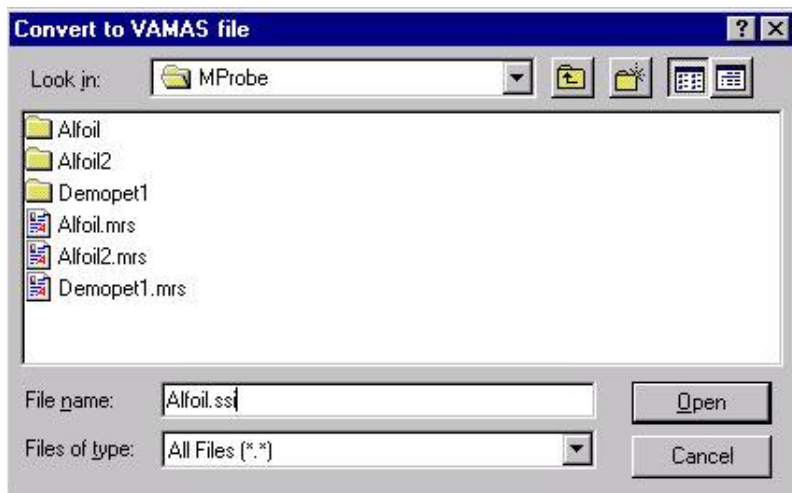


Figure 97: The binary data held in the file/directory pair named Alfoil will be converted to ISO format.

PHI MultiPak ASCII files

PHI MultiPak writes an ASCII format. Data in this format can be converted to ISO format through CasaXPS.

The first step to creating a new VAMAS file is to enable the Convert toolbar button. The Convert option is disabled unless the selected sub-frame window representing a file contains no VAMAS regions. If no empty Experiment Frame is

amongst the list of windows then the New toolbar button must be pressed to create an empty Frame.

The ASCII files may contain single regions, so called multiplexed regions, or depth profile data. To convert a set of files in any of these formats into a single VAMAS file, the .asc files must be collected into a sub-directory that contains only valid .asc data files and then a new filename must be entered with an extension of .qua (Figure 98) A VAMAS file containing all the regions from the files in the directory with an .asc extension will be created.

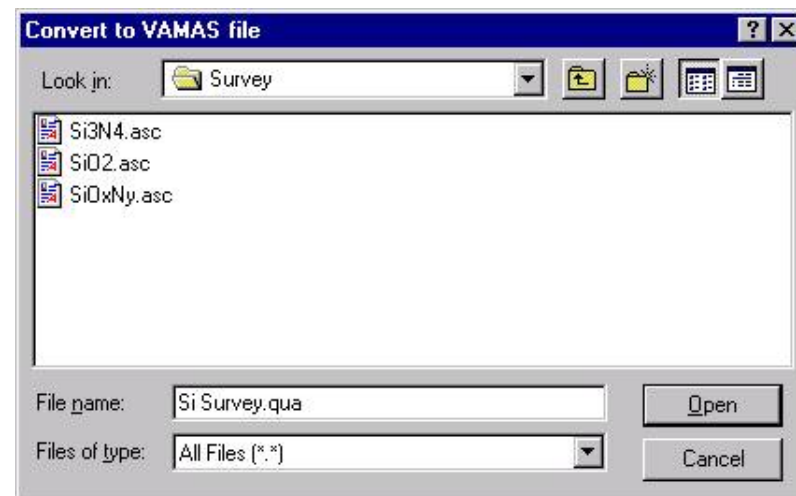


Figure 98: MultiPak ASCII files are converted by entering a filename with an extension of .qua.

The CasaXPS system has three “levels” of control - the familiar Windows style “Menu bar”, a system specific “Tool-bar” and “Options bar” (providing commands or command windows accessed by clicking “button” icons) and keyboard shortcuts. This chapter provides brief “one line” descriptions indicating where a command is - how it is accessed, what the command does, and the part of the system or data to which it applies.

Main menubar

The main system window Menu bar contains five menus, File, View, Window, Options, and Help .



File

New opens a fresh (blank) Experiment Frame inside the CasaXPS main window (Programme Frame).

Open... brings up a “file dialog” window so that a chosen (existing) file is opened and displayed in the current

(top) tile.

Convert (“greyed out” (disabled) unless the current Experiment Frame is empty) brings up the “Convert” file browser (see “Using Different File Formats” on page 123)

Close dismisses (closes) the current Experiment Frame, providing the file it represents is saved.

Save As... enables changing a file name and/or saving to a different directory

Save Picture... brings up a “save” file browser enabling the displayed graphic to be saved as a “Windows Enhanced MetaFile” (.EMF)

Print opens the standard Windows print dialog box for the system default printer.

Print Setup invokes the standard Windows print set up box for the system default printer.

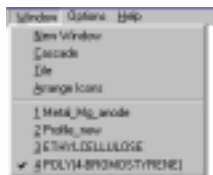
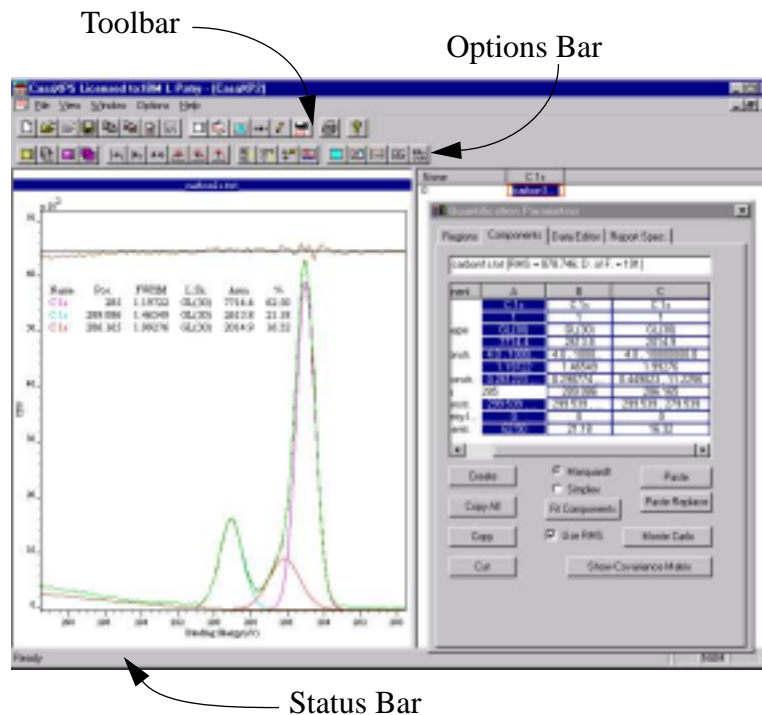
(recent files) the system indicates the four most recently opened files (a “push down” list). Clicking on one of these opens a new Experiment Frame containing the file. **Exit** quits CasaXPS entirely.

View

Toolbar toggles (sets on or off) visibility of the Toolbar

Options Bar toggles visibility of the Options bar

Status Bar toggles visibility of the Status bar



Window

New Window launches a fresh (blank)processing window (a tile) within the CasaXPS work area

Cascade overlaps a number of processing windows

tidily, so that the top and left edges are visible “behind” the front one.

Tile fills the work area uniformly with the open windows so that all their contents are visible. The usefulness of this command depends on the screen area available, but it works best with a even number and is seldom effective with more than four tiles, on an average (800x600 pixel) screen

Arrange Icons tidies any minimised processing windows, “docking” them in rows along the lower edge of the work area.

(File List) provides a numbered list of the file names for the open windows

Options

Launches one of six dialog windows for the major command sets and processing activities:-

Page Layout... deals with the way in which the graphical display area of the active tile is arranged (in rows and columns).

Tile Display... deals with the way that parameters (e.g.energy scale, intensity) are displayed within a tile and sets attributes such as colour.

Quantify... provides access to the **Regions** and **Components** specifications, enables definition of **Report** format (standard or custom, with defined names and tags) and also provides a numerical **Data Editor** (for spike removal)

Elements... controls the setting and display of the basic line position database for the system, accessing **Element Table**, **Periodic Table**, and **Load File** dialog

Processing... provides access to the major processing routines of the system: **Smoothing, Integration, Differentiation**, and Principal Component Analysis (**PCA**), as well as **Test Data** (including spectrum synthesis), (Energy) **Calibration, Intensity Calibration**, a spectrum **Calculator** (e.g. for addition and subtraction of whole spectra) and the **Processing History** for the selected tile.

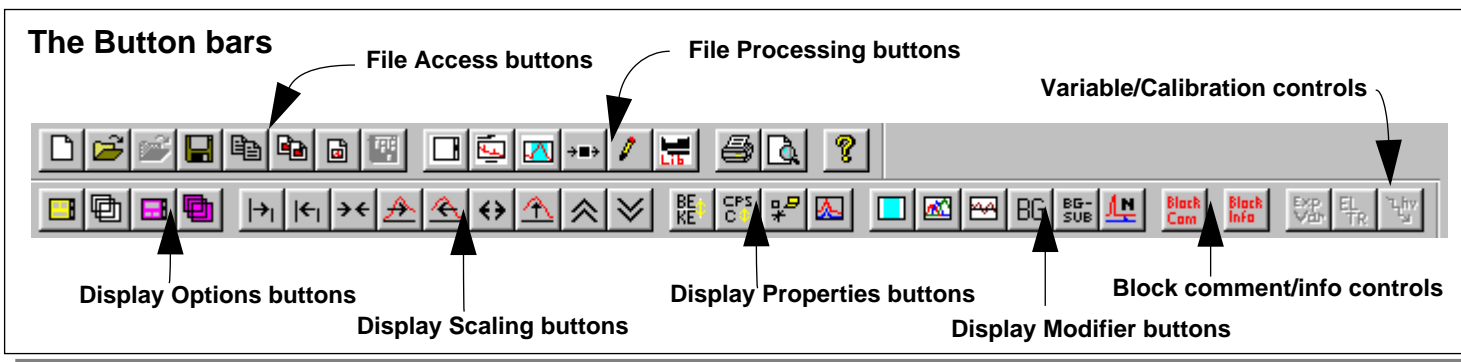
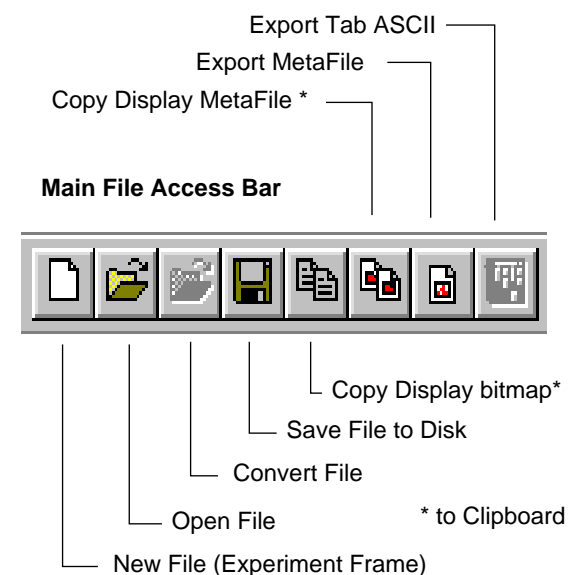
Annotation... provides **Peak Labels**, and (general) **Text** annotation for the display, control for **Regions, Components**, and **Quantification** (Table) storage and presentation and access to the **Annotation History** for the selected tile.

Help

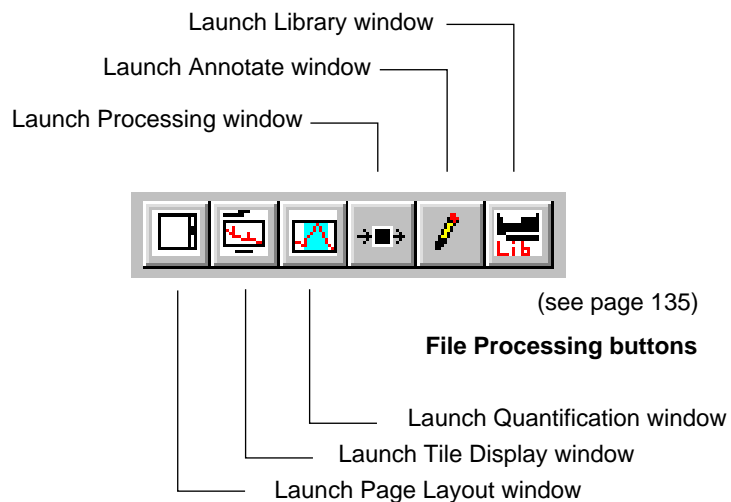
About launches a window giving the system version number, license status, and a dialog enabling change or update of license.

Help launches on screen (html file based) help.

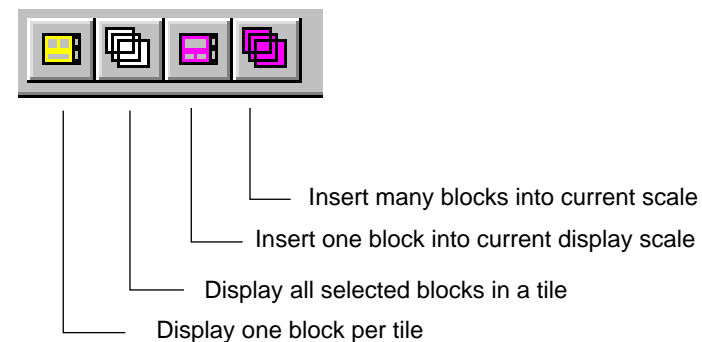
Toolbar



Options Bar



Display Options buttons



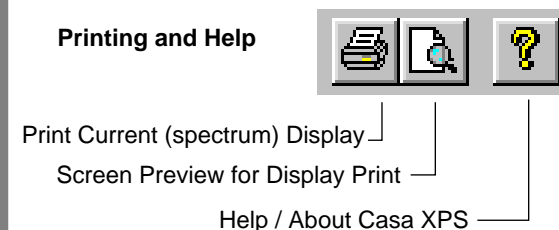
The **Toolbar** (upper) buttons in general provide access to menus or dialog boxes.

The **Options** bar (lower) buttons in general execute actions immediately or (reversibly) change the format of the display.

“Hovering” over a button (placing the mouse screen pointer on an icon without clicking) produces a descriptive label for that button (a “Tool Tip”), and a slightly longer description in the **status bar**

Function Buttons F5 - F10 mirror the **Processing** buttons (Page Layout - Element Library)

Printing and Help



Decrease intensity scale maximum

Increase intensity scale maximum

Reset Intensity scale to original maximum

Increase Energy scale range*†

Step zoomed energy scale left*

Display Scaling buttons

Step zoomed energy scale right*

Reset to original scale

Zoom out (step back round history)

Zoom in (requires selection box)

* rescales intensity to suit
† adds 20% left and right

Toggle normalised display

Toggle subtracted display

Display Modifier buttons

Toggle background

Toggle residuals display

Toggle components display

Toggle (shaded) region display

NB. Modifiers require prior function definition

Display Properties buttons

BE KE †

CPS C †

† "tri-state" button

Toggle Binding and Kinetic Energy scales

Toggle Counts and CPS Intensity scales

Toggle 2D, 3D, and Factor Space Display †

Toggle Offset for multiple traces †

Comment, Information & Variable control

Edit Block Comment

Edit Block Information

Block Com

Block Info

Exp Var

EL TR

hw

For Display

For Browser

Set Experiment variable (linear)

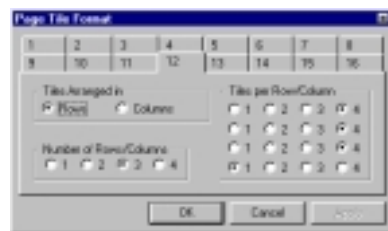
Edit Species/Transition parameters

Edit Source/Analyser parameters

Processing Dialog windows

Page Layout

(see “Tiles of Spectra” on page 26)



Sixteen predefined layouts are provided, corresponding to arranging spectra in symmetrical rows up to the maximum (16) allowed. All predefined formats may be changed as required, for example to provide a project specific layout set, the three controls in each layout functioning as “radio buttons”, allowing selection of only one alternative or combination for each case.

Tile Display Parameters

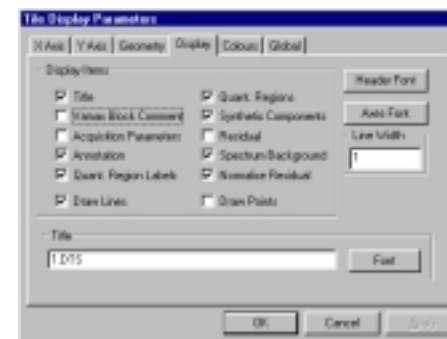
(see “Tile Display” on page 27)

The six “tabs” give access to tile (individual spectra) parameters in a straightforward way. Settings may be applied to individual tiles or all tiles in an Experiment (via “Global” tab).

X-Axis

Independent variable provides a “radio button” choice between BE and KE for x-axis display, and a means of labelling the axis as desired (the type-in boxes) and setting the

typeface (“Font”).



Min and max enable setting the span of the display (which may be greater or less than the data range for the selected tile). Type in decimal values: values relate to the scale selected above, so take care with max and min values for BE!)

Set Scale enables a linear remapping of the x-axis for use in e.g. depth profiles, Most usefully employed when the output report is saved to a fresh file with “.vms” extension, then re-opened in the system and the scale changed.

Display Minor Tick-Marks is a simple “on - off” control

Display Axis Scale similarly turns off the x-axis scales and labels

Rescale Axis changes the scale to KE and also changes the label to whatever is typed in the adjacent box

Y-Axis

Dependent variable provides a “radio button” choice between “counts per second” and “counts” (“counts per bin” i.e. counts accumulated in each energy interval acquired) for y-axis display, a means of labelling the axis as desired (the

type-in boxes) and of setting the typeface (“Font”) for this label.

Min and max enable setting the span of the display (which may be greater or less than the data range for the selected tile). Type in decimal values, which relate to the scale selected above.

Offset spectra provides an automatically calculated shift in the y-axis for each overlaid spectrum. The “%” value sets the size of the y-axis scale in relation to the available area (*not* to the offset). The offset is then calculated to divide the remaining space equally.

Display Minor Tick-Marks is a simple “on - off” control

Display Axis Scale similarly turns off the y-axis scales and labels

Normalise Display sets y values equal at indicated point on the x-axis. To move normalisation point, use “shift-left-click” in the display tile (uses the foreground trace scale in a multi-trace display). Control mirrors the “normalise” button.

Display Exp. Variable Left/Right labels each trace at its right or left extremity with the corresponding value of the experimental variable (e.g. etch time) Non - exclusive.

Display Background Subtracted includes or excludes background from display. Greyed when not available

Geometry

Tile Display Geometry selects either 2D, 3D, or Factor Space. The parameters immediately below apply to 3D (Back Plane and Front Plane) and Factor Space (Scatter) dis-

plays only. 2D has no further options.

Reverse Display mirrors the *order* of plotting traces in a 2D or 3D display (“top” to bottom” or “front” to “back”) but not the traces themselves (spectra are always plotted with increasing KE left to right).

Back Plane Parameters provide scaling and offset (and thus “perspective”) for the “furthest” (“back”) trace of a series of traces visualised as slices of a cube, subject to plotting within the tile area. Values are percentages. Offsets may be positive or negative and relate to the leftmost or rightmost point in the display subject to the constraint of plotting within the viewable area in the tile.

Front Plane Parameters are similar except that $y=0$ is always the front baseline (no offset). Colours for 3D background and fill are set in the “Colours” tab

Scatter Parameters provide for rotation (in degrees, positive or negative) about the three axes (x, y, and z), and shift (percentage of display area) in x and y, constrained as above.

Display

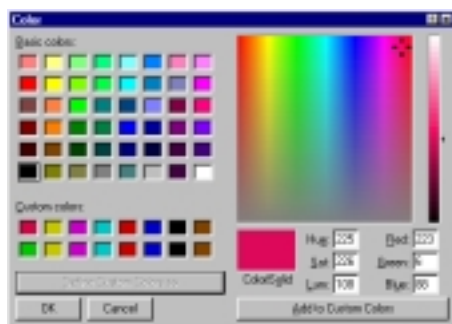
Display provides simple “checkbox” (on - off) controls for items to be included in the display. Traces are normally displayed as “lines” (continuous curves) unless “Draw Points” is checked, when actual point-by-point values are plotted with no joining lines (useful, e.g., for curve fitting, overlaying envelopes (lines) and points). Typefaces for the Header, Title, and Axes labels are controlled here (a fresh header may be added in the input box here) and the line width (in pixels) of the display defined (e.g. thicker lines for clearer

AV presentation or graphical capture and scaling)

Colours

provides a means of setting colour for Spectra, Spectrum Background, Tile Background, 3D Background, Region Background, 3D Fill, Residual Trace, and Synthetic Components.

Clicking any of these items brings up a Standard + Custom Colour swatch for selection and a further click on “Add to Custom Colours” extends this to a Colour Picker with HLS, RGB and visual presentation



When more than one item of the same type is present in a display (e.g. spectrum traces), then custom colours are cycled through in order

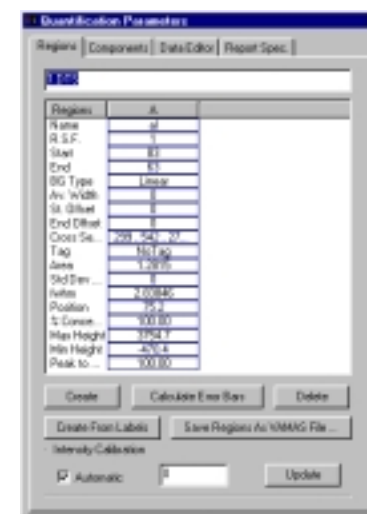
Global

This tab provides a means of applying the changes made with the other tabs to all the tiles in the current Experiment Frame, including (or otherwise) x and y axis ranges and the title. The Tabs also include buttons to enable saving the default parameters for colours, settings, and fonts (typefaces). A directory called “casaxps.def” should be created at the

same logical level on the currently operational system disk as that at which the file “casaxps.exe” is stored. CasaXPS will then create and store files for all the default parameters for acquisition, colours and so on here.

Quantification

(see “Quantification” on page 42)



Regions

This tab provides access to the mostly self explanatory parameters for any new or existing region for the displayed block or blocks.

Note that all the parameters in the region list are editable. A selected region as a whole (click on the header letter) has a blue background: an individual parameter selected by a mouse left click has a plain white background and the exist-

ing value is left justified. Changing an entry requires a “return” (enter) key at the end of the entry to validate it: merely clicking on a new parameter restores the old value. Note also that correct spelling is required for background type (linear, Shirley, Tougaard - though case is insignificant).

Create brings up a new region with (as far as possible) the parameters entered for the displayed peak.

Create From Labels provides a region whose parameters are taken from the library entry with the same label as the displayed peak (see “Step 5: Create quantification regions” on page 11)

Delete deletes the selected region and its parameters from the Quantification window (and thus from calculations and display).

Calculate Error Bars completes the standard deviation calculation and related procedures.

Save Region As VAMAS File records the region parameters in the block header (see page 145)

Intensity Calibration provides either automatic incorporation of the transmission function data held in the (Kratos style) VAMAS file (see also page 145) or manual entry of a specific factor (transmission exponent) in the type-in box. If CasaXPS detects the presence of transmission function correction values in the data block (an appropriate ‘corresponding variable’) then the ‘automatic’ checkbox will be ticked and the quantification will proceed using the correction

(note that the original data are not affected by this correction, only the quantification). The manual entry (type-in) box in general should only be used when automatic correction is not available. The Update button applies a new manual value or the automatic setting after any change in the status of the checkbox.

Note that the width of the region list may be altered in the usual fashion by dragging on the edge of the divider in the header bar.

Components

(see “Tile Display” on page 27, and also see “Line Shapes and Backgrounds” on page 55)

This tab defines the parameters for each synthetic component for a particular region and also controls peak fitting.

Note that the **Cross Section** is the Tougaard “3-parameter” constant set (consisting of four (!) comma separated values) (see “U 4 Tougaard (short form U 4) :” on page 67).

Offset provides a numerical adjust for spectrum overlay in display after background subtraction. the value entered is a percentage (of the largest peak to be displayed) and may only be positive. Positive values result in movement of the region ‘downwards’, offset towards the baseline.

Max Height and **Min Height** are derived from the region under consideration (signal values above / below background) and used e.g. with AES signals (or other derivatives) in determining peak to peak concentration.

See page 11 for notes on other button use.

The **Report Spec** tab enables customised output of quanti-



fied values, with the ability to apportion and aggregate areas and set up mathematical relationships, as well as produce 'standard' reports. The **Data Editor** provides access to individual channel counts to enable, for example, spike removal. Note that this is a potentially dangerous procedure which can modify original data.

Processing

(see "Processing" on page 34)

The **Smoothing**, **Integration**, and **Differentiation** tabs have similar controls (and use related procedures). **Zero adjust** (in **Integration**) provides for a shift of the minimum ordinate value after integration to the axis zero. **PCA** (see "Principal Component Analysis" on page 80) is an entirely

different and specialised procedure. The two **Calibration** procedures are straightforward (see "Energy Calibration" on page 35, "Intensity Calibration" on page 36) as is the **Calculator**. The **Test Data** tab provides not only standardised files for benchmarks, but also access to the Monte Carlo procedures (see "Monte Carlo Methods - Uncertainties in Intensity Calculations" on page 97).

Processing History provides a complete or selective "undo" function as well as a record of applied methods.



Reset removes all processing from the selected block and returns it to the original state.

Apply selection resets the displayed block then reapplies only the selected items from the existing history. Use 'control' and 'shift' keys with the mouse to select non-contiguous or continuous items, respectively.

Propagate flag sets the 'propagate' flag to apply the current selection from the history to a series of (selected) blocks.

Annotation

(see “Graph Annotation” on page 31)



Controls placement and style of text for annotation (**Text**), **Peak Labels**, **Regions** and **Component** summaries, and **Quantification** tables in the display tile and printed output. An **Annotation History** panel functions in a similar fashion to the Processing History, enabling removal of unwanted items or selective application. Each tab has provision for changing of typeface (font, size and colour). Position for annotation items may be referenced to the display tile frame or the data itself (for convenient placement when multiple blocks are displayed) and labels may be orientated vertically or horizontally. Positioning targets (small squares with central dots) may be set with the mouse anywhere within the display tile and disappear when the annotation window is dismissed.

Library

(see “Element Library” on page 38)



In order to function correctly, CasaXPS should have an element library (table of defined line positions, widths, shapes, sensitivity factors, excitation source and labels) present in the system. This table may list any (single or composite) transition appropriate for the analysis undertaken (it is not limited to chemical elements) and may be changed at any time (so that several files may be held on disk), but only one library (of unlimited size) may be active at any time.

Input File (above) enables loading, changing or combining the line position data file which forms the basis for feature identification. File selection (of either CasaXPS or Kratos basic type) may be direct, typing in a file name or using a file browser, and the selected file may be merged with or may replace the existing file (if any).

The **Element Table** window provides a scrolled list of the contents of the Library File: clicking on a name (toggle) transfers a labelled position marker to the display for the selected item. Position markers may be removed from the dis-

play with the **Clear All Elements** button. **Find Peaks** adds coloured markers to the display for all major peaks in the displayed block, according to a defined algorithm (see XREF xxx): the **Clear Markers** button removes these.



The **Periodic Table** window provides a simpler means of selecting elemental transitions (from the current element library, if any) for display labelling. Each element symbol is a toggle (click on - click off) button).



Keyboard & Mouse Shortcuts

Function	Shortcut*
New Experiment Frame (Browser)	Ctrl+ N
Open ISO File	Ctrl+ O
Print (graphic)	Ctrl+ P
Copy (graphic to clipboard)	Ctrl+ C
Set 'Normalise' point	Shift+ (left click)
Zoom out (cycles)	Ctrl+ (right click)
Context menu	(right click)
Edit a data point	Ctrl+ Shift+ (left click)
Display (spectrum)	F1
Overlay (spectra)	F2
Zoom out	F3
Reset Zoom	F4

* F1 - F4 are single press "Function" keys: for the other commands, hold down the "control" (Ctrl) and/or "shift" key and press the listed key or mouse button.

"**Drag and Drop**" - you may load ISO 14976 files by dragging them from Windows™ Explorer and dropping onto an open CasaXPS Programme Frame.

Appendix 1: ISO 14976 format files annotated

1. Simple single block XPS (region) spectrum

Item Description	Line	ISO 14976 Format Item (in PET-C1s.vms file)
format identifier	(1)	VAMAS Surface Chemical Analysis Standard Data Transfer Format 1988 May 4
institution identifier	(2)	Acolyte
instrument model i.d.	(3)	Kratos AXIS-HS
operator i.d.	(4)	AC
experiment i.d.	(5)	PET C1s Test Spectrum
# comment lines	(6)	1
(comment lines)	(7)	Acolyte copyright 1999
experiment mode	(8)	NORM
scan mode	(9)	REGULAR
# regions	(10)	1
# experiment variables	(11)	0
# parameter exclusion entries	(12)	0
# manual items	(13)	0
# future experiment items	(14)	0
# future block entries	(15)	0
# blocks	(16)	1
block i.d.	(17)	C1s 1
sample i.d.	(18)	PET
year (4 digits)	(19)	1999
month	(20)	06
day	(21)	24
hour	(22)	14
minute	(23)	45
second	(24)	00
# hours + GMT	(25)	1
# lines in block comment	(26)	0
technique	(27)	XPS
source	(28)	Al
source energy	(29)	1486.6
source strength	(30)	300

source width x, (mu)	(31)	500
source width y, (mu)	(32)	500
source polar angle	(33)	54
source azimuth	(34)	180
analyser mode	(35)	FAT
analyser resolution characteristic	(36)	20
magnification of analyser	(37)	1
analyser work function	(38)	4.5
target (sample) bias	(39)	0
analysis width x, (mu)	(40)	200
analysis width y, (mu)	(41)	200
analyser axis polar angle	(42)	0
analyser azimuth	(43)	180
species	(44)	C
transition state	(45)	1s
charge of analysed particle	(46)	-1
abscissa label	(47)	kinetic energy
abscissa units	(48)	eV
abscissa start	(49)	1191.6
abscissa increment	(50)	0.05
# corresponding variables	(51)	1
corresponding variable label	(52)	counts per channel
corresponding variable units (d= none)	(53)	d
signal mode	(54)	pulse counting
signal collection time/channel	(55)	0.5
# scans for this block	(56)	1
signal time correction	(57)	400E-9
sample normal tilt	(58)	0
sample normal azimuth	(59)	0
sample rotation angle	(60)	0
# additional params	(61)	0
# ordinate values	(62)	401
min. y	(63)	42
max. y	(64)	2250
(first data point)	(65)	179
(399 data points)	()	()
(last data point)	(465)	59
terminator	(466)	end of experiment

2. Multiple block depth profile composed of XPS (wide scan) spectra

Note: the line numbers for the “experiment header” and “block 1” have been deliberately left the same as in example 1, so that differences stand out clearly. Extra lines (variables and values) in this example are denoted with a + sign.

Item Description	Line	ISO 14976 Format Item (in DATA2.vms file)
format identifier	(1)	VAMAS Surface Chemical Analysis Standard Data Transfer Format 1988 May 4
institution identifier	(2)	Not Specified
instrument model i.d.	(3)	AXIS-165 (SAC)
operator i.d.	(4)	kratos (VISION User)
experiment i.d.	(5)	/export/home/kratos/data/2043it1.dset
# comment lines	(6)	0
experiment mode	(8)	NORM
scan mode	(9)	REGULAR
# regions	(10)	1
# experiment variables	(11)	1
experiment variable label	+	Etch Time
experiment variable units	+	s
# parameter exclusion entries	(12)	0
# manual items	(13)	0
# future experiment items	(14)	0
# future block entries	(15)	0
# blocks	(16)	117
block i.d.	(17)	OSnIn a a/2
sample i.d.	(18)	OSnIn/2
year (4 digits)	(19)	1999
month	(20)	12
day	(21)	22
hour	(22)	17
minute	(23)	40
second	(24)	2
# hours + GMT	(25)	0
# lines in block comment	(26)	3
“comment” lines used internally	+	Lens Mode:Magnetic Resolution:Pass energy 80 Anode:Al(180 W)

by Kratos Vision system	+	Step(meV): 300.0	Dwell(ms): 136	Sweeps: 1
CasaXPS adds "history" here	+	Acquired On: Tue Dec 22 17:40:02 1998		
technique	(27)	XPS		
experiment variable value	+	0		
source	(28)	Al		
source energy	(29)	1486.6		
source strength	(30)	180		
source width x, (mu)	(31)	1E+37		
source width y, (mu)	(32)	1E+37		
source polar angle	(33)	1E+37		
source azimuth	(34)	1E+37		
analyser mode	(35)	FAT		
analyser resolution characteristic	(36)	80		
magnification of analyser	(37)	1E+37		
analyser work function	(38)	1E+37		
target (sample) bias	(39)	0		
analysis width x, (mu)	(40)	1E+37		
analysis width y, (mu)	(41)	1E+37		
analyser axis polar angle	(42)	1E+37		
analyser azimuth	(43)	1E+37		
species	(44)	Wide		
transition state	(45)	none		
charge of analysed particle	(46)	-1		
abscissa label	(47)	kinetic energy		
abscissa units	(48)	eV		
abscissa start	(49)	926.7		
abscissa increment	(50)	0.3		
# corresponding variables	(51)	2		
corresponding variable label	(52)	Intensity		
corresponding variable units (d= none)	(53)	d		
	+	Transmission		
	+	d		
signal mode	(54)	pulse counting		
signal collection time/channel	(55)	0.136		
# scans for this block	(56)	1		
signal time correction	(57)	0		
sample normal tilt	(58)	1E+37		

```

sample normal azimuth (59) 1E+37
sample rotation angle (60) 1E+37
# additional params (61) 0
# ordinate values (62) 878
min. y (63) 427
max y (64) 15744
(first data point) (65) 2967
(first transmission value) + 1
(second data point) (66) 2954
(second transmission value) + 1
(875 total data & transmission points) ( + ) ( )
(last data point) (950) 458
(last transmission value) (951) 1
block i.d. (952) OSnIn a a/6
sample i.d. (953) OSnIn/6
year (4 digits) (954) 1999
month (955) 12
day (956) 22
hour (957) 17
minute (958) 43
second (959) 27
# hours + GMT (960) 0
# lines in block comment (961) 3
+ Lens Mode:Magnetic Resolution:Pass energy 80 Anode:Al(180 W)
+ Step(meV): 300.0 Dwell(ms): 136 Sweeps: 1
+ Acquired On:Tue Dec 22 17:43:27 1998
technique (965) XPS
experiment variable value + 30
source (967) Al
source energy (968) 1486.6
source strength (969) 180
.
and so on, as above for this block (block 2)
.
(last data point) (1884) 504
(last transmission value) (1885) 1
block i.d. (952) OSnIn a a/11

```

sample i.d. (953) OSnIn/11
.
and so on, as above for this block (block 3)
.
and also for another 114 blocks,
.
.
until
.
(last data point) (15894) 504
(last transmission value) (15895) 1
terminator (15896) end of experiment

Appendix 2: ISO 14976 and the World Wide Web

The ISO 14976 file format is designed for simple, unambiguous communication of information derived from surface chemical analysis experiments. The World Wide Web is one of the most important communications links for experimentalists, and so it is appropriate that means are provided for transport and display of ISO data over the web.

Standard methods of text file compression and transfer work well with the format because it is text based. Display of spectra in a web environment - on a browser page for example, is a (non trivial) matter which is of interest in many contexts, and there exists (at least one) simple Java applet which enables display of XPS spectra “live” on a web page. The applet is also designed to provide a basic practical mechanism for “platform independent” (that is “universal”) transfer and display of XPS spectra across the world wide web.

See, for example, “<http://www.acolyte.co.uk/JISO>”

“MIME” type. MIME types (Multimedia Internet Mail Extensions) are the designators which enable applications such as web browsers and their “helpers” to decide how to handle (interpret) an otherwise “unknown” file type. The “.vms” file extension is used for the ISO 14976 spectra and is in process of submission through the IETF as a “Chemical MIME” type. The present status is “experimental” (i.e. the type/sub-type) “chemical/x-vamas-iso14976” should be set. Since the format is solely text based, however, http systems (and browsers communicating with them) will in general react correctly if configured to serve (or read) unknown MIME types as “text/plain”. In most cases this will be the default action

Appendix 3: Definitions and Formulae

Glossary of terms

Block (ISO). A sub unit of an Experiment (file) (q.v.) which consists of a header - a set of parameters that only apply to that block - followed by a series of ordinate values which may represent a curve (e.g. a depth profile), a spectrum, or a map

Experiment (ISO). A complete ISO 14976 file which defines the context of and data acquired during a surface chemical analysis determination. Consists of a header containing of a series of parameters which apply to the measurement procedure as a whole, followed by one or more blocks (q.v) of data

VAMAS. the Versailles project on Advanced Materials And Standards

Formulae: Peak shapes

Gaussian/Lorentzian Product Form

$$GL(x, F, E, m) = \frac{\exp\left(-4\ln 2(1-m)\frac{(x-E)^2}{F^2}\right)}{1 + 4m\frac{(x-E)^2}{F^2}}$$

Gaussian/Lorentzian Sum Form

$$SGL(x, F, E, m) = \exp\left(-4\ln 2\frac{(x-E)^2}{F^2}\right)(1-m) + \frac{m}{1 + 4\frac{(x-E)^2}{F^2}}$$

Doniach Sunjic:

$$DS(x, \alpha, F, E) = \frac{\cos\left(\frac{\pi\alpha}{2} + (1 - \alpha) \operatorname{atan}\left(\frac{x - E}{F}\right)\right)}{(F^2 + (x - E)^2)^{(1 - \alpha)/2}}$$

Appendix 4: References and other Resources

World wide web links:

CasaXPS	http://www.casaxps.com
UK Surface Analysis Forum	http://www.uksaf.org
Kratos Analytical	http://www.kratos.com
Physical Electronics	http://www.phis.com
Thermo-VG	http://vacuumgenerators.com
American Vacuum Society (AVS)	http://www.vauum.org
ECASIA	http://dmxwww.epfl.ch/ecasia/
QSA	http://www.surrey.ac.uk/MME/QSA/
Queen Mary/Westfield College	http://www.chem.qmw.ac.uk/surfaces/
Institute of Physics(UK)	http://www.iop.org
Royal Society of Chemistry(UK)	http://www.rsc.org
American Chemical Society (ACS)	http://www.acs.org
American Institute of Physics (AIP)	http://www.aip.org
IUVSTA	http://www.vacuum.org/iuvsta/default.asp
Materials Research Society (USA)	http://www.mrs.org/

References from the Text Footnotes:

4. Tanaka A., *J. Surf. Analysis*, **1** 189 (1995)
5. M.P. Seah and M.T. Brown, *J. Elec. Spec.*, **95** (1998) 71-93
6. Shirley D. A., *Phys. Rev.*, **55**, 4709 (1972)
7. Doniach S. and Sunjic M., *J. Phys.* **4C31**, 285 (1970)
8. Evans S., *Surf. Interface Anal.*, **17**, 85 (1991)
9. Castle J. E. et al., *J. Electr. Spectr. Related Phenom.*, **106**, 65 (2000)
10. G. Wertheim, *J. Electron Spectrosc.* **6**, 239 (1975)
11. Tougaard S., *Surf. Interface Anal.*, **25** 137 (1997)
12. Tougaard S., *Surf. Interface Anal.*, **11** 453 (1988)
13. Jo M., *Surface Science*, **320** 191 (1994)
14. Fiedor J.N., Proctor A., Houalla M. and Hercules D.M., *Surf. Interface Anal.*, **20** 1 (1993)
15. Do T., McIntyre N.S., Harshman R.A., Lundy M.E. and Splinter S.J., *Surf. Interface Anal.*, **27**, 618 (1999)
16. (to be added)
17. K. Harrison and L.B. Hazell, *Surf. Interface Anal.*, **18**, 368 (1992)
18. P.J. Cumpson and M.P. Seah, Random Uncertainties in AES and XPS: Peak Energies, Areas and Quantification. NPL Report DMM(A) 26 ,May 1991; *Surf. Interface Anal.*, **18**, 345 (1992) and **18**, 361 (1992).
19. S. Evans, *Surf. Interface Anal.*, **18**, 323 (1992).
20. Beamson G. and Briggs D., The XPS of Polymers Database - CD-ROM , Surface Spectra Ltd (2000)
21. Jones F. et al, Fluoride uptake by glass ionomer cements: a surface analysis approach , submitted to *Biomaterials*.
22. Press W.H. et al, Numerical Recipes in C , Cambridge University Press (1988)
23. Cumpson P. J. and Seah M. P., Random Uncertainties in AES and XPS , *Surface and Interface Anal.*, **18** 361 (1992)

Appendix 5: Quick Reference Card

(following two pages - for printing)

Keyboard & Mouse Shortcuts

Function	Shortcut*
New Experiment Frame (Browser)	Ctrl+ N
Open ISO File	Ctrl+ O
Print (graphic)	Ctrl+ P
Copy (graphic to clipboard)	Ctrl+ C
Set 'Normalise' point	Shift+ (left click)
Zoom out (cycles)	Ctrl+ (right click)
Context menu	(right click)
Edit a data point	Ctrl+ Shift+ (left click)
Display (spectrum)	F1
Overlay (spectra)	F2
Zoom out	F3
Reset Zoom	F4

* F1 - F4 are single press "Function" keys: for the other commands, hold down the "control" (Ctrl) and/or "shift" key and press the listed key or mouse button.

"**Drag and Drop**" - you may load ISO 14976 files by dragging them from Windows™ Explorer and dropping onto an open CasaXPS Programme Frame.

Printing and Help



Print Current (spectrum) Display

Screen Preview for Display Print

Help / About Casa XPS

To Contact Us

Phone:

+44 (0) 1625 535346

E-Mail:

info@casaxps.com

World Wide Web:

<http://www.casaxps.com>



Casa Software Ltd.,

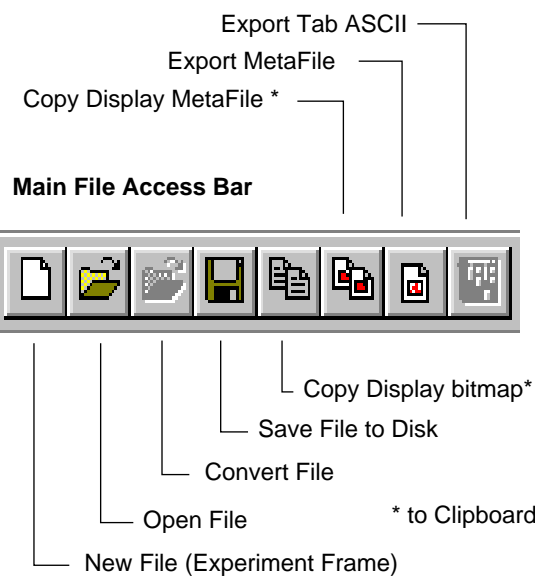
26 Burford Crescent
Wilmslow
Cheshire SK9 6BN
United Kingdom

CasaXPS

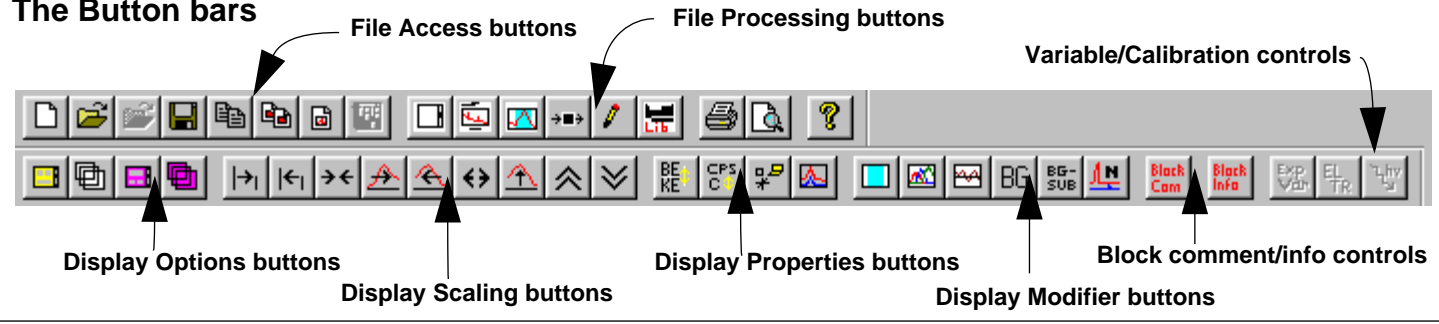
A Complete System for XPS Spectra

Version 2.0

QUICK REFERENCE CARD



The Button bars

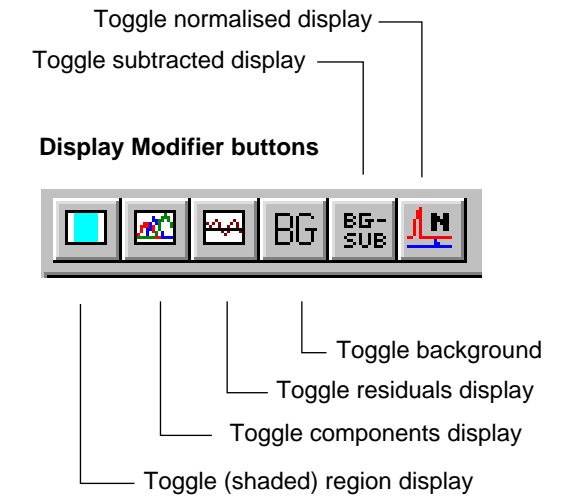
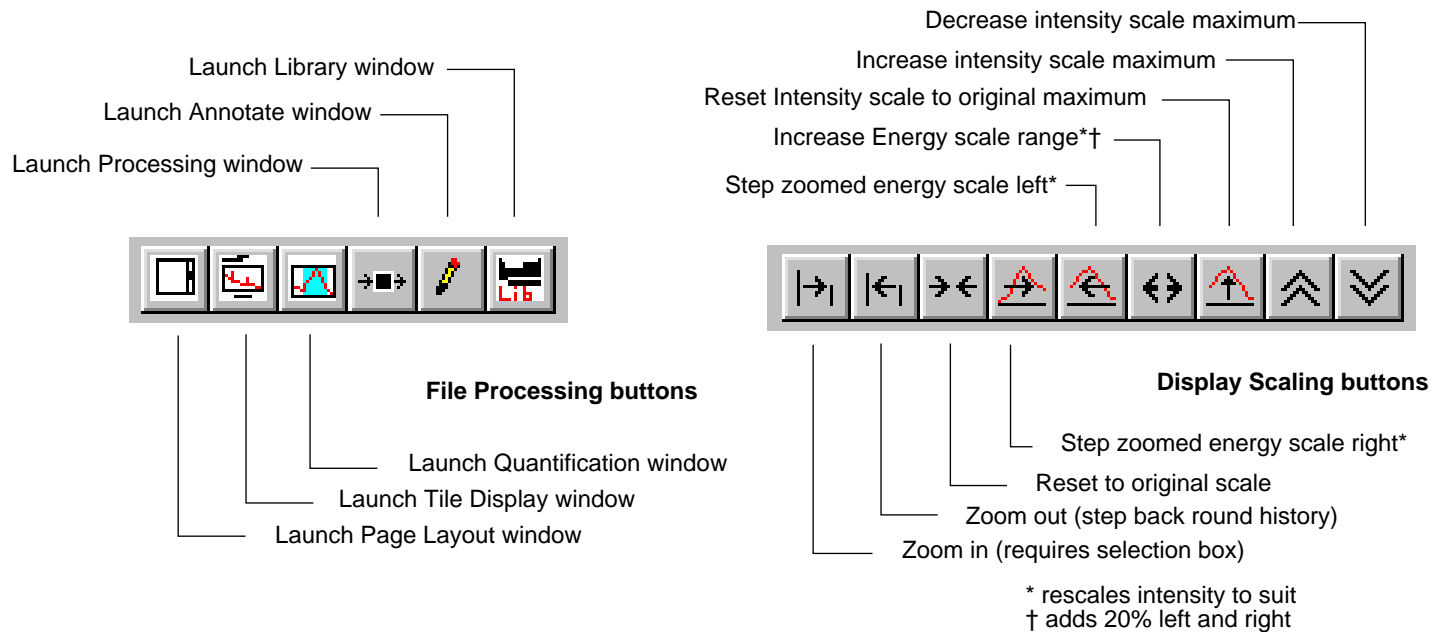


The **Toolbar** (upper) buttons in general provide access to menus or dialog boxes.

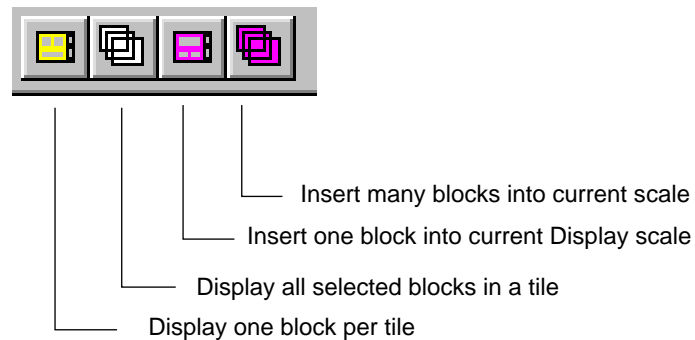
The **Options bar** (lower) buttons in general execute actions immediately or (reversibly) change the format of the display.

“Hovering” over a button (placing the mouse screen pointer on an icon without clicking) produces a descriptive label for that button (a “Tool Tip”), and a slightly longer description in the **status bar**

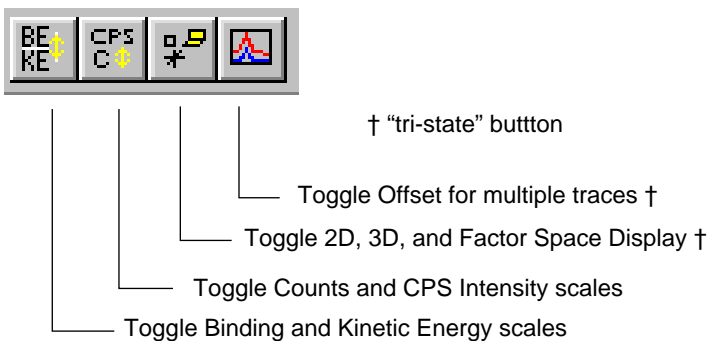
Function Buttons F5 - F10 mirror the **Processing** buttons (Page Layout - Element Library)



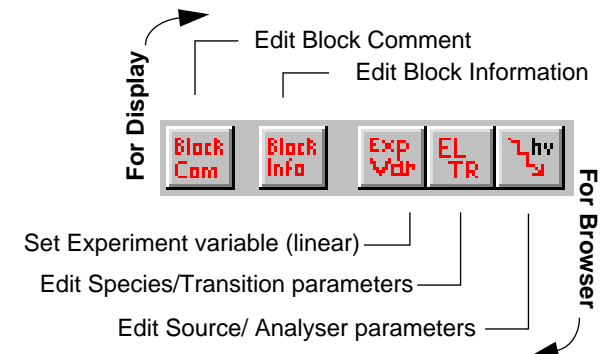
Display Options buttons



Display Properties buttons



Comment, Information & Variable control



Index

Symbols

~GL(50) 99
“Dayta” System Files 127
'GL(p)' 60
'GL(p)K(b0,b1)' 60
'GL(p)T(k)' 60
'SGL(p)' 60
'SGL(p)K(b0,b1)' 60
'SGL(p)T(k)' 60

Numerics

1E37 21

A

AET
 apparent error in the test vector 84
Analyser
 response function 55
annotation
 history 32
 moving 32
 repositioning 15
annotation dialog 11
artificial peaks
 test data 47
asymmetric line shapes,alternative 58
Auger lines 39

B

Background Subtraction 65
Backgrounds 43
Beamson and Briggs 107
block
 window 7
Briggs and Seah 61

Browser 15, 16
browser 7

C

CasaXPS 5
 installing 6
 starting 6
 terminology 6, 7
Colours 28, 138

D

data
 Zooming 26
Data Display 15
Data Editor 49
Define Custom Colors 28
Derivatives 50
Differential charging 55
display
 3-D plot 28
 Geometry 27
 window 7
Display Parameters 29
Doniach Sunjic 56, 58, 59, 61, 150
DS800 Binary Files 123

E

Eclipse Files 124
Element Library 16
 loading 40
Element Library dialog 9
Enhanced Metafiles 31
Excel 54
experiment file 8
Experiment Frame 9

Index

experimental variable 22

F

F line shape 64

F.W.H.M 40

File Formats 123

File formats

 DS800 Binary Files 123

fonts 29

G

Gaussian/Lorentzian

 Product 57

 Sum 57

Gelius, Ulrich 58, 61

Geometry 137

Graph Annotation 17

H

H line shape 63

I

IBM 5

Identifying Peaks 41

IE 86

IND 86

Installing CasaXPS 6

integration regions 13

intensity calibration 23

ISO 14976 5

 Binding vs Kinetic Energy 22

 blocks 16, 20

 Data Blocks 21

 Experiment header 21

 line separator 20

 parameter exclusion list 20

 Partially Encoded 19

 Selecting 26

 spectrometer geometry 20

 experiment. 20

 Experimental Variable 22

 File Format 19

 File Structure 20

 Quantification 24

 Transmission Functions 23

K

Kappa 59

L

Line shapes 44

 Asymmetric Blend 58

 Doniach Sunjic 58

 F profile 64

 Gaussian/Lorentzian 57

 H form 63

line shapes, including background 59

M

Macintosh computer 20

Microsoft Foundation Class 5

Microsoft Windows 95 5

Multiple Document Interface 8, 25

N

Name/Formula List. 53

NPL 20

Index

O

offset spectra 27

P

Page Layout dialog 15

peak labels 32

Peak Parameters

optimisation 47

PHI MultiPak ASCII files 130

Phonon broadening 55

polymer database 109

Processing 17

Q

Quantification 17

Calculation 45

Custom Report 46

Propagating 49

Report 45

Quantification Parameters dialog 11, 13

quantification regions 31

R

R.S.F 40

REELS 66

Regions 42

name 42

Regions Property Page 12, 13

REGULAR scan 22

Relative Sensitivity Factors 22

REP

real error in the predicted vector 84

RET

real error in the target vector 84

RSF 112

RSF values 24

RUSTI 69

S

Shirley 42, 43, 59

Shirley background 74

Simulating Spectra 69

splitter 25

bar 7, 25

window 25

SSI M-Probe Files 129

Starting CasaXPS 6

Synthetic Component 44

T

Tables

quantification 17

Tag field 113

test data 47

text annotation 33

Tile Display 27

Tiles 26

tiles

preferred layouts 27

reducing scrolled list 27

rows & columns 27

Tougaard 42, 43, 66

transition

name 39

transmission correction 24

transmission encoding 146

trend analysis 114

typefaces 29

Index

U

universal cross-section
adjusting 68

V

VAMAS
definition 18
VAMAS Files 17
VGX900 files 125
Vision 1.x/2.x ASCII files 128
Voigt 57, 59, 61

W

Windows 98 7

X

X-Axis 136
x-ray line shape 55

Y

Y-Axis 136

Z

Zoom Reset 14
zoom states 26



© 2001 **Casa Software Ltd.**, 26 Burford Crescent, Wilmslow, Cheshire SK9 6BN, United Kingdom

Tel: +44 (0) 1625 535346 **E-Mail:** info@casaxps.com **World Wide Web:** <http://www.casaxps.com>
