

Ultrathin SiO₂ on Si IV. Intensity measurement in XPS and deduced thickness linearity

M. P. Seah* and S. J. Spencer

Centre for Optical and Analytical Measurement, National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

Received 11 January 2003; Revised 27 March 2003; Accepted 27 March 2003

The procedures for measuring the intensities and for subsequent calculation of the thickness of thermal SiO_2 layers on Si in the range 0.3–8 nm have been evaluated to determine the best measurement protocols. This work is based on earlier work where the measurements for (100) and (111) Si surfaces indicate the need to work at a reference geometry. In the spectra, the Si 2p peaks may be separated clearly into the substrate Si and the overlayer SiO₂ but it is recommended here that, for accuracies better than 1%, the interface oxides are also analysed. The analysis here is for thermal oxides. Oxides grown by other routes may require a modification of this analysis. It is shown that, in evaluating the data to determine the layer thickness, the failure to remove x-ray satellites or the use of a straight-line background will both lead to unacceptable errors that may exceed 5%. On the other hand, if a Shirley background is used consistently for both the peak area analysis and for evaluation of the ratio of intensities for bulk SiO₂ and Si, R₀, the results should be linear over the above range to within ± 0.025 nm. This excellent result includes the non-linearities arising from elastic scattering effects and the data reduction method. Equations are provided, together with a value of R₀ for Mg and Al x-rays, to calculate the oxide thicknesses with the above linearity. In order to determine the oxide thickness accurately, the relevant inelastic mean free paths also must be known. Theoretical evaluations are only accurate to 17.4% and so better values need to be obtained by calibration. This paper provides the infrastructure to do this. Crown Copyright © 2003 Published by John Wiley & Sons, Ltd.

KEYWORDS: attenuation lengths; calibration; gate oxide; layer thickness; linearity; quantification; silicon; silicon dioxide; XPS

INTRODUCTION

The quantitative measurement of film thicknesses by x-ray photoelectron spectroscopy (XPS) has not, in the past, been considered a very accurate or precise procedure. The thicknesses given depend sensitively on the values of the inelastic mean free paths (IMFPs) of the relevant photoelectrons and these are estimated to involve uncertainties of 17.4% for selected elemental solids where direct calculations can be used¹ or a total of 20.4%² if the TPP-2M equation needs to be used³ to calculate this parameter. Added to these uncertainties are problems that, even if the IMFP were accurately known, recent calculations involving both elastic and inelastic scattering show that the attenuation length parameter, deduced from the IMFP, is not constant for all film thicknesses.⁴⁻⁶ Finally, in the important example of SiO₂ layers on Si, which should be both an ideal example and an ideal test of the XPS method, there is one further parameter, R_o (to be discussed later), that is used. The calculated value^{7,8} of this parameter appears to differ from that obtained

*Correspondence to: M. P. Seah, Centre for Optical and Analytical Measurement, National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK. E-mail: martin.seah@npl.co.uk experimentally by a factor of \sim 2,8 and the experimental values in turn range from 0.67 to 0.94. The situation, therefore, looks poor and supports the commonly held view that XPS, at best, may be only semi-quantitative. We shall show, in this and other papers in this series, that XPS can be used to obtain measurements of film thickness with both high precision and high accuracy. There are many other methods for measuring film thicknesses accurately and with precision but XPS has significant advantages in the 0–10 nm thickness range that are important in many areas of microelectronics and nanotechnology.

In the present work we evaluate the errors that occur in different procedures for measuring the XPS intensities, and in using different equations for calculating the oxide thicknesses, that may be found in the literature or that are used in different laboratories. This leads to an evaluation of their errors and also to procedures that can be shown to be linear to ± 0.025 nm for the thickness range 0.3-8 nm. This provides all the parameters to measure the oxide thickness accurately, except for the absolute value of the relevant attenuation length. The attenuation length needs to be obtained by calibration, which is dealt with elsewhere. For the present, calculated values of this parameter are used.



The present assessment is critical for laboratories trying to achieve both high accuracy and precision.

Ellipsometry has been the method of choice for measuring the thicknesses of many transparent layers and it can do this rapidly with detailed maps covering large wafers and with layers up to fractions of a micron thick. However, for thinner layers, the method cannot easily distinguish changes in a required film thickness from changes in composition or unwanted layers such as contamination. This arises because ellipsometry is sensitive to these effects but is not specific to them in a way that analytical methods are. This may cause a bias of up to 1 nm and uncertainties in ellipsometric determinations of thickness. With an analytical method such as XPS, the effects for different chemical species are all separated and the thicknesses for each composition may be determined. Other analytical methods that have the required sensitivity in this range are the ion scattering methods of medium-energy ion scattering spectrometry (MEIS) and Rutherford backscattering spectrometry (RBS). Both of these methods are excellent and generally give good accuracy and good precision but neither method distinguishes the chemical state so the amounts of oxygen in any contamination and the oxygen in silicon dioxide are simply added. Furthermore, at the present time, there is no facility to provide for maps or for efficient laboratory-sized commercial MEIS or RBS units. Both methods, however, do have major roles to play where they have little competition and both have better accuracies, at the present time, than XPS unless the latter is calibrated. It is likely that MEIS will be able to give more accurate composition-depth profiles for the major elements present over the 0-10 nm range and RBS can give excellent composition-depth profiles in the important range up to and beyond 100 nm. However, amongst general analytical methods XPS is unique in that it can be used to measure quantities in specified chemical states and, for thin films, has the capability and linearity for thicknesses down to small fractions of a monolayer with no offset. The accuracy of this linearity and lack of offset will be defined in this paper.

In a previous work⁹ we have shown that XPS can provide efficient maps over large areas with a thickness precision of 0.01 nm at 4 nm thickness. In both that study and most work by ourselves and others analysing the thickness, d_{oxide} , of oxides at surfaces, a very simple equation is used, exemplified by that for SiO₂

$$d_{\text{oxide}} = L_{\text{SiO}}, \cos\theta \ln(1 + R_{\text{expt}}/R_{\text{o}}) \tag{1}$$

where $L_{\rm SiO_2}$ is the attenuation length for the Si 2p photoelectrons in the SiO₂ overlayer, θ is the angle of emission of the photoelectrons from the surface normal, $R_{\rm expt}$ is the measured ratio of the oxide and substrate photoelectrons and $R_{\rm o}$ is the ratio of electrons for these two states from bulk samples. In this work, we shall consider the issues concerning the linearity of Eqn. (1) and the measurements of $R_{\rm expt}$ and $R_{\rm o}$. If it can be shown that Eqn. (1) is linear within, say, $\pm \delta$ for $d_{\rm oxide}$ up to a certain limit and if absolute values of $L_{\rm SiO_2}$ and $R_{\rm o}$ can be established then XPS instruments worldwide could be quantitative for SiO₂ thicknesses, within $\pm \delta$, without the need for certified reference materials. We need to determine the factors contributing to δ and the value of δ . Reference materials may be needed to confirm that all instruments can work to this accuracy, i.e. method validation, but they would not be needed thereafter. If this methodology is valid, it may be extended to other systems measured by XPS. The methodology is restricted, at present, to oxide overlayers because, in these, the analysis is made using the peak areas for the substrate element in the oxide and elemental states. These photoelectrons have very similar energies so the equations are simplified and the number of parameters to be established is reduced to those shown in Eqn. (1).

Although we have discussed the equations as if there were only two peaks, for the case of thermal SiO_2 on Si there are actually five separate chemical states to consider, leading to five peaks after some data reduction. This leads, in effect, to four equations like Eqn. (1) but rather more complex. We shall deal with this in detail later because this can lead to significant variations in practice between laboratories. Thus, the equations are more complex but the principles behind them are as described above.

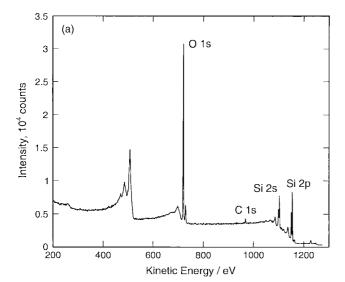
In the next section we consider the issues of measuring R_{expt} and R_{o} together with the linearity of Eqn. (1).

CONSIDERATIONS OF EQN. (1)

Measuring R_{expt} and R_{o} and evaluating d_{oxide}

The peak structure that we are dealing with is shown in Fig. 1(b) as a detail of the widescan in Fig. 1(a). These data are taken from our earlier work8 on thermal SiO2 where, prior to peak fitting, the Mg x-ray satellites have been removed, followed by removal of the 2p_{1/2} spin-orbit split component, which is measured to be 50% of that of the 2p_{3/2} peak and at 0.6 eV higher binding energy. After this, a Shirley background is removed and the peaks are fitted with five peaks representing the Si elemental peak together with four other peaks for SiO₂, Si₂O₃, SiO and Si₂O at 3.3-4.8, 2.48, 1.75 and 0.95 eV higher binding energies than the Si elemental peak, as defined by Himpsel et al. 10 and Keister et al.11 The binding energy of the SiO2 peak shifts with the oxide thickness and, being an insulator, moves with the potentials of local charge control mechanisms.¹² The peak intensities lead to a value for d_{oxide} as described later. Note that oxides formed by deposition and other methods may lead to different peaks appearing and then the present analysis may need slight modification for subsequent application. The measurements here for thermal oxides are recorded at reference geometries for (100) surfaces of 34° from the surface normal in an azimuth midway at 22.5° between the [010] and [011] azimuths, and for (111) surfaces at 25.5° from the surface normal in the $[10\overline{1}]$ azimuth. These particular geometries are chosen because forward focusing of the Si 2p photoelectrons in the crystalline wafer substrate can lead to variations of the deduced d_{oxide} values over a range of $\pm 20\%$ if Eqn. (1) is used with one fixed R_0 value at all geometries. The forward focusing is strongest along the lowindex lattice directions and so the above reference geometries are chosen to reduce these effects.8 For both (100) and (111) surfaces the electrons are thus emitted in the same direction with respect to the Si lattice. In our work, a spectrometer with





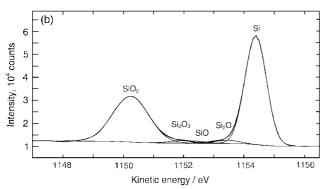


Figure 1. The XPS data for a sample with \sim 2 nm of oxide using Mg x-rays: (a) widescan, (b) Si 2p peaks after satellite subtraction and spin-orbit splitting removal, showing the Shirley background and peak synthesis.⁸

an input lens accepting a cone of electrons of semiangle $\sim\!\!6^\circ$ is used. Smaller acceptance solid angles lead to less angular averaging and may pick up fine structure in the emission distribution, leading to errors in the deduced thicknesses and hence poorer results.⁸

In the determination of $R_{\rm expt}$ for Eqn. (1) many analysts use different methods, which may give poorer results but are simpler and may be more precise or more acceptable in their instruments. We shall investigate these different approaches. In general, if the work is conducted with poorer signal-to-noise ratio, fitting of the three small peaks for the suboxides shown in Fig. 1(b) is uncertain and so analysts adopt an approach fitting just the two main peaks. We call this approach '2P', to distinguish it from '5P' earlier. Depending on the fitting program, the areas of the Si and SiO₂ peaks may change a little and so this is not the same as doing the five-peak fit and simply ignoring three of the intensities. We analyse both options.

Many commercial software systems do not have the spin-orbit removal software and so analysts may fit the Si 2p peak to two peaks and the oxide peak to one peak, totalling three peaks. We call this method '3P'. We analyse both options. Further variations arise through the use of Mg or Al unmonochromated x-rays or monochromated Al x-rays. Analysts using the unmonochromated sources may or may

not have x-ray satellite subtraction software that they trust so they may or may not use it. We analyse both options. Finally, we should consider the background subtraction method. The Shirley background, as used in 5P, is reliable and is not dependent on the precise channels for the end points, but averaging over 1.5 eV at the end points of 96.8 and 107.8 eV binding energy does improve the precision. The Shirley background has no validity in theory but is a recognized and popular method available in all commercial data capture and analysis computer systems. Some analysts prefer to use a straight-line background because that seems reasonable and it may appear not to be readily distinguished from the Shirley background in Fig. 1(b). This is clearly true at a certain level of accuracy but it is not clear what error this will cause. Indeed, with all of the above choices one may suspect that a modest change will arise in the relevant value of R_0 but, because experimental R_o values vary over such a wide range, analysts have accepted the imprecision of the method as being compatible with the imprecision of the parameters involved. Later we will consider these effects experimentally but it is useful first to analyse the straight-line background with a simple model. This is done in a later section.

Measuring Ro

For an overlayer of SiO_2 on Si with no interfacial oxides we have the simple equations

$$I_{SiO_2} = I_{SiO_2}^{\infty} \left[1 - \exp(-d_{SiO_2}/L_{SiO_2} \cos \theta) \right]$$
 (2)

and

$$I_{\text{Si}} = I_{\text{Si}}^{\infty} \exp(-d_{\text{SiO}_2}/L_{\text{SiO}_2} \cos \theta)$$
 (3)

so that

$$\frac{I_{\text{SiO}_2}}{I_{\text{SiO}_2}^{\infty}} + \frac{I_{\text{Si}}}{I_{\text{Si}}^{\infty}} = 1 \tag{4}$$

and because

$$\frac{I_{\text{SiO}_2}^{\infty}}{I_{\text{ci}}^{\infty}} = R_{\text{SiO}_2} = R_{\text{o}}$$
 (5)

we find the useful relation8

$$I_{\text{SiO}_2} = R_0 [I_{\text{Si}}^{\infty} - I_{\text{Si}}] \tag{6}$$

This relation is useful for determining R_o experimentally⁸ from a plot of I_{SiO_2} versus I_{Si} for a number of 'clean' samples. The values of R_o may be obtained conveniently from the gradient of a plot of I_{SiO_2} versus I_{Si} . Alternatively, the pure materials may be used but the surfaces then need to be cleaned equivalently without sputtering.

Effect of a straight-line background

In Fig. 2 we show a simple pair of Gaussian peaks with full widths at half-maxima of 1 eV, sufficiently separated to have a flat background between them. As shown, the peaks have relative intensities of 1:3 and their peaks are separated by 4 eV. The background is a Shirley background with a height in counts at each channel equal to 0.05 of the peak areas above the background at higher energy in counts eV. Figure 2(b) shows more detail of the background with a straight line AB



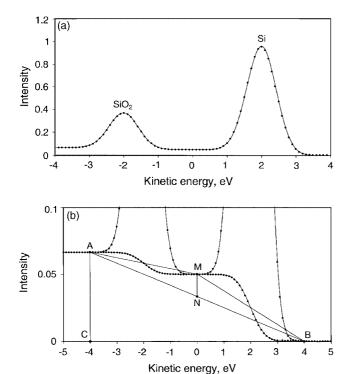


Figure 2. A synthetic spectrum for two Gaussian peaks of relative intensity 1:3 with full widths at half-maxima of 1 eV separated by 4 eV on a Shirley background with an intensity in counts of 0.05 times the peak areas to higher energy in counts eV: (a) the spectrum; (b) detail of the background with the symmetrical straight line fitted.

added symmetrically. It is not clear how a fitting program would fit the two Gaussian peaks with the straight line subtracted but an estimate of the peak areas in this case would be the true peak area plus the areas of the triangles AMN and BMN. The choice of symmetrical conditions here makes the Shirley background symmetrical around AM and BM and the areas of our two triangles equal. If other energies for A and B are chosen, the details of this calculation change but not the general conclusions.

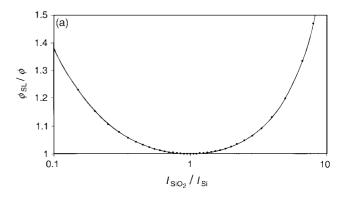
Using the above approach we see that the areas of the oxide peak, $I_{\rm SiO_2}$, and the substrate peak, $I_{\rm Si}$, are increased by using the straight-line background, but if $I_{\rm SiO_2}$ and $I_{\rm Si}$ are equal then the increases are the same and the ratio is unchanged. It is easy to show that if $I_{\rm SiO_2}/I_{\rm Si}$ is given by the ratio ϕ , the measured value of this ratio using the straight-line background, $\phi_{\rm SL}$, is given by

$$\frac{\phi_{\rm SL}}{\phi} = \frac{1 + k(\phi^{-1} - 1)}{1 + k(1 - \phi)} \tag{7}$$

and the thickness, $d_{\text{oxide SL}}$, deduced from Eqn. (1) is related to the true thickness, d_{oxide} , by

$$\frac{d_{\text{oxide SL}}}{d_{\text{oxide}}} = \frac{\ln[1 + (R_{\text{expt}}/R_{\text{o}})(\phi_{\text{SL}}/\phi)]}{\ln(1 + R_{\text{expt}}/R_{\text{o}})}$$
(8)

In Eqn. (7) k is a parameter describing the strength of the Shirley background and is discussed in the next section. Here, $AC = k(I_{SiO_2} + I_{Si})$. Figure 3(a) shows ϕ_{SL}/ϕ as a function of I_{SiO_2}/I_{Si} for a k value of 0.05 eV⁻¹. We see that the error is <5%



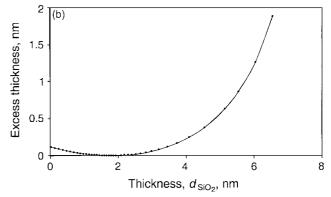


Figure 3. The errors associated with the straight-line background removal for Fig. 2 compared with the Shirley background: (a) the error in $I_{\text{SiO}_2}/I_{\text{Si}}$; (b) the excess thickness deduced for d_{oxide} .

for $0.33 < I_{SiO_2}/I_{Si} < 3$. More important is the effect on d_{oxide} using Eqn. (1). For consistency with later work in this study when using Mg x-rays, we use an L_{SiO_2} value of 2.964 nm⁸ and an R_o value of 0.9329.¹³ The recommended value of R_o may be changed later by up to 2% as a result of ongoing work but we shall retain this value for the present. Figure 3(b) shows the excess thickness for the values of d_{oxide} obtained using the straight-line background minus the true d_{oxide} as a function of the true d_{oxide} used to generate Fig. 2 with the Shirley background. Although the errors in ϕ_{SL} can be large, what we see is a very small effect in d_{oxide} for thicknesses of <5 nm, which leads to a slight non-linearity of result. Above 5 nm, the error grows rapidly. This occurs because I_{Si} is now small and the point M in Fig. 2(b) has fallen below N. In this regime, few analysts would continue with the straight-line and would start using the Shirley background. In discussions with analysts, it is clear that some do adopt this ambivalent view. The straight-line background clearly generates unwanted errors and is not recommended as a reliable approach for the system of SiO₂ on Si. We therefore recommend the Shirley background.

Shirley background

At the present time, we do not know the correct background to remove but it could be argued that, provided the peaks retain their shape from the pure bulk condition, any consistent background subtraction may be valid. We need, therefore, to analyse the consistency of this method. The Shirley background is shown in Figs 1(b) and 2(b). In both cases a single Shirley background has been applied over the



energy range shown. The Shirley background B_i at energy channel E_i may be described by the equation

$$B_i = k \sum_{j=i+1}^{i_{\text{max}}} (N_j - B_j)$$
 (9)

where N_j are the measured counts in the jth channel and i_{\max} is a channel above the peak where $N_{i\max}$ and $B_{i\max}$ are equal. In order to define the constant k at some channel i_{\min} below the peak, we again make $B_{i\min}$ and $N_{i\min}$ equal. Earlier we suggested for Fig. 2 a k value of $0.05 \, \mathrm{eV}^{-1}$. This is a typical value for many materials. However, for studying pure, clean Si and pure, clean SiO₂ we find, for several samples of SiO₂

$$k_{\rm SiO_2} = 0.007 \pm 0.001 \,\mathrm{eV^{-1}}$$
 (10)

and for sputter-cleaned Si

$$k_{\rm Si} = 0.038 \pm 0.001 \,\,{\rm eV}^{-1}$$
 (11)

In terms of our Fig. 2(b), this moves the point M upwards, close to the height of point A. In this case, the use of a single k value, k^* , to determine the peak areas will lead to added intensity for both $I_{\rm Si}$ and $I_{\rm SiO_2}$ to give intensities $I_{\rm Si}^*$ and $I_{\rm SiO_2}^*$, where, for the example of Fig. 2

$$I_{SiO_2}^* = I_{SiO_2} [1 + 2I_{Si}(k_{Si} - k_{SiO_2})/(I_{SiO_2} + I_{Si})]$$
 (12)

$$I_{\text{Si}}^* = I_{\text{Si}}[1 + 2I_{\text{SiO}_2}(k_{\text{Si}} - k_{\text{SiO}_2})/(I_{\text{SiO}_2} + I_{\text{Si}})]$$
 (13)

For these relations we may evaluate the effect that this has on Eqn. (6). If the k values are equal, the effect disappears. If the k values differ, Eqns (4), (5) and (6) are still valid but the measurement values with an asterisk have a more complex relation than given by Eqn. (6). This causes the measured d_{oxide}^* from Eqn. (1) using R_{expt}^* in place of R_{expt} to differ from the true d_{expt} .

Figure 4 shows the result for $I_{\rm SiO_2}^*/I_{\rm SiO_2}^\infty$ versus $I_{\rm Si}^*/I_{\rm Si}^\infty$ using Eqns (12) and (13) with the $k_{\rm Si}$ value given in Eqn. (11) and for $k_{\rm SiO_2}=0$, 0.007, 0.019 and 0.038 eV $^{-1}$. The equivalent plot for the true areas measured with the separate k values would be the straight line as for $k_{\rm SiO_2}=0.038$ eV $^{-1}$. For data taken

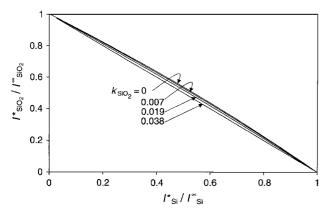
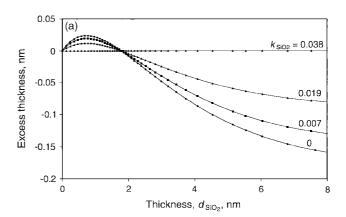


Figure 4. Normalized plot of $I_{\text{SiO}_2}^*$ and I_{Si}^* for a model as in Fig. 2 but where the Shirley background k value for the Si peak is 0.038 eV⁻¹ and a single k^* value is used to measure the peak areas. The four curves are for k_{SiO_2} values of 0, 0.007, 0.019 and 0.038 eV⁻¹. For these plots an R_0 value of 0.9329 is used.



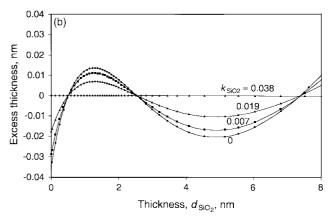


Figure 5. Values of the excess thickness, $d_{\mathrm{SiO}_2}^* - d_{\mathrm{SiO}_2}$, deduced using peak areas with a single value k^* for the k_{SiO_2} values of Fig. 4, and using appropriate R_0 values from fits to data for oxides >1.5 nm thick: (a) absolute values; (b) values after a small adjustment in L_{SiO_2} .

over the whole range of values in Fig. 4 a straight-line fit to the result will generate the correct value of R_o , but for a restricted set of data taken with oxides >1.5 nm thick the value of R_o could be 6% too low.

If we use these slightly revised R_o values from the oxides >1.5 nm thick and deduced using the single k^* value, we may use Eqn. (1) to evaluate d_{oxide}^* from the relevant intensities measured with constant k^* . Figure 5(a) shows the result for each choice of k_{SiO_2} . For thickness <2 nm the added area increases $I_{SiO_2}^*$ significantly, whereas at thickness >2 nm the situation is reversed. In terms of a calibration where L_{SiO_2} is to be deduced from known values of d_{oxide} , the effects would be reduced as shown in Fig. 5(b). Here the effective values of L_{SiO_2} are increased by 2.55%, 2.1%, 1.3% and 0%, respectively, for the four k_{SiO_2} values given above. We see that if $k_{SiO_2} = 0.007$ this correction is likely to be 2.1% and will lead to the data being linear within ± 0.015 nm for thickness >0.3 nm. As noted in the introduction, the calculations for the attenuation length are only expected to be accurate to 17.4% and so a value within 2.1% may be defined accurately only by a calibration using oxides of known thicknesses or thickness differences. If, instead of using one fixed R_0 value for these studies, an effective R_0 is evaluated each time from Fig. 4 for the data for films >1.5 nm thick, we find that, for instance, for $k_{SiO_2} = 0.007$ the value of R_0 falls to 0.8751. This leads to small changes in Fig. 5 with the effective value of the



attenuation length increased only 1.0% and with data linear as above.

In the above we have modelled the spectrum as an effective sum of SiO_2 and Si spectra with their respective backgrounds. The real background between the peaks, however, will not be the simple Si background because we know that the loss spectrum following the peak for a buried material rises more with loss energy than for the same material extending to the surface. This is well established and the loss region up to 50 eV is used to establish the depth of the emitting source. 14,15 No data exist for details in the first few electron-volts of losses for Si but we may expect a small effect that increases linearly with the thickness of the SiO_2 overlayer. This will lead to a proportionate change in the SiO_2 intensity of <1% but an effect that, as above, will be cancelled by the equivalent effect occurring in the determination of R_0 .

Choice of equations for d_{oxide}

In this study we model the surface as shown in Fig. 6. The angle-resolved work of Keister $et~al.^{11}$ shows the interface oxide states to be correctly below the SiO₂ layer. For oxides formed by deposition and other methods, the stoichiometry may not be as shown and extra layers may be involved. Over the oxides, here, is a layer of carbonaceous contamination. ¹⁶ If the Si 2p electrons for the peaks analysed were all at the same energy, we could ignore the effect of the contamination. In fact, the 4 eV separation between the SiO₂ and Si peaks should be included with additional terms to allow for the contamination. Without this, the calculated oxide thickness would be lower than the correct value by \sim 0.25% for 2 nm of the carbonaceous overlayer. Measurements show, however, a small apparent increase in thickness of \sim 0.015 nm per nm of carbonaceous overlayer⁸ that may arise from slight

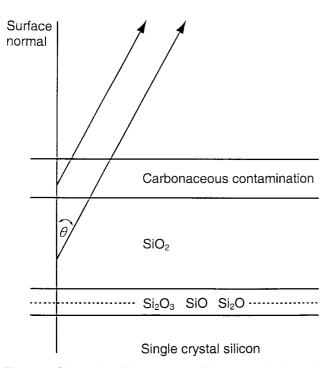


Figure 6. Schematic of the structure of the contamination and of the oxide layers analysed.

changes in the backgrounds following the peaks that lead to a small bias in $R_{\rm expt}$. We ignore this effect here, because typical contamination levels are in the range 0.15–0.3 nm and, from the data given earlier, would only involve an increase of <0.005 nm in the oxide thickness.

It is probable that the suboxide peaks are all from atoms at the interface that are side by side rather than layered one over the other. We shall make this assumption here but the reader can easily show that, for the following evaluation, it makes no significant difference what assumption is used. From before

$$d_{SiO_{2}} = L_{SiO_{2}} \cos \theta \ln \left[1 + \left(\frac{\frac{I_{SiO_{2}}}{R_{SiO_{2}}}}{\frac{I_{Si_{2}O_{3}}}{R_{Si_{2}O_{3}}} + \frac{I_{SiO}}{R_{SiO}} + \frac{I_{Si_{2}O}}{R_{Si_{2}O}} + I_{Si}} \right) \right]$$
(14)

$$d_{\text{Si}_2\text{O}_3} = L_{\text{Si}_2\text{O}_3}\cos\theta\ln\left[1 + \left(\frac{I_{\text{Si}_2\text{O}_3}}{R_{\text{Si}_2\text{O}_3}I_{\text{Si}}}\right)\right]$$
(15)

$$d_{\text{SiO}} = L_{\text{SiO}} \cos \theta \ln \left[1 + \left(\frac{I_{\text{SiO}}}{R_{\text{SiO}}I_{\text{Si}}} \right) \right]$$
 (16)

$$d_{\rm Si_{2}O} = L_{\rm Si_{2}O} \cos \theta \ln \left[1 + \left(\frac{I_{\rm Si_{2}O}}{R_{\rm Si_{2}O}I_{\rm Si}} \right) \right]$$
 with

$$d_{\text{oxide}}$$
 (5P) = $d_{\text{SiO}_2} + 0.75 d_{\text{Si}_2\text{O}_3} + 0.5 d_{\text{SiO}} + 0.25 d_{\text{Si}_2\text{O}}$ (18)

The 5P in parentheses denotes calculation by, in this case, the five-peak method. Equation (18) effectively says that a layer of, say, $d_{\rm Si_2O_3}$ is equivalent to a layer of thickness 0.75 $d_{\rm Si_2O_3}$ of SiO₂ and a layer of thickness 0.25 $d_{\rm Si_2O_3}$ of Si, which is precisely what any analytical method concerned with the amount of oxygen reacted with Si would do. In terms of a physical thickness, the difference between $d_{\rm oxide}$ and $d_{\rm SiO_2}$, measured for 10 thermal oxides from different sources and for different orientation substrates, is 0.118 nm. This is close to half a monolayer of SiO₂ calculated for the bulk density.

There are no calculations or data for values of R and L for the intermediate oxides and so, here, we merely interpolate linearly from SiO_2 to Si. In this way

$$R_{\rm Si_2O_X} = 1 + 0.25X(R_0 - 1) \tag{19}$$

and for Mg x-rays, if R_0 is taken as 0.9329^{13}

$$R_{\text{Si}_2\text{O}_X} = 1 - 0.016775X \tag{20}$$

For Mg K α x-rays, $L_{\rm Si}=2.415\,{\rm nm}$ and $L_{{\rm SiO}_2}=2.964\,{\rm nm}$ so that

$$L_{\rm Si_2O_X} = 2.415 + 0.13725X \tag{21}$$

The above gives the five-peak approach. In the two-peak approach (2P) where the Si spin-orbit splitting is subtracted or where it is not subtracted but the elemental Si intensity is measured using two consistent peaks (3P), we only have data for

$$d_{\text{oxide}}(2P) = L_{\text{SiO}_2} \cos \theta \ln \left[1 + \left(\frac{I_{\text{SiO}_2}}{R_{\text{SiO}_2} I_{\text{Si}}} \right) \right]$$
 (22)



This value would be too high because the denominator in square brackets has no compensation for the interface oxides. This error is offset by ignoring the thickness contribution from these oxides but we shall see later what the final effect on the calculated thickness is. In practice, with different resolution spectrometers some of the interface oxide intensities may be incorporated into the measured $I_{\rm SiO_2}$ and $I_{\rm Si}$ peaks. If we apportion these interface intensities to the two main peaks, we may write from the five-peak analysis

$$\begin{split} d_{\text{oxide}}(5\text{P}^*) \\ &= L_{\text{SiO}_2}\cos\theta\ln\left[1 + \frac{I_{\text{SiO}_2} + 0.75I_{\text{Si}_2\text{O}_3} + 0.5I_{\text{SiO}} + 0.25I_{\text{Si}_2\text{O}}}{R_o(I_{\text{Si}} + 0.75I_{\text{Si}_2\text{O}} + 0.5I_{\text{SiO}} + 0.25I_{\text{Si}_2\text{O}_3})}\right] \end{split}$$

This equation is simpler than Eqns (14)–(18) but appears almost as accurate, as we shall see later. We shall compare these three evaluations of d_{oxide} later to assess the errors that would be obtained in practice.

Linearity of Eqn. (1)

The interface oxides noted above provide an important but small contribution to d_{oxide} . The main contribution arises from I_{SiO_2} and we now need to look at the linearity of Eqn. (1) or Eqn. (22). The lack of linearity arises from the lack of constancy of L_{SiO_2} with d_{SiO_2} , which has been ignored so far. Of course, as noted earlier, the estimated uncertainty in L_{SiO_2} is \pm 0.6 nm but we retain the extra digits to clarify the number traceability, to retain precision and to avoid rounding errors. We shall see later how valuable this is.

Detailed calculations of the value of $L_{\rm SiO_2}/\lambda_{\rm SiO_2}$ for SiO₂ film thicknesses in the range 0–15 nm are given by Powell and Jablonski^{4,5} for the Si 2p photoelectrons of Mg and Al x-rays. They use a geometry where the angle between the x-ray beam and the outgoing electrons is 55° and in which the emission angle is either 0° or 55°. For Mg x-rays, the variation in this ratio for thicknesses in the 0.3–8 nm range is ~17% more than that for Al. The average of the

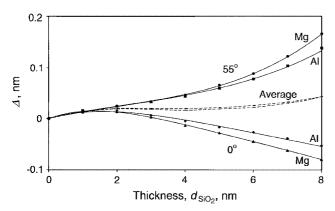


Figure 7. Plot of Δ (the thickness calculated using Powell and Jablonski's⁵ varying values of L/λ minus the thickness calculated using Seah and Spencer's⁸ fixed value of L/λ) versus the oxide thickness for 0° and 55° angles of emission (solid lines) and their average (dashed lines).

0° and 55° results for Mg falls from 0.934 for very thin films to 0.917 at 2 nm and to 0.913 at 8 nm. These are reasonably consistent with our separate calculations,8 which give a value of 0.908 for $\theta \leq 58^{\circ}$. We are not concerned, at present, with the precise value for L_{SiO_2} but with the calculated variation with depth. Figure 7 shows a plot of Δ (the thickness calculated using Powell and Jablonski's varying value of L/λ minus the thickness calculated using Seah and Spencer's fixed value of L/λ) versus the oxide thickness in the range 0-8 nm for 0° and 55° angles of emission for both Mg and Al x-rays. Also shown (dotted) are the averages for these x-rays. As noted earlier, for (100) and (111) surfaces, we recommend measurement at 34° and 25.5° angles of emission, respectively, and these will show effects closer to the average than the extremes calculated in Fig. 7. For these geometries, therefore, it is likely that by using the true value of L_{SiO} , the errors will be <0.04 nm and by using a value of L_{SiO_2} evaluated by calibration any non-linearities will generally be <0.01 nm. In the latter case, a value of L_{SiO_2} that is 0.5% too high will be obtained.

EXPERIMENTAL

All XPS analyses in this work were conducted in a VG Escalab II instrument with a five-channel electron multiplier detector, using 20 eV or 50 eV pass energy and 6 mm input and output slits. Both Mg and Al x-rays were used and these were incident at such an angle that the angle γ between the x-ray direction and the emitted electron direction was, on average, 54.7° (the magic angle).

Samples of both (100) and (111) polished Si wafers were used, with thicknesses of oxide in the range 1.5-8 nm. Samples from six (100) wafers were cut into 1 cm squares bounded by (111) planes defining [110] directions so that the azimuthal zero of angle was along the [110] direction. As recommended elsewhere,8 these samples were analysed at 34° from the surface normal in an azimuth at 22.5° to an edge of the square. Samples from four (111) wafers were cut into equilateral triangles of 15 mm edges bounded by (111) planes defining [110] directions. As recommended elsewhere,8 these samples were analysed at 25.5° from the surface normal in an azimuth given by one of the triangle edges. These geometries are arranged so that the emission directions of the electrons are the same with regard to the lattice structures for both surfaces in an orientation that minimizes the effects of forward focusing.

All samples originally had been cleaned using a 16 h soak in isopropyl alcohol (IPA), followed by a 1 min ultrasonic clean in fresh IPA with the solvent removed by a jet of argon gas. This method routinely removes most of the carbonaceous contamination and leaves the surface with a carbonaceous film ~ 0.12 nm thick without changing the underlying oxide. The samples then were stored in polypropylene containers for 6 months to establish the oxide stability as part of the programme of work on this material. This ten sample set of oxide layers is thus one set from some 40 sets being used in an extensive interlaboratory study under the auspices of the Consultative Committee for the Quantity of Material (CCQM).



RESULTS AND DISCUSSION

The method for measuring peaks is as described earlier. We start with the five-peak method using satellite subtraction, spin-orbit removal and a Shirley background with subsequent peak fitting using Gaussian/Lorentzian product function peaks. Of course, when we discuss subtraction or removal this is done by spectral deconvolution and not by a simple mathematical subtraction. The result of the fitting is shown in Fig. 1(b). From Eqn. (6), a plot of $I_{\rm SiO_2}$ versus $I_{\rm Si}$ for our ten samples gives $R_{\rm o}$ directly from the slope. The samples all have some contamination and it is assumed that this attenuates all Si 2p signals equivalently for all samples. This is not strictly true and leads to a rather larger scatter in the plot than we see for freshly cleaned samples, but the effect is random among these samples.

To simulate common practice, we have analysed the intensities and hence the thickness using seven procedures with the single Shirley background and the same seven procedures with a straight-line background over the same energy range. The seven procedures are as follows: (1) satellite subtraction (SS) followed by removal of Si spinorbit splitting using the Si 2p_{1/2} peak at 50% intensity and 0.6 eV higher binding energy (SO), with the final spectrum analysed using five peaks (5P), calculated via Eqns (14)–(18) and denoted SSSO5P; (2) a repeat procedure but calculated using Eqn. (23); (3) satellite subtraction followed by analysis with six peaks because many users do not have the spin-orbit option and could fit the Si 2p peak with two peaks, calculated via Eqns (14)-(18) and denoted SS6P; (4) a direct analysis with six peaks because satellite subtraction is often omitted, calculated via Eqns (14)–(18) and denoted 6P; (5), (6) and (7) often analysts ignore the suboxides and so procedures (1), (3) and (4) becomes SSSO2P, SS3P and 3P, respectively, and d_{SiO_2} is calculated using Eqn. (22). In each case the fitting is made separately using commercial software that subtracts a single Shirley background and fits the peaks with functions that are a product of Gaussian and Lorentzian functions. These procedures are completed for spectra measured using both Mg and Al K α x-rays.

Figure 8 gives an example fit of ten samples in the thickness range 1.5–8 nm to the I_{SiO_2} versus I_{Si} plot for Mg x-rays using SSSO5P. The gradient gives R_0 as 0.9154.

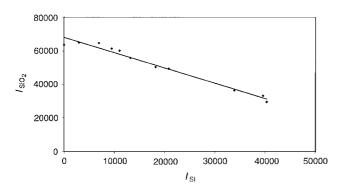


Figure 8. Data for $I_{\rm SiO_2}$ and $I_{\rm Si}$ summed for the five-peak approach after satellite subtraction and spin-orbit removal as shown in Eqn. (23) from measurements using Mg K α x-rays for the 10 samples with oxide thicknesses in the range 1.5–8 nm.

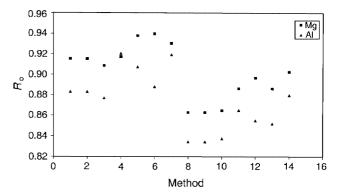


Figure 9. Values of R_0 from the data as in Fig. 8 for the seven procedures (1–7) with Shirley background removal and for the same procedures (8–14) with straight-line background removal for (\blacksquare) Mg x-rays and (\triangle) Al x-rays. The seven procedures are denoted SSSO5P, SSSO5P, SS6P, 6P, SSSO2P, SS3P and 3P (for explanation, see text).

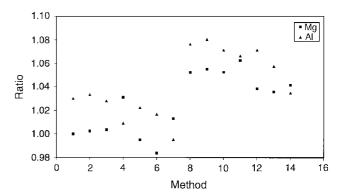


Figure 10. Average values of the ratio of d_{oxide} by one of the 14 procedures divided by d_{oxide} for Mg K α x-rays using the SSSO5P procedure and the individual R_{o} values of Fig. 9 versus the 14 procedures. The values are averages for the 10 samples: (\blacksquare) Mg x-rays; (\blacktriangle) Al x-rays.

Figure 9 shows the R_o values for the seven measurement procedures listed above for both the Shirley and straight-line background methods for Mg and Al x-rays, respectively. The theoretical R_o values for Al and Mg are effectively the same.⁸ Here we see that the average value for Mg is 0.028 ± 0.013 higher than for Al and that the values range from 0.83 to 0.94. This range may explain part of the scatter seen in the literature.⁸ In general, the straight-line backgrounds give R_o values 0.044 lower than those using the Shirley background. We now need to see if this variation in R_0 values is selfcompensating when each separate R_o is used to evaluate d_{SiO_2} . Figure 10 shows the average, for the ten samples, of the ratio of d_{SiO_2} for a given method and x-ray source to the corresponding value for SSSO5P using Mg x-rays for the 14 methods (seven with Shirley background and seven with straight-line background). This ratio is shown as Ratio on the ordinate axis and for these calculations we have used the R_o values from Fig. 9 and a value of 3.448 nm for L_{SiO_2} for Al x-rays.8

From Fig. 10 we see that all of the straight-line background removal results for d_{SiO_2} are higher (typically by 5%) than those for the Shirley background, as predicted in



Fig. 3(b). Methods (4) and (7) ignore satellite subtraction so inevitably, on the low-binding-energy side of the Si 2p peak for the substrate, the intensity has the small $K\alpha'$ satellite, which raises the background and makes the substrate peak too small. This leads to an enhancement in $R_{\rm o}$ that approximately cancels the apparent thickening of the oxide. The averages of the five remaining Mg and five remaining Al measurements (methods 1, 2, 3, 5, 6, etc.) are, respectively, 0.997 ± 0.008 and 1.026 ± 0.007 . In all cases there is no significant difference between SSO5P and SS6P and little difference between Eqns (14)–(18) and Eqn. (23), so that good quality data may be recorded using commercial software systems without access to spin-orbit subtraction.

If, instead of using different values of R_o , we use the single value of 0.9329 for both Mg and Al x-rays determined elsewhere, 13 we get the result of Fig. 11 where the overall scatter is reduced by a factor of 2. Theoretically we expect that the R_o values to be very close. The averages for the five Mg and five Al measurements are now, respectively, 1.001 ± 0.004 and 1.007 ± 0.007 . This excellent consistency indicates that the establishment of a universal value for R_o will lead to an accurate calibration of thickness using either Mg or Al x-rays when accurate values of the attenuation length can be established. These methods can be used with simple equations available for data acquired with commercial spectrometers using commercial data processing software.

Finally, in Fig. 12 we show second-order curves for the ratios of the Mg data using these five methods to the thickness

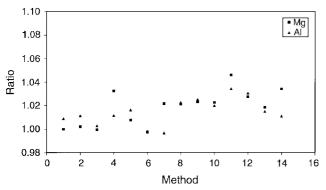


Figure 11. Repeat of Fig. 10 using $R_0 = 0.9329$ throughout.

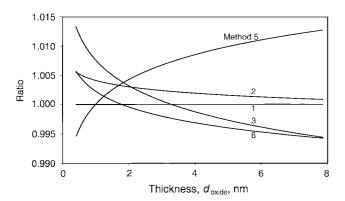


Figure 12. The ratio of d_{oxide} values for methods (2), (3), (5) and (6) for Mg x-rays to method (1) as a function of d_{oxide} for method (1).

calculated using method (1) all data being calculated for $R_{\rm o}=0.9329$. The scatters of the data about the lines are omitted for the sake of clarity but are random and exhibit standard deviations of 0.04%, 0.39%, 0.30% and 0.25% for methods (2), (3), (5) and (6), respectively. The scatter for method (1) in this plot is, of course, zero because the data are all related to method (1). Methods (1), (2) and (3) using all five peaks are better than methods (5) and (6) but even with the simple approach ignoring the suboxides the results are generally consistent to within 1%. These results are extremely good and support the above conclusions.

CONCLUSIONS

Evaluation of the thicknesses of SiO₂ layers on Si by XPS has been analysed in the thickness range 1.5–8 nm. It is shown that the way in which the data are measured and the procedure for subsequent calculation of the thickness lead to variations in the values of the thicknesses obtained. Gross variations occur if arbitrary measurement geometries are used arising from the effects of forward focusing of the Si substrate electrons along the low-index directions. To avoid this, a reference analytical geometry is used for both (100) and (111) surfaces.⁸ In the present experiments, this allows thickness measurements to be repeated after a 6-month interval, with an average increase in thickness of 0.001 nm for the ten samples. Comparisons with other methods show that the XPS repeatability using Mg x-rays and is typically 0.025 nm at one standard deviation.

In this work, simple equations are used to evaluate the thicknesses, involving the attenuation length for electrons in SiO_2 and Si, L_{SiO_2} , and the ratio of intensities for bulk SiO_2 and Si, R_o . These are Eqns (14)–(18), (22) or (23), which involve the elastic scattering effects through an approximation that, in this work, involves errors of <0.04 nm. If the value of L_{SiO_2} is determined by calibration using known thicknesses of oxides, a value 0.5% higher will be obtained and the linearity of the calibration will be valid to ± 0.01 nm.

In evaluating R_0 using plots of I_{SiO_2} versus I_{Si} , the value obtained may range between 0.83 and 0.94 at the reference geometry, depending on the inclusion of the interface suboxides, the equation used and the source of the x-rays. Along crystalline low-index directions R_0 will be lower (e.g. 0.7), but elsewhere it could be higher (e.g. 1.1). The value of R_o measured here for Mg x-rays is 0.028 ± 0.013 higher than that for Al x-rays. The use of a straight-line background gives R_0 values 0.044 higher than those using the Shirley background but, because the straight-line background gives an error of >0.25 nm for thicknesses in the range 4-8 nm, this background is not recommended. By restricting the data to the methods using the Shirley background, for each x-ray source the R_0 values are restricted to a zone of ± 0.02 about the mean. If these values are then used with the relevant equations for calculating the thicknesses, the results using Al x-rays require an extra reduction of 1.5% in the theoretical values of L_{SiO_2} in order to match the thicknesses deduced by Mg x-rays. If a single value of R_0 is used for both Mg and Al x-rays, the difference between the thicknesses reduces to 0.6% and the results for five methods using either source are scattered by 0.6%.



We have assumed here that the Shirley background is correct. There is no theoretical validity for this but, provided that any error introduced remains constant for each of the peaks, the error is removed by the consistent use of the Shirley background in evaluating $R_{\rm o}$. One small inconsistency still remains in that the k values used in the Shirley background may be different for different peaks, whereas in commercial software a single background is used. This leads to absolute errors where the oxide thickness may be underestimated by 0.13 nm at 8 nm if the true value of $L_{\rm SiO_2}$ is used or, if $L_{\rm SiO_2}$ is obtained by calibration, the thicknesses may be correct within ± 0.017 nm. In this case the effective value of $L_{\rm SiO_2}$ is raised by 2%.

The reference method that we have used involves satellite subtraction, removal of the Si $2p_{1/2}$ spin-orbit peak and the resolution of five peaks. The thickness is then calculated using Eqns (14)–(18). Two other methods that give thicknesses within \pm 0.5% of this method are to replace Eqns (14)–(18) by Eqn. (23) or, if spin-orbit subtracting software is not available, to add the extra peak solely for the Si $2p_{1/2}$ intensity for the substrate Si, linked to the Si $2p_{3/2}$ peak at 50% of its intensity and at 0.6 eV higher binding energy. The use of similar analyses, but ignoring the interface suboxides, leads to equivalent thicknesses. Evaluations in which the x-ray satellites are not subtracted can lead to errors of up to 3% on average and up to 10% for the thicker oxides.

It is therefore recommended that analyses for thermal oxides may be made using Eqns (14)–(18) or Eqn. (23) and that the spectra have x-ray satellites removed, a Shirley background removed and the interface suboxides included. Linear results then should be obtainable in the thickness range 0.3-8 nm within ± 0.025 nm, with the

main non-linearities arising from the effects of background subtraction and from variation of attenuation length with oxide thickness. In this work and for these conditions, the values of $R_{\rm o}$ for Mg and Al x-rays are 0.9154 and 0.8830, respectively. In work elsewhere we obtain an $R_{\rm o}$ value of 0.9329 for Mg x-rays. The use of a single value for Mg and Al leads to the best consistency. In this work the values of $L_{\rm SiO_2}$ given in Ref. 8 have been are assumed to be correct.

Acknowledgements

This work forms part of the Valid Analytical Measurement programme of the UK National Measurement System of the Department of Trade and Industry. The authors would like to thank I. S. Gilmore for providing the satellite subtraction and spin-orbit removal software and Dr S. Suzer for material prior to publication.

REFERENCES

- 1. Powell CJ, Jablonski A. J. Phys. Chem. Ref. Data 1999; 28: 19.
- 2. Powell CJ, Jablonski A. Surf. Interface Anal. 2000; 29: 108.
- 3. Tanuma S, Powell CJ, Penn DR. Surf. Interface Anal. 1994; 21: 165.
- 4. Powell CJ, Jablonski A. J. Electron Spectrosc. 2002; 114-116: 1139.
- 5. Powell CJ, Jablonski A. Surf. Sci. Lett. 2001; 488: L547.
- 6. Jablonski A, Powell CJ. Surf. Sci. Rep. 2002; 47: 33.
- Gross Th, Lippitz A, Unger W, Guttler B. Surf. Interface Anal. 2000; 29: 891.
- 8. Seah MP, Spencer SJ. Surf. Interface Anal. 2002; 33: 640.
- 9. Seah MP, White R. Surf. Interface Anal. 2002; 33: 960.
- 10. Himpsel FJ, McFeeley FR, Taleb-Ibrahim A, Yarmoff JA. *Phys. Rev. B* 1988; **38**: 6084.
- Keister JW, Rowe JE, Kolodziej JJ, Niimi H, Tao H-S, Madey TE, Lucovsky G. J. Vac. Sci. Technol. A 1999; 17: 1250.
- 12. Suzer S. J. Phys. Chem B (in press).
- 13. Seah MP, Spencer SJ. to be published.
- 14. Tougaard S. J. Vac. Sci. Technol. A 1990; 8: 2197.
- 15. Tougaard S. J. Vac. Sci. Technol. A 1996; 14: 1415.
- 16. Seah MP, Spencer SJ. J. Vac. Sci. Technol. A 2003; 21: 345.