Anal. Chem. **2002,** *74,* 1104−1110

Whole-Surface Analysis of Semiconductor Wafers by Accumulating Short-Time Mapping Data of Total-Reflection X-ray Fluorescence Spectrometry

Yoshihiro Mori,*[,†,‡] Kenichi Uemura,[†,‡] and Yoshinori Iizuka[§]

Advanced Technology Research Laboratories, Nippon Steel Corporation, c/o Wacker-NSCE Corporation, 3434 Shimata, Hikari, Yamaguchi 743-0063, Japan, and School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Total-reflection X-ray fluorescence (TXRF) spectrometry with no chemical preconcentration, often called "straight-TXRF", is now widely used in the semiconductor industry. The small detection area of TXRF enables mapping measurement of contamination of the semiconductor surface, which is very useful in process characterization. However, the small detection area had been believed to limit rapid whole-surface analysis. Contrary to this general understanding, in this study we demonstrated that a new method, called "sweeping-TXRF", which is essentially short-time multipoint mapping by straight-TXRF, can rapidly provide an average concentration. A considerable problem of this method is the contribution of errors in glancing angle and areal element distribution to the fluorescence. Using statistics, we examined the errors and demonstrated that most of them are canceled and are not significant in actual semiconductor applications. The results of an experiment that measured localized 6×10^{10} atoms cm^{-2} nickel contamination supported the above conclusion. Applying sweeping-TXRF to existing TXRF instruments is easy—the only requirement is a small software modification. We believe that sweeping-TXRF will be utilized for rapid whole-surface analysis in many fields, especially in the semiconductor industry.

Total-reflection X-ray fluorescence spectrometry (TXRF) is widely used in many fields.[1,2] One of the major applications of TXRF is semiconductor analysis, because the features of this method meet the requirements and major improvements specifically for semiconductor analysis have been introduced by some manufacturers.[3−5] Today, both "straight-TXRF" and "vapor-phase decomposition (VPD)-TXRF"[6,7] are used in this field. The former, which directly analyzes the sample surface, is the most common, while the latter combines chemical preconcentration to achieve a higher degree of sensitivity comparable to VPD-inductively coupled plasma mass spectrometry (ICPMS).[8]

Straight-TXRF is an easy and useful application of TXRF. Its mapping capability is one of its major merits. Straight-TXRF can be used to analyze the exact location of contamination if the coordinates are given by appropriate methods such as a light point defect (LPD) counter.[9] If the analyte is isotropic, small-number mapping measurement can reasonably estimate the average concentration of the whole surface. For nonisotropic contamination, however, this method is disadvantageous in conducting whole-surface analysis because the detection area is too small. The typical detection area of TXRF is 1−2 cm^2, while that of an analyte wafer is several hundred square centimeters (e.g., 314 cm^2 for mainstream 200-mm-diameter wafers, 707 cm^2 for state-of-the-art 300-mm-diameter wafers). The integration time of straight-TXRF measurement is typically several hundred seconds for each spot. To analyze the whole surface, we must map more than 100 points, meaning a mapping time of about one day or more for a single wafer. Therefore, the general understanding has been that straight-TXRF is not practical for analyzing the average concentration of a surface.[10]

It should be noted, however, that the detector cannot distinguish between stationary and moving samples. If the sample under the detector is moved continuously, the accumulated signal may reflect the average concentration. We named this method of moving the sample, "sweeping-TXRF". In this paper, we will first show that, in principle, sweeping-TXRF provides an average concentration. Then, using statistics, we will examine the issue of errors in glancing angle and areal distribution and will prove that the errors are not significant in actual application. Finally, we will demonstrate the actual capability of sweeping-TXRF

* Corresponding author. E-mail: moriy@nsce.co.jp. Fax: +81-833-72-3098.
† Nippon Steel Corp.
‡ Present affiliation: R&D Group, Wacker-NSCE Corp., 3434 Shimata, Hikari, Yamaguchi 743-0063, Japan.
§ The University of Tokyo.

(1) Klockenkämper, R. *Total-Reflection X-Ray Fluorescence Analysis*; John Wiley & Sons: New York, 1997.
(2) Hockett, R. S. *Adv. X-ray Anal.* **1994**, *37*, 565.
(3) Nishihagi, K.; Yamashita, N.; Fujino, N.; Taniguchi, K.; Ikeda, S. *Adv. X-ray Chem. Anal. Jpn.* **1991**, *22*, 12.
(4) Utaka, T.; Sako, Y.; Kojima, S.; Iwamoto, K.; Kouno, H.; Atsumi, J. *Adv. X-ray Chem. Anal. Jpn.* **1992**, *23*, 225.
(5) Funahashi, M.; Utaka, T.; Arai, T. *Spectrochim. Acta, Part B* **1997**, *52* 887.
(6) Streckfuss, N.; Frey, L.; Zielonka, G.; Kroninger, K.; Ryzlewicz, C.; Ryssel, H. *Fresenius' J. Anal. Chem.* **1992**, *343* 765.
(7) Funahashi, M.; Matsuo, M.; Kawada, N.; Yamagami, M.; Wilson, R. *Spectrochim. Acta Part B* **1999**, *54* 1409.
(8) Shimazaki, A. *Proceedings of the 179th Meeting of The Electrochemical Society* PV91-9; The Electrochemical Society: Pennington, NJ, 1991; p 47.
(9) *International Technology Roadmap for Semiconductors 1999 Edition*; http://public.itrs.net/.
(10) Fabry, L.; Pahlke, S.; Kotz, L.; Wobrauschek, P.; Streli, C. *Fresenius' J. Anal. Chem.* **1999**, *363*, 98.

10.1021/ac0112061 CCC: $22.00
Published on Web 01/26/2002

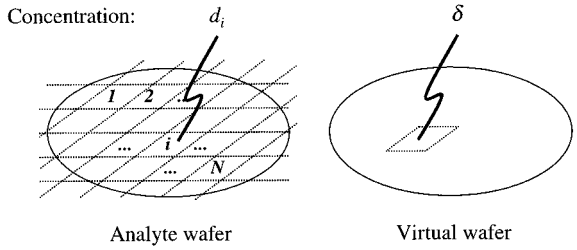

Figure 1. Schematic illustration of the concept of our theoretical treatment.

through an experiment that analyzes localized 6×10^{10} atoms cm^{-2} nickel contamination.

PRINCIPLES AND PROBLEMS

To simplify the mathematical treatment, we assume a procedure of step-by-step movement instead of continuous movement. The basis of the theoretical treatment is shown in Figure 1. The sample wafer surface is divided into N regions, each of which is equal in size to the detector view. The analyte concentration at region i is d_i, and the integration time for each spot is t. If we ignore counting statistics, the fluorescence intensity from point i, w_i, can be described as

$$w_i = hd_it \tag{1}$$

where h is a proportionality factor. Then the accumulated intensity of all points, $\sum w_i$, is

$$\sum w_i = \sum hd_it = ht\sum d_i \tag{2}$$

Here, let us suppose a virtual reference wafer. Its concentration is assumed uniform over the entire surface, and the value is the same as the average concentration of the above analyte wafer. The average concentration δ is defined as

$$\delta = \sum d_i / N \tag{3}$$

When we measure a spot on this wafer for Nt s, which corresponds to the normal straight-TXRF measurement, the fluorescence intensity w_{av} becomes

$$w_{av} = h\delta nt \tag{4}$$

Substituting eq 3 for δ yields

$$w_{av} = h\frac{\sum d_i}{N}Nt = ht\sum d_i \tag{5}$$

Equations 2 and 5 are identical. This means that, in principle, sweeping-TXRF provides the average concentration of the analyte wafer. The same result would also be obtained if the sample was moving continuously.

When we try to put sweeping-TXRF into practical use, however, a major problem concerning the error of the glancing angle must be considered. The accuracy of TXRF measurement strongly depends on the accuracy of the glancing angle.[11] Figure 2 is a typical angle scan profile of a wafer intentionally contaminated by immersion in alkaline hydrogen peroxide solution (IAP) method.[12−15] In this case, only a 0.01° shift in the glancing angle around 0.08° causes a ±14% error in fluorescence, so a high-precision angle adjustment is required for accurate analysis. Because it is difficult to make a perfect angle adjustment mechanically, a feedback control to make the Si Kα intensity from the substrate consistent is commonly employed in commercially available TXRF instruments. This process consumes a certain amount of time before each individual measurement. For example, the time required to accomplish such feedback control in our instrument (TXRF300, Rigaku Industrial Corp.) is ∼70 s for each measuring point. This means that whole-surface sweeping-TXRF will take several hours, because the angle adjustment must be repeated for each measuring point. Figure 3 is a visual representation of total measurement time versus target precision of the glancing angle. This figure assumes 1.67 s × 300 mapping points. Achieving a precision of nearly 0% requires a total measurement time of ∼6 h. A method with such a low throughput cannot be used in actual semiconductor process characterization. Therefore, sweeping-TXRF still seems to be impractical for actual use. Of course, the measurement time would be shortened if we sacrificed angle precision, but such an action has been believed to deteriorate the accuracy of the measured value. In addition, there is another potential error factor that originates in the surface elemental distribution. In the semiconductor fabrication processes, nonuniform contamination caused by such as particle adhesion often occurs. Such nonuniformity may also be a potential error factor in sweeping-TXRF.

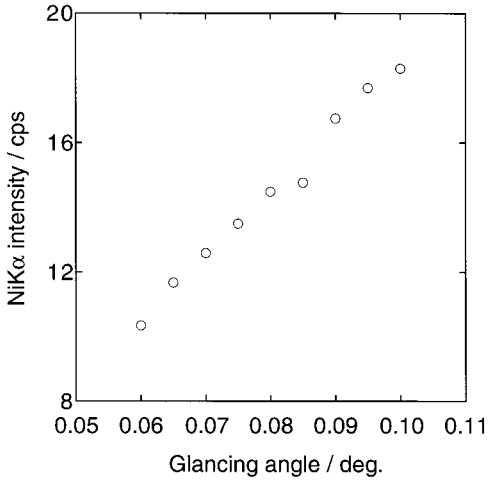

Figure 2. Angle scan profile of an IAP wafer for Ni.

We think, however, that the above two types of errors become small naturally, because independent errors cancel each other out when many data are accumulated, according to statistics. In the next section, we will demonstrate this by means of statistics.

THEORY

Square Root *n* Law. It is known that, if n different variables follow a common probability distribution, the standard deviation

(11) Mori, Y.; Uemura, K. *X-Ray Spectrom.* **1999**, *28*, 421.
(12) Mori, Y.; Shimanoe, K.; Sakon, T. *Anal. Sci.* **1995**, *11*, 499.
(13) Mori, Y.; Shimanoe, K. *Anal. Sci.* **1996**, *12*, 141.
(14) Mori, Y.; Kubota, K.; Shimanoe, K.; Sakon, T. *Anal. Sci.* **1998**, *14*, 275.
(15) Mori, Y.; Uemura, K. *Anal. Sci.* **2000**, *16*, 987.

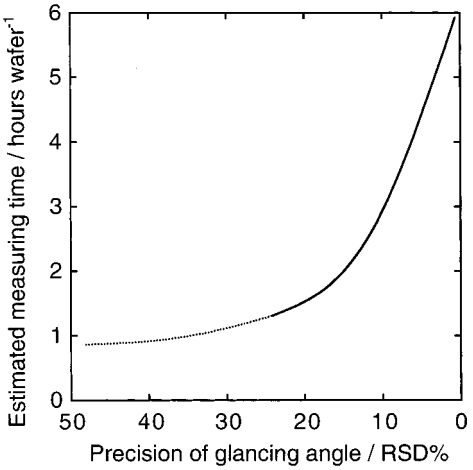

Figure 3. Schematic relationship of sweeping-TXRF measurement time versus glancing angle precision. A measurement of 1.67 s × 300 points is assumed. The solid portion of the line approximately represents the actual measurement time of our instrument, while the dotted portion is a rough estimate.

of their sum is inversely proportional to $\sqrt{n}$, because the individual dispersions cancel each other out. This is sometimes called the "square root *n* law".[16] The mathematical expression of this law is as follows. When *n* variables, X_1, X_2, ..., X_n, follow a common probability distribution with a relative standard deviation (RSD) of ρ, the RSD of their sum, ρ_{sum}, is

$$\rho_{sum} = \rho/\sqrt{n} \tag{6}$$

In sweeping TXRF, a similar cancellation of errors should occur, at least in principle.

Finite Population Correction. Equation 6 is valid for infinite population. For the samples sampled from a finite population, the RSD of their sum, $\rho_{sum,fin}$, is known to be described as follows.

$$\rho_{sum,fin} = \frac{1}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\,\rho \tag{7}$$

The $(N - n)/(N - 1)$ is called "finite population correction". This means that, for a finite population, the sampling error becomes small as the sampling number increases, because the act of sampling decreases the uninvestigated portion of the population. Because all samples are sampled when sweeping-TXRF is conducted (i.e., $N = n$), the finite population correction becomes zero, so that the contribution of nonuniformity of analyte will vanish.

Premises and Deduction. However, the actual situation is not so simple as to allow direct application of the above square root *n* law and finite population correction. In sweeping-TXRF, three types of errors (or dispersions) must be considered at a time—glancing angle, concentration, and counting statistics. These errors may interact with each other, making the overall error more complex. We examined this issue by applying statistical theories.

Here, for simplicity, let us examine step-by-step sweeping-TXRF. In the measurement, the following assumptions can be made: (1) Each individual measurement is statistically independent. (2) The glancing angle of each point follows a common probability distribution. (3) The concentration of each point follows a common probability distribution. These assumptions simplify the statistical treatment.

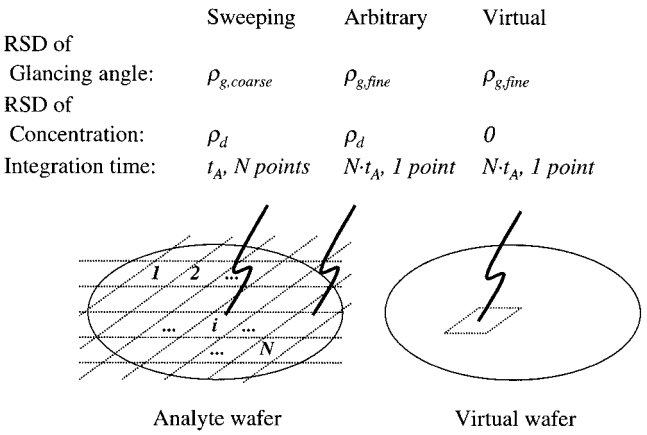

Figure 4. Schematic illustration of the concept of our detailed theoretical treatment.

The expectation of fluorescence intensity, *w*, can be expressed by this equation:

$$w = kgdt \tag{8}$$

where *k* is a constant, *g* is a glancing-angle-dependent term, *d* is areal density (or concentration) of the analyte element, and *t* is integration time. Because fluorescence emission follows a Poisson process, the actual fluorescence, *f*, includes "counting statistics". Then *f* is described as

$$f = \mathrm{Po}(w) \tag{9}$$

where Po symbolizes Poisson distribution.

The glancing-angle-dependent term, *g*, is the change of signal intensity per unit shift of the glancing angle for a certain sample. This term seemingly works as excitation efficiency. It is apparent that the value of *g* depends on the shape of the angle scan profile. It is smallest for a particle-type analyte in which the angle scan profile is relatively flat, and it increases as the depth distribution changes to a film type.[17] Although *g* is not always symmetrical around a given glancing angle (i.e., *g* is not always in direct proportion to the glancing angle), this does not affect our examination, because our statistical treatment does not need to assume any specific probability distribution for *g*, but only requires variance (or standard deviation).

We assume an analyte wafer and divide the surface into *N* regions, as illustrated in Figure 4. Each point has concentration d_i and is measured at a certain glancing angle where the glancing-angle-dependent term is g_i. The average concentration is δ, and the average glancing-angle-dependent term is γ. The relative standard deviation of the concentration and the glancing-angle-dependent term are ρ_d and ρ_g, respectively. From the *N* regions, *n* points are measured for each *t* seconds and the data are

(16) *Toukeigaku Nyuumon*; University of Tokyo Press: Tokyo, Japan, 1991 (in Japanese).

(17) Weisbrod, U.; Gutschke, R.; Knoth, J.; Schwenke, H. *Fresenius' J. Anal. Chem.* **1991**, *341*, 83.

accumulated. On the basis of these assumptions, we deduced the following general expression of the RSD, ρ, for accumulated fluorescence (The detailed procedure of deduction is published as Supporting Information.).

$$\rho = \sqrt{\frac{1}{nk\gamma\delta t} + \frac{1}{n}\left\{\rho_g^2(1+\rho_d^2) + n\frac{N-n}{N-1}\rho_d^2\right\}} \quad (10)$$

Comparison of Errors. By using eq 10, we deduced the equations that express the errors of following three measurement methods under the assumption that the total integration time is the same (Nt_A s):

(1) Arbitrary point measurement: measuring an arbitrarily selected point on the analyte wafer ($n = 1$, $t = Nt_A$). In this method, the glancing angle is finely adjusted ($\rho_g = \rho_{g,fine}$).

$$\rho_{Arb} = \sqrt{\frac{1}{Nk\gamma\delta t_A} + \rho_{g,fine}^2(1+\rho_d^2) + \rho_d^2} \quad (11)$$

(2) Virtual wafer measurement: measuring a point on a virtual wafer that has a completely uniform areal element distribution ($n = 1$, $t = Nt_A$, $\rho_d = 0$). The glancing angle is finely adjusted ($\rho_g = \rho_{g,fine}$).

$$\rho_{Virt} = \sqrt{\frac{1}{Nk\gamma\delta t_A} + \rho_{g,fine}^2} \quad (12)$$

(3) Sweeping-TXRF measurement: measuring all the points of the analyte wafer in a step-by-step manner and accumulating all the data ($N = n$, $t = t_A$). In this method, the glancing angle is not finely adjusted to shorten the measurement time ($\rho_g = \rho_{g,coarse}$).

$$\rho_{Sweep} = \sqrt{\frac{1}{Nk\gamma\delta t_A} + \frac{1}{N}\rho_{g,coarse}^2(1+\rho_d^2)} \quad (13)$$

The common term, $1/Nk\gamma\delta t_A$, which appears in each equation, represents the contribution of counting statistics. The other terms concern N, ρ_g, and ρ_d. Although these equations are somewhat complex, we can provide an intuitive explanation—all the equations simply represent that the independent dispersions are additive. Since the counting statistics, glancing angle, and concentration distribution are independent of each other, their variances are additive, so the contents of each square root symbol are basically in linear combination. The term $1/N$ appearing only in eq 13 means the contribution of the square root n law. In eq 13, the second ρ_d included in eq 11 disappeared, because of the finite population correction. The term $1 + \rho_d^2$ in eqs 11 and 13 is a nonuniformity factor of surface analyte concentration, ranging from 1 to N, as is shown in the Supporting Information.

The differences of these equations are determined by $\rho_{g,fine}$, $\rho_{g,coarse}$, ρ_d, and N. The RSD of arbitrary point measurement, ρ_{Arb}, may be the greatest. This shows the risk of single point measurement—if the concentration dispersion of a given wafer is large, an arbitrarily selected point from the whole surface may hardly represent the average of the sample. The RSD of the virtual sample, ρ_{Virt}, is apparently smaller than ρ_{Arb}, which is a natural conclusion based on the assumptions. What we are interested in here is the difference between ρ_{Virt} and ρ_{Sweep}. However, since both ρ_{Virt} and ρ_{Sweep} depend on the actual values of the variables in each equation, it is not clear at this stage whether their difference is significant. By applying actual values of these variables obtained experimentally, we estimated the actual ρ_{Virt} and ρ_{Sweep} as well as ρ_{Arb}.

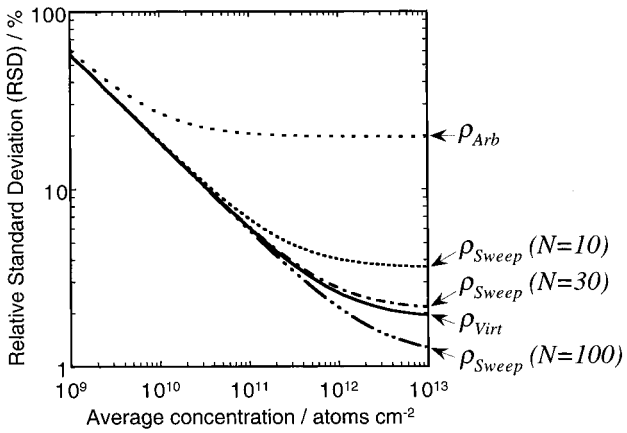

Figure 5. Comparison of calculated ρ_{Arb}, ρ_{Virt}, and ρ_{Sweep} as a function of analyte concentration. Different N are applied for ρ_{Sweep}.

The glancing-angle-dependent term, ρ_g, depends on the slope of the angle scan as mentioned before. In this paper, we used the angle scan data in Figure 2 (IAP wafer, surface layer contamination), which is relatively sensitive to the glancing angle. For single-point measurement, we applied 1.9% to $\rho_{g,fine}$, which is the actual RSD when the glancing angle is accurately adjusted by feedback control in our TXRF instrument (corresponding to 0.8% RSD for the glancing angle). For sweeping-TXRF, 11.2% was applied to $\rho_{g,coarse}$, which is the experimental value when feedback control of the glancing angle is not performed in our instrument (corresponding to 5.2% RSD for the glancing angle). As for the RSD of the surface concentration distribution, ρ_d, 19.6% was applied, which is the highest value for IAP wafers so far.[12] Based on the assumption that a 200-mm-diameter wafer (314 cm^2) is measured by a detector with a view of ~3 cm^2, the default N was set to be 100. Other N, 10 and 30, were also tested for comparison. The constant, k, was determined from the actual measurement of an IAP wafer with our instrument. The integration time, Nt_A, was 500 s, which is typical in semiconductor applications.

The calculation results are shown in Figure 5. The x-axis represents the surface analyte concentration. The y-axis represents the RSD of total signal intensity, being equivalent to the error of each measuring method. The value of ρ_{Arb} for arbitrary point measurement is larger than that of the others in all concentration regions, as was predicted. As the concentration increases, the relative contribution of counting statistics decreases, and ρ_{Arb} converges on 19.7% ($= \sqrt{\rho_{g,fine}^2(1+\rho_d^2) + \rho_d^2}$). This means that ρ_{Arb} will never fall below the RSD of the surface concentration distribution (ρ_d), even if the error of the glancing angle ($\rho_{g,fine}$) is zero. This indicates the risk of treating a single-point measurement as representative of the analyte wafer. From eq 12, ρ_{Virt} in the virtual reference wafer depends on counting statistics and angle dispersion. It does not contain a surface concentration distribution term, because $\rho_d = 0$ as assumed. In this case, as the concentration increases, the relative contribution of counting statistics decreases, so that ρ_{Virt} converges on the value of $\rho_{g,fine}$ (1.9%). As for sweeping-

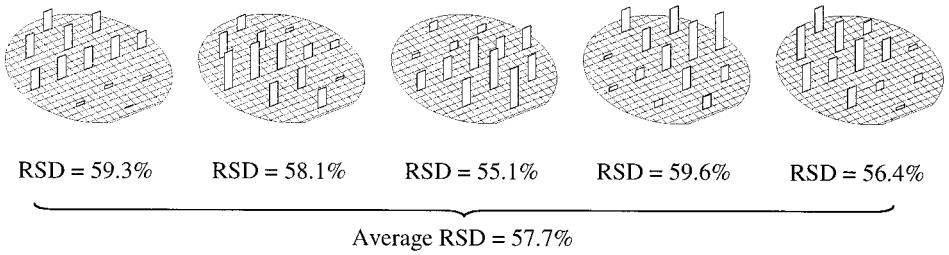

Figure 6. Thirteen-point mapping data of the prepared five samples to show the areal element distribution of each wafer. The integration time is 500 s/point.

TXRF, ρ_{Sweep} strongly depends on the number of measuring spots, N, because of the square root n law. As N increases, the $1/N$ in eq 13 works to make the contribution of $\rho_{g,coarse}$ smaller. Now, we should pay attention to the case where $N = 100$, because N in actual applications is 100 or larger. A value of $N = 100$ is large enough to make the contribution of $\rho_{g,coarse}$ and ρ_d in eq 13 smaller than that of $\rho_{g,fine}$ in eq 12, so ρ_{Sweep} becomes smaller than ρ_{Virt}, as shown in Figure 5. This means, surprisingly, that sweeping-TXRF is superior to virtual wafer measurement. Therefore, despite the fact that the glancing angle is not finely adjusted, sweeping-TXRF is highly accurate in estimating average concentration.

EXPERIMENTAL SECTION

To verify the above theory, we performed some actual measurements. In estimating the errors, both the angle dispersion (g_i) and the concentration distribution (d_i) must be considered. However, because g_i and d_i are independent, it is not necessary to vary the two parameters at the same time, so that we can fix one of the two in this experiment. In this study, we chose to fix g_i. We prepared several samples that have different d_i patterns but have almost consistent values of δ and ρ_d, and measured them under a single g_i pattern. To prepare such samples, we conducted a special IAP[12] procedure, in which only half of each wafer was immersed in a Ni-doped solution. Five wafers were prepared from a single solution at one time. By shifting the orientation of each wafer in the immersion cassette, we prepared five wafers that had different areal concentration distributions while their average surface concentrations as well as areal concentration dispersions were almost the same as each other.

Standard 13-point straight-TXRF mapping (500 s/point) measurements were first conducted to estimate the RSD of each areal concentration distribution. A TXRF300 (Rigaku Industrial Corp.) was used for the TXRF measurement. The excitation source was a rotating anode Au-Lβ (11.44 keV). The applied voltage was 35 kV; the current was 255 mA. The measuring glancing angle was 0.08° where the best detection limit is achieved. The 13-point mapping results are shown in Figure 6. The samples, as expected, have different elemental distributions from each other, while their ρ_d is almost the same, which averages 57.7% in RSD. Then, for each wafer, 1 of the 13 points was randomly sampled to simulate arbitrary point measurement and to calculate ρ_{Arb}.

The wafers were then transferred to the sweeping-TXRF test. To test sweeping-TXRF, we had to decide on the sampling number, N. The value of N depends on the mapping pattern, which determines the distance between neighboring measurement spots. We decided on a spot-to-spot distance of 7 mm, by which the 50% sensitivity circles of the detector just overlap each other. With this mapping pattern, the original number of measuring points was 129. However, 4 of these points were discarded for the convenience of data analysis, and a 125-point mapping pattern was actually used (Figure 7). Built-in mapping software was used without modification for testing sweeping-TXRF. The integration time for each point was 4 s (totally, 4 s × 125 points = 500 s). To make the glancing-angle-dependent term (ρ_g) larger (11.2%), the feedback control function for the glancing angle was intentionally turned off. This also saved measurement time. Under these conditions, the total measurement time was ~90 min/wafer. After the sweeping measurement of each wafer, 125 individual spectra were accumulated to obtain a 500-s integration spectrum, which was then quantified by the usual procedure to estimate the trueness and precision of sweeping-TXRF. The calibration curve for quantification was prepared by measuring a reference standard IAP sample of known concentration with the same procedure.

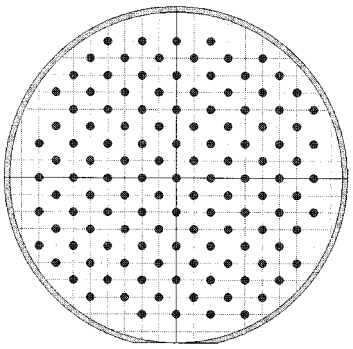

Figure 7. A 125-point mapping pattern used for testing sweeping-TXRF. The lattice lines in the background are drawn at 10-mm intervals.

All the wafers were then analyzed with VPD-atomic absorption spectrophotometry (AAS) to check the average surface Ni concentration of each sample. The entire VPD procedure was performed with a VRC-300T automatic VPD residue collector (SES Corp.). The VPD solution was 50% HF, and the decomposition time was 60 min. The recovery solution was a mixture of 5% HF + 10% H_2O_2, with a typical volume of 105 μL. The sample surface was scanned twice for complete recovery. The recovered solution was analyzed with a SIMAA6000 (Perkin-Elmer Corp.).

RESULTS AND DISCUSSION

Trueness of Sweeping-TXRF. The first and second rows of Table 1 compare the results of both VPD-AAS and sweeping-TXRF. All of the sweeping-TXRF determination values are around the average of reference VPD-AAS value. This demonstrates that

Table 1. Statistical Summary of the Present Experiment

	sample						
	1	2	3	4	5	av	% RSD
VPD-AAS/$\times 10^{10}$ atoms cm^{-2}	4.55	6.41	6.05	6.11	6.69	5.96	13.9
sweeping-TXRF/$\times 10^{10}$ atoms cm^{-2}	5.57	6.15	6.73	6.51	6.05	6.20	7.3
13-point TXRF/$\times 10^{10}$ atoms cm^{-2}	5.21	6.62	6.87	6.41	6.46	6.31	10.2
arbitrary one-point TXRF/$\times 10^{10}$ atoms cm^{-2}	6.03	1.93	4.80	3.03	8.43	4.85	52.7

sweeping-TXRF provides whole-surface average concentration with a high degree of trueness, even when the surface elemental distribution is different and its dispersion is very large.

Precision of Sweeping-TXRF. The second, third, and fourth rows of Table 1 are statistical summaries of the experimental results for TXRF. Despite the large dispersion of concentration on each wafer surface (average of 57.7% as mentioned above) and the difference of localization patterns, the results of sweeping-TXRF are almost the same between the test wafers (7.3% in wafer-to-wafer RSD). The small RSD indicates that the cancellation of errors through multipoint accumulation works well. In addition, surprisingly, the RSD of sweeping-TXRF (7.3%) is smaller than that of a standard 13-point measurement (10.2%). It should be emphasized here that the former essentially requires only a 500-s integration time while the latter requires a much longer integration time (500 s $\times$ 13 points = 6500 s for each wafer). In contrast to the above, the results for arbitrary one-point measurement show a very large wafer-to-wafer RSD (52.7%). This fact, according to the terminology of statistics, strongly demonstrates the risk of a small sampling number to estimate the population mean. Although small-number measurement (1−5 points/wafer) is popular in actual TXRF applications for semiconductor process characterization, this result indicates that such a measurement method should be avoided if the purpose is to estimate whole-surface average concentration.

CONCLUSIONS

Straight-TXRF has not been considered suitable for measuring the average concentration of nonuniform contamination on large-area samples such as semiconductor wafers. In this paper, however, we demonstrated that sweeping-TXRF, which is essentially short-time multipoint mapping by straight-TXRF, is able to provide the average concentration. We examined the errors of sweeping-TXRF by using statistics and showed that the errors are not significant in actual semiconductor applications. The result of an experiment that analyzes localized 6 $\times$ 10^{10} atoms cm^{-2} nickel contamination also supported the above conclusion. Applying sweeping-TXRF to existing TXRF instruments is easy—the only requirement is to make a small software modification.

One of the concerns in actually applying this method is measurement time. The total measurement time was ∼90 min/wafer in this test study. This level of throughput, corresponding to 16 wafers/day, is not too bad when compared with the traditional wet techniques such as VPD-AAS, VPD-ICPMS, and VPD-TXRF. Moreover, it may not be difficult to improve throughput. In the 90-min measurement time, the essential data acquisition time is less than 10 min. The 80-min remainder consists of glancing-angle adjustment, stage movement, and sample height adjustment for each measuring point. These operations could be shortened by preliminarily memorizing the warp of the stage. In addition, the introduction of a continuously moving stage instead of step-by-step one may save even more time. Such improvements can be achieved by modifying the control software. If these improvements are implemented, the throughput of sweeping-TXRF may be more than 100 samples/day, which is much faster than any conventional wet methods.

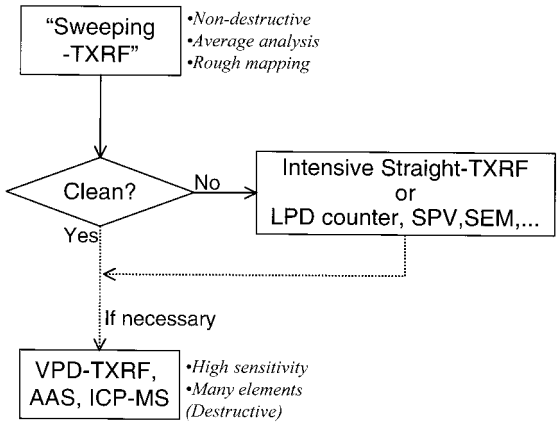

Figure 8. Schematic diagram explaining our proposal for utilizing sweeping-TXRF.

How can we utilize sweeping-TXRF in semiconductor applications? Thus far, straight-TXRF has been recognized as a method of "sampling from the surface". Since rapid measurement of the average whole-surface concentration had not been believed possible, other complementary wet techniques such as VPD-AAS, VPD-ICPMS, and VPD-TXRF have been used. However, the original areal and lateral information of the contaminant is lost after such wet methods are applied, because the VPD procedure is destructive. In contrast, because sweeping-TXRF is nondestructive, we still have the opportunity to apply other methods of analysis afterward. This is the greatest merit of sweeping-TXRF. Figure 8 is a schematic illustration of our proposal for utilizing sweeping-TXRF. The first step is the application of sweeping-TXRF itself. Since this method is automatic and fast, it provides a very rapid feedback of analytical results to the production line than do other methods. Its detection limit, which is typically 2 $\times$ 10^9 atoms cm^{-2} for transition metals such as Ni, is the same as single-point measurement—the detection limit of sweeping-TXRF is of the whole-surface average while that of single-point measurement is of one small spot. In addition, because sweeping-TXRF is essentially multipoint short-time mapping, it provides rough mapping as well as average concentration. In this study, the integration time of each unit point is 4 s, which achieves a detection limit of ∼2 $\times$ 10^{10} atoms cm^{-2}. Although the actual detection limit is considered worse than 2 $\times$ 10^{10} atoms cm^{-2} due to the error of

the glancing angle, the rough mapping may be helpful in process characterization. If any contamination is detected by sweeping-TXRF, we can move to the second step. To investigate the contamination more closely, other methods such as an intensive straight-TXRF mapping, LPD counter,[9] surface photovoltage (SPV),[9] secondary electron microscope (SEM),[18] etc., can be used. This is possible because sweeping-TXRF is nondestructive and preserves the original areal and lateral information. Then, if still necessary, we can go on to a third step. In this step, destructive techniques such as VPD-AAS, VPD-ICPMS, and VPD-TXRF can be applied for a final analysis.

On the basis of the above discussions, we believe that sweeping-TXRF is a promising method for rapid surface analysis. We trust that this method will be utilized in many fields, especially in the semiconductor industry. In addition, the statistical treatment used in this paper may have wide applications in other fields of analytical chemistry.

GLOSSARY

Variables

f	fluorescence intensity
w	expectation of fluorescence intensity
k	proportionality coefficient
g	glancing angle
γ	average glancing angle
d	areal density (concentration) of analyte element
δ	average areal density (concentration) of analyte element
N	number of total measuring points (size of population)
n	number of sampling points (sample size)
t	integration time of each point
ρ	relative standard deviation

Subscripts

Arb	arbitrary single-point measurement
Virt	virtual wafer measurement
Sweep	sweeping measurement
fine	glancing angle is finely adjusted
coarse	glancing angle is not finely adjusted

(18) Hattori, T.; Koyata, S. *Proc. SPIE* **1991**, *1464*, 367.

ACKNOWLEDGMENT

The authors thank Dr. Y. Gohshi (National Institute for Environmental Studies) for his helpful suggestions.

SUPPORTING INFORMATION AVAILABLE

A detailed description of the deduction of the theoretical error equations. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review November 21, 2001. Accepted November 27, 2001.

AC0112061