ELSEVIER

Journal of Chromatography A, 850 (1999) 107–117

JOURNAL OF CHROMATOGRAPHY A

Low-level calibration study for a new ion chromatographic column to determine borate in deionized water

L.E. Vanatta[a,*], D.E. Coleman[b], R.W. Slingsby[c]

[a]*Air Liquide Electronics Chemicals and Services, Box 650311, M/S 301 Dallas, TX 75265, USA*
[b]*Alcoa Technical Center, AMCT-D-10, 100 Technical Drive, Alcoa Center, PA 15069, USA*
[c]*Dionex Corporation, Box 3603, Sunnyvale, CA 94088, USA*

Abstract

In the semiconductor industry, there is interest in determining borate at sub-ppb levels in ultrapure water, since borate is an early breakthrough ion from ion-exchange resin beds. Although dissolved silica is the most common species currently used to monitor the breakdown of the deionization systems, it is thought that borate probably breaks through earlier than silicate. To be of use as an early-warning indicator, borate must be determined at ppt levels. This paper discusses benchtop results with several new column products designed to deliver low-ppt detection limits for boron as borate. The system uses a prototype borate-specific concentrator column that is coupled to an ion-exclusion separator and suppressed-conductivity detection. The acidic eluent, containing mannitol, quantitatively elutes the borate from the concentrator. The analytical separation is performed using a specially designed ion-exclusion column. Data presented are from two multilevel calibration studies. Included is a discussion of detection-limit calculations and recommended formats for reporting results.

Keywords: Calibration; Detection limit; Statistical analysis; Water analysis; Borate; Inorganic ions; Quantitation

1. Introduction

In the semiconductor industry, trace level determination of boron in ultrapure water is critical [1,2]. Boron often is doped into specific locations on the silicon wafers. Consequently, presence of the element at undesired times will result in contaminated products. In addition, tracking of boron levels is used in the water purification plants, since this element is one of the first to break through when resins approach exhaustion. Currently, industry specifications for boron are 50 parts-per-trillion (ppt) (w/w) [3].

Existing procedures utilize the reaction of borate, $B(OH)_4^-$, with diols in water to form polyol borate esters, which behave like monovalent anions. For example, mannitol can be added to boric acid and the resulting 1:1 anion can be titrated with NaOH–phenolphthalein [4]. The ionic character of the complex can be used in an ion chromatographic method; such a procedure is based on the ion-exclusion separation of borate–mannitol, followed by electrical conductivity detection. No concentrator column is used, nor is a time-consuming pre-column derivatization needed. However, sub-ppb levels of boron cannot be reliably detected [5].

In another analytical scheme, the borate-binding diol is formed using color reagents such as chromotropic acid. The borate–chromotropic acid complex

*Corresponding author.
E-mail address: lynn.vanatta@airliquide.com (L.E. Vanatta)

0021-9673/99/$ – see front matter
PII: S0021-9673(99)00263-0

is separated from the excess reagent by reversed-phase HPLC and detected by fluorescence detection [6]. The reaction time is about 40 min. Although this method is an improvement over other published UV and VIS methods, reliable detection at sub-ppb levels still is not possible.

The method studied in this paper utilizes a polyol-based borate concentrator column for preconcentration of borate from deionized water prior to injection onto an ion-exclusion (ICE) analytical column. The ion-exclusion column is compatible with the acid–mannitol mobile phase that is used to elute the borate from the concentrator. After the borate is bound on the concentrator column from deionized water, the eluent's acid hydrolyzes the borate–resin polyol bonds; the eluent's mannitol then is available to form borate–mannitol. The Anion MicroMembrane Suppressor (AMMS)-ICE is used to lower the conductivity contribution from the eluent and to enhance the conductivity of the borate–mannitol analyte for detection by conductivity. The suppressor regenerant is a mixture of tetramethylammonium hydroxide (TMAOH) and mannitol.

The purposes of this paper, then, are: (1) to discuss the columns (a prototype concentrator and an ion-exclusion analytical column) and (2) to present the results of two multilevel calibration studies performed on a benchtop instrument.

2. Experimental

2.1. Materials

Methanesulfonic acid (99%) (MSA) and mannitol (A.C.S. reagent) were purchased from Aldrich (Milwaukee, WI, USA). Tetramethylammonium hydroxide (25%) (TMAOH) was obtained from Air Liquide (Dallas, TX, USA). The eluent was 2.5 m*M* MSA–93 m*M* mannitol; regenerant for the suppressor was 25 m*M* TMAOH–22 m*M* mannitol. The mobile phase was prepared by first diluting 24.025 g of the pure acid to 1000 ml. Then 40.0 g of this solution and 68.0 g of mannitol were diluted to 4000 ml. The eluent was kept under pressure with high-purity helium throughout its life. The regenerant was made by diluting 36.46 g of the TMAOH and 16 g of mannitol to 4000 ml. A 1000-ppm-boron stock standard was prepared from boric acid (99.99%), also ordered from Aldrich. Deionized water (18 mΩ cm) was provided by a point-of-use water purification system (Ahlfinger Water, Dallas, TX, USA). Note: All reagent ware must be plastic, since borosilicate glass is a significant source of borate contamination. In addition, glass-distilled water should not be used for any preparations.

2.2. Apparatus and columns

A Dionex (Sunnyvale, CA, USA) DX500 standard-bore ion chromatograph with a rear-loading 9126 Rheodyne injection valve was utilized for all work. Unless otherwise noted, all instrument modules and consumables were from Dionex. The analytical column was an IonPac ICE-borate analytical column (250×9 mm). A GP40 Gradient Pump delivered the eluent at a flow of 1.0 ml/min.; the pump was fitted with a high-pressure in-line filter housing that contained both 35- and 5-μm filters. Post-column eluent suppression was accomplished with an AMMS-ICE II MicroMembrane Suppressor with regenerant flow at 0.25 ml/min; detection was via a CD20 conductivity detector at an output range of 10 μs. End-line filters were installed on the eluent and regenerant lines. All tubing in the chromatography path (from the outlet of the pump to the exit of the suppressor) was polyether ether ketone (PEEK) (0.005 in. (0.125 mm; 1 in.=2.54 cm) I.D.).

For preconcentrating samples, a prototype IonPac TBC-1 (35×3 mm) concentrator column was used. To load the TBC-1, high-purity helium pressurized the sample bottle in a modified Dionex reagent-delivery module [7] to 5 p.s.i. The delivery tube was 0.063 in. I.D. PTFE tubing. The liquid was pumped through the concentrator by a Dionex GPM-II gradient pump at a flow-rate of 8.0 ml/min until the desired amount had been loaded.

Instrument control and data collection were performed with a personal computer and Dionex PeakNet software. Statistical calculations were carried out using JMP software (SAS Institute, Cary, NC, USA).

2.3. Standards preparation

Polypropylene Mason jars (32 oz) (Bel-Art Products, Pequannock, NJ, USA) were used for all standards and blanks. Polyethylene transfer pipets (Fisher) were used to deliver starting standards and

to add the final few milliliters of the diluting water. A total of 1000.0±0.1 g of solution was made for each preparation.

A stock standard containing 1000 ppm of boron was prepared from the boric acid powder and used throughout the investigation. This standard was made by weighing out 5.72 g of the acid on a Sartorius BP211D analytical balance; dilution to 1000.0 g was accomplished using an XT top loading balance (Fisher). The first mass was recorded to four decimal places and the second to one place. Each day, the stock standard was used to make a 10-ppm solution, which was then used for a 100-ppb preparation. This last solution was utilized throughout the day to make the working standards that were actually injected into the instrument. In each case, the analytical balance was used to weigh out the appropriate amount of the starting standard; the top loading balance was used to dilute to 1000.0 g.

Dilution errors in the daily working standards were estimated by conducting a Monte Carlo simulation. This exercise was based on the upper bounds on the magnitude of weighing error for the scales (0.1 g for the XT balance and 0.0001 g for the analytical balance). In the simulation, weighing errors were randomly drawn from a Normal distribution with mean equal to zero and standard deviation equal to the upper bound. The distribution of these relative concentration errors was found never to exceed 0.1% relative error, which was considered negligible.

At the end of each day, the jars were emptied, rinsed thoroughly, and then filled with deionized water. The same jar was used for the same standard each time. System blanks were prepared in a separate jar, using the same procedure as that used for making the working standards, except that deionized water was added instead of 100-ppb standard. Each day, the standards and system blank were prepared and analyzed in random order. Twelve working standards (all in ppt) were used in this study: blank, 12.5, 25, 37.5, 50, 62.5, 75, 100, 125, 150, 175 and 200.

3. Results and discussion

3.1. Overview of the analytical method

The concentrator packing is a 10-μm, styrene-based resin bearing hydroxyl functional groups. The column specifically concentrates borate and allows subsequent elution with the acid–mannitol eluent in the ion-exclusion separation step. Carbonate also is retained and elutes as a poorly shaped peak with a retention time of 12 to 15 min. If the carbonate is present above ~100 ppb, the concentration of borate will be inhibited by a constant percentage that depends on the carbonate level. Since this method was developed for ultrapure water, the water used to make the calibration-study standards should be well below 100 ppb in carbonate. A similar situation should exist for any ultrapure-water samples that are analyzed using the resulting calibration curve. However, if the carbonate level exceeds the recommended limit, the method of standard addition should be used to determine ppt levels of borate.

The borate–mannitol anion is more conductive than borate itself and is therefore the preferred form for ion-exclusion chromatography. In ion exclusion, the separator resin bears anionic (negative) charge. A weak-acid analyte is partitioned into the resin bed of the separator column to the extent that it is protonated; this mechanism is the basis for retention. As an anion, the analyte is excluded from the separator resin by the anionic charges on the resin, due to electrostatic repulsion. The AMMS-ICE II reduces the background conductivity of the eluent, while maintaining borate as the borate-mannitol anion for detection by conductivity. The suppression reactions can be summarized as follows:

$$\underset{\text{acid eluent}}{H^+A^-} + \underset{\text{regenerant}}{TMA^+OH^-} = \underset{\text{suppressed eluent}}{TMA^+A^-} + \underset{\text{water}}{H_2O} \quad (1)$$

$$\underset{\text{analyte}}{H^+[\text{Borate–Mannitol}]^-} + TMA^+OH^- = \underset{\text{analyte detectable by conductivity}}{TMA^+[\text{Borate–Mannitol}]^-} + H_2O \quad (2)$$

Since the ion-exclusion eluent is not suppressed to water, but to a lower-conducting salt (than the acid), the background conductivity in ion exclusion is directly proportional to the eluent concentration. The background conductivity in this method is approximately 200 μS. Mannitol is used in the regenerant to maintain equilibrium across the suppressor membrane. Although not directly involved in the suppres-

sion reaction, the polyol is necessary to minimize noise.

A typical chromatogram for a 50-ppt standard is shown in Fig. 1a; the corresponding system blank is given in Fig. 1b. The borate eluted in ≈9 min.

3.2. Statistical theory

This section explains the statistical theory that was used in this study. Protocols for detection-limit calculations (DLs) have been described in a previous work [8] and are not repeated here. Quantitation theory is discussed below. (Application of these theories to borate data begins in Section 3.3.)

When a measurement is made, the ideal goal is to report 'the truth' – the actual, exact value. However, in the real world, the true number is almost never known. Instead, the quantity must be estimated; therefore, there will be some uncertainty associated with the number.

Various formats for reporting data are found throughout the literature. In many cases, only the calculated value itself is stated, with no mention of the associated uncertainty. Other reports give the standard deviation of the responses, calculated from

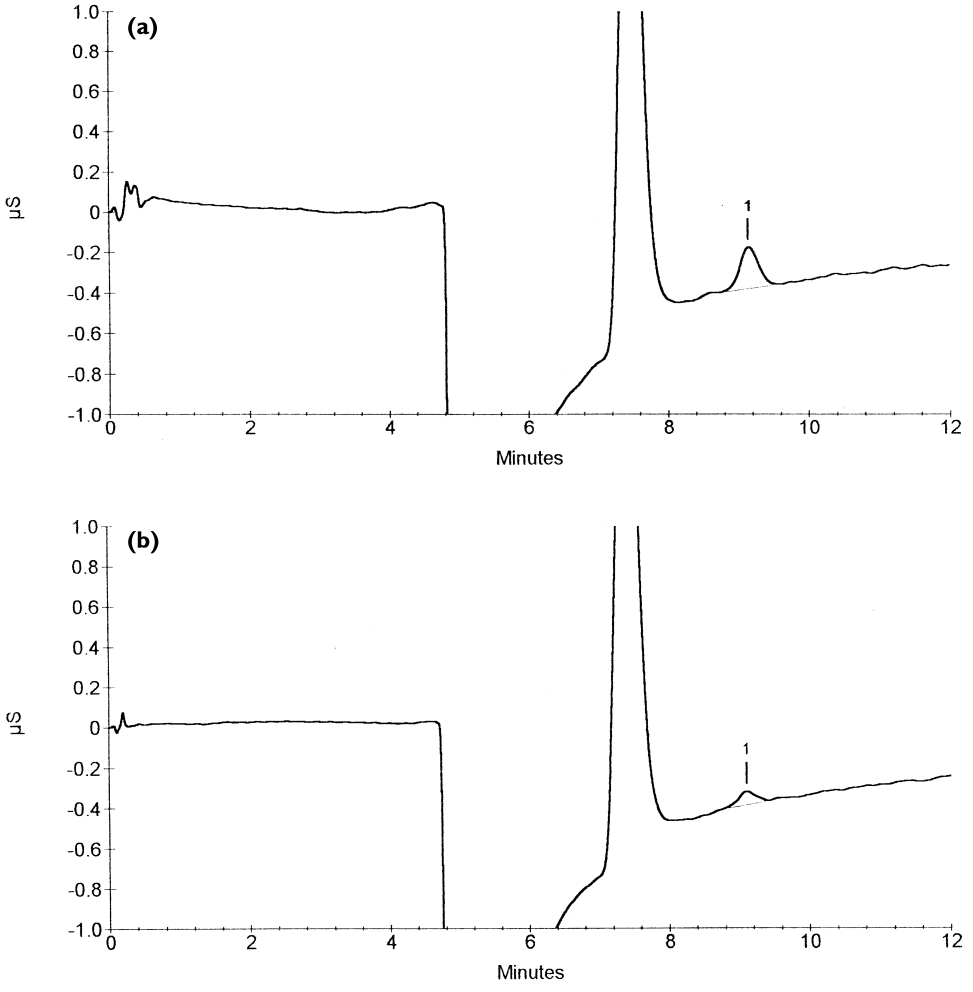


Fig. 1. (a) Chromatogram of a 50-ppt boron standard, where 250 g has been concentrated. (b) A chromatogram of the accompanying system blank (250 g concentrated). Boron (as borate) is peak 1 in both cases.

a given number of replicates at a chosen concentration. This second format ignores the uncertainty connected with the regression process. In other instances, it is not clear how the uncertainty was calculated.

Uncertainty intervals are used to quantify this lack of precision and have a level of confidence (chosen by the analyst) associated with them. These plus or minus intervals depend on: (1) the concentration itself, (2) the standard deviation at that concentration, (3) the value of Student's t for the given degrees of freedom and confidence level and (4) the calibration design. Three types of uncertainty intervals are used: (1) confidence, (2) prediction and (3) statistical tolerance. The first gives the uncertainty in a population parameter. The second type deals with one future measurement. The final interval quantifies the uncertainty in a chosen percentage of 'm' (possibly infinite) future measurements. As one would expect, the width of the interval increases as one goes from (1) to (3).

The ±prediction interval is the most widely used and should be reported with any published measurement, along with the degrees of freedom for the uncertainty. (This format is known as MSD, for measurement–standard deviation–degrees of freedom. The standard deviation is the uncertainty interval that has been chosen.) Then the analyst can decide if the measurement is 'tight' enough for his or her purposes.

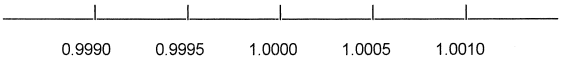
The MSD format is the most desirable, since it conveys the most amount of information about the measurement and its associated uncertainty. However, this protocol is not widely used and may not be accepted readily. An alternate reporting format, then, is to state the measurement along with the relative measurement uncertainty (%RMU), where, at a given level of confidence:

$$\%\mathrm{RMU} = (\pm \text{ uncertainty interval (u.i.)}/\text{concentration}) \cdot 100 \quad (3)$$

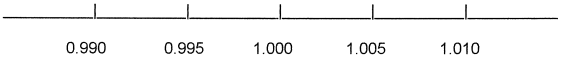
The %RMU is analogous to %RSD. The latter uses the standard deviation of the responses (at a given concentration); %RSD=(standard deviation/concentration)·100.

Still another alternative is available for those preferring significant digits. The concept, though, allows for fractions of a significant digit. The development of the idea is as follows. Assume that the reported number is $v.xyz \cdot 10^k$, where all four digits have information content (hence, one would commonly say there are four significant digits). There is a ±u.i. associated with the number. This interval can be as small as $\pm 0.0005 \cdot 10^k$ (if z has maximum information content; i.e., it is known exactly), but no larger than $\pm 0.005 \cdot 10^k$ (if z has minimum information content). To illustrate why these bounds apply, let $v.xyz \cdot 10^k = 1.000 \cdot 10^k$. (For convenience, 10^k will be dropped in the next two paragraphs.)

To examine the least uncertainty (±0.0005), consider the number line below. If the uncertainty is between 0.9995 and 1.0005, the number would round to 1.0000, always. Thus, there is enough information to warrant five significant digits. Once the uncertainty reaches the minimum of just above ±0.0005, the number will sometimes round to either 1.0010 or 0.9990. Thus, it is not certain if the number is 0.9990, 1.0000, or 1.0010. Now, the 'cut-off line' has just been crossed (i.e., the fifth digit was just lost) and only four significant digits (exactly) are warranted. As the uncertainty continues to increase, the number of significant digits falls below four.



To examine the most uncertainty (±0.005), consider a second number line, given below. If the uncertainty is between 0.995 and 1.005, the number would round to 1.000, always. Thus, there is enough information to warrant four significant digits. Once the uncertainty reaches the maximum of ±0.005, the number will sometimes round to either 1.010 or 0.990. Thus, it is not certain if the number is 0.990, 1.000, or 1.010. Again, the 'cut-off line' has just been crossed (i.e., the fourth digit was just lost) and only three significant digits (exactly) are warranted.



If w=the number of significant digits, then when $w=4.0$, the ±u.i. is ± 0.0005, or $\pm 0.5 \cdot 10^{-3}$, as shown with the first number line. When $w=3.0$, the

interval is ± 0.005, or $\pm 0.5 \cdot 10^{-2}$ (second number line).

In general, then, the $\pm$u.i. will be $\pm 0.5 \cdot 10^{-w+1} \cdot 10^{k}$. Substituting in Eq. (4) gives:

$$\%\text{RMU} = \frac{0.5 \cdot 10^{-w+1} \cdot 10^{k} \cdot 100}{\text{conc.}} \quad (4)$$

Rearranging gives:

$$10^{-w+1+k} = (2 \cdot \%\text{RMU} \cdot \text{conc.}) \div 100$$

Then:

$$\log 10^{-w+1+k} = \log [(2 \cdot \%\text{RMU} \cdot \text{conc.}) \div 100]$$

$$-w + 1 + k = \log [(2 \cdot \%\text{RMU} \cdot \text{conc.}) \div 100]$$

and, finally:

$$w = 1 + k - \log [(2 \cdot \%\text{RMU} \cdot \text{conc.}) \div 100] \quad (5)$$

Eq. (5) also can be stated as: $w = 1 + k - \log (2 \cdot \pm\text{u.i.})$.

Using Eq. (5), a table can be constructed to show the number of significant digits for various %RMU values and ranges of measurements (see Table 1). If the %RMU rises above 50%, w may be negative, depending on the size of the measurement; in all such cases, w will be no greater than 1. Note also that even if %RMU is as low as 1% (a value only rarely claimed, even for an analytical instrument itself), the number of significant digits cannot be greater than 2.7.

If the MSD format is used, or the %RMU or number of significant digits is stated along with the measurement value, the user can decide if the uncertainty is narrow enough for the situation at hand. A quantitation limit then can be set to be the highest acceptable uncertainty for the measurement.

If $\alpha = \beta$, then for any set of data, %RMU is 50% at the H−V DL (see Fig. 2). Assume the $\pm$prediction intervals and calibration line are linear in the region between 0 and conc. = DL. Then all of the line segments labeled x are equal in length. The line of length $2x$ is the H–V DL. Thus, the $\pm$uncertainty interval is DL/2, and the %RMU = (DL/2)/(DL) · 100 = 50%. Therefore, at the H–V DL, one can be assured of having at least zero significant digits; in other words, one is able to assert, with a given level of confidence, that something is there. However, for most users, such a measurement would be too noisy to report a number.

Table 1
Number of significant digits available for selected %RMU values, given a specific measurement value

%RMU	Measurement ($\cdot 10^{k}$)	w
1	1.0	2.7
1	9.9	1.7
5	1.0	2.0
5	9.9	1.0
10	1.0	1.7
10	9.9	0.7
50	1.0	1.0
50	9.9	0.0

3.3. Calibration-study design

Two ICE–borate/TBC-1 column sets (hereafter designated as Set 1 and Set 2) were chosen for the

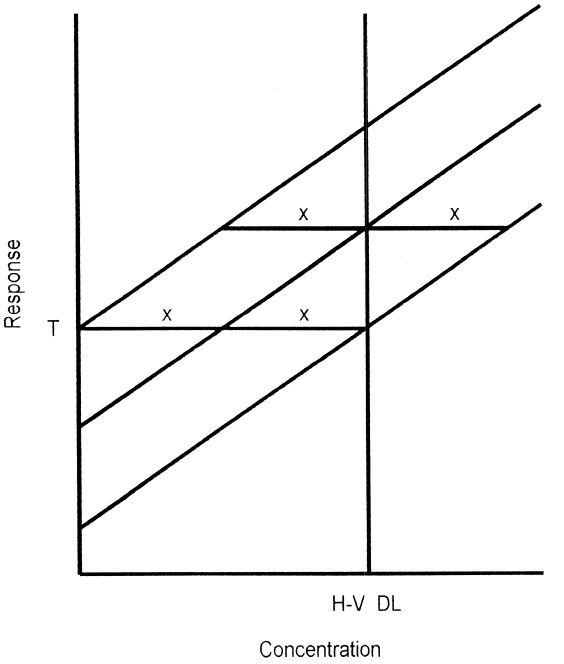


Fig. 2. Plot of response vs. concentration, showing response threshold (T) and Hubaux–Vox detection limit (H–V DL). If $\alpha = \beta$, then all lines x are equal in length and the %R.M.U. is 50% at the H–V DL. See text for details.

study. A straight-line model was proposed for resulting calibration curves, with ordinary least squares (OLS) to be used as the fitting technique. α and β (average proportion of false positives and false negatives, respectively), were each set to 2.5%. For Set 1, a nine-level, equi-spaced design was chosen; six replicates were analyzed at each level (blank, 25, 50, 75, 100, 125, 150, 175 and 200 ppt). A total of 250.0 ± 0.1 g of standard was concentrated each time. For Set 2, 12 levels (blank, 12.5, 25, 37.5, 50, 62.5, 75, 100, 125, 150, 175 and 200) were chosen; the additional three concentrations were added to investigate more fully the behavior at the low end. Eight replicates were analyzed for each level; 160.0 ± 0.1 g of standard was concentrated each time.

3.4. Evaluation of calibration curves

Each set of data first was investigated to see if peak areas trended with time. For each concentration, the response was plotted vs. the day it was obtained. A straight line was fitted through the data and the *P* value of the slope examined. (The null hypothesis: an adequate model was a zero-slope straight line through the mean of the responses.) For Set 1, the 25-, 175- and 200-ppt standards had a significant *p* value (i.e., less than 0.05), which indicated a trend upwards with time (slope was positive). However, with the second column, such a phenomenon was seen only for the 150-ppt standards. With both column sets, the 'water dip' increased with time, thereby causing the baseline around the borate peak to begin sloping upwards. The problem was more severe with the first columns, and probably caused the increased number of peak-area trends.

Using the proposed straight-line model with OLS fitting, a calibration curve was generated for each column set. Each graph was analyzed statistically to see if the chosen model was appropriate. R^2_{adj} values were 0.9916 and 0.9911, respectively. The *p* values for the slope and intercept were less than 0.01 and therefore significant. (Null hypothesis: the same as in the preceding paragraph.) Also, the *P* values for the lack-of-fit test were above 0.05 (0.7716 and 0.2279, respectively), indicating no lack-of-fit problems. (Null hypothesis: there was no lack of fit; i.e., the proposed model was adequate.) When a quadratic fit was tried, the x^2 term was not significant, indicating that this term was not needed in the equation.

To investigate the behavior of the standard deviation of the responses, the residual patterns were examined (see Fig. 3a and b) and the standard deviation of the responses was plotted vs. concentration. For Set 1, the standard deviation appeared to be increasing with concentration (i.e., the spread of the residuals appeared to increase with concentration). For Set 2, some curvature was detected at low concentrations and the standard deviation showed signs of increasing with concentration. The standard-deviation graphs both had significant slopes (i.e., *p* values less than 0.05). This latter finding was crucial, because it indicated that the standard deviation was increasing with concentration, for both column sets. Consequently, weighted least squares (WLS) was needed instead of OLS for fitting the straight line.

With WLS, weights are calculated for each true concentration and are used in the regression. The technique 'favors' the low end, where the precision is best. To reflect the increasing variation with concentration, the resulting prediction intervals will flare out as *x* increases. To determine the weights, the following procedure is used. For each concentration, the regression line from the above standard deviation vs. concentration plot is used to predict the standard deviation. The reciprocal of each result is

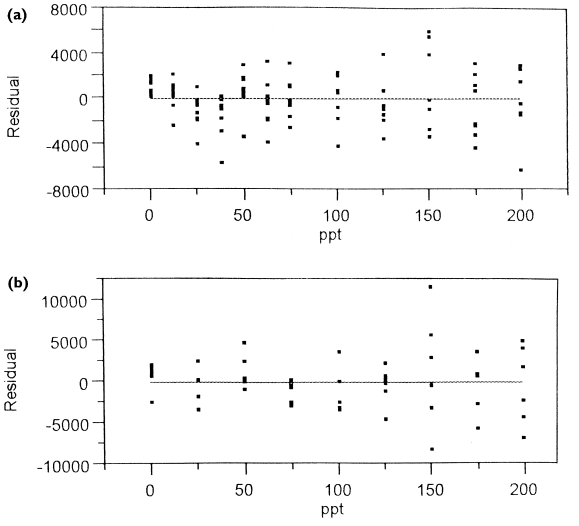


Fig. 3. For straight-line/OLS curve, a plot of residuals vs. concentration; (b) is for Set 1 data and (a) is for Set 2 data.

squared, and then divided by the mean of all the squared reciprocals. These 'normalized' values are the weights.

Once the weights were determined, WLS was used on each data set. The R^2_{adj} values were 0.9934 and 0.9895 for Sets 1 and 2, respectively. The lack-of-fit test gave p values of 0.4058 and 0.0081, respectively.

The latter p value indicated that there was a problem with the fit for Set 2 (i.e., the WLS curve was not adequate to explain the data). To try to resolve the issue, the behavior at low concentrations was examined more carefully (since some curvature had been detected earlier). A plot (Set 2–low) was constructed, using all replicates of the levels from the blank through 75 ppt. A quadratic fit was found to be an adequate model; R^2_{adj} was 0.9664 and the p value for the lack-of-fit test was 0.0645. Note that a decreased R^2_{adj} is to be expected because of the more limited concentration range; the decrease does not necessarily mean a poorer fit. (To compare fits, the prediction intervals should be compared for the low-concentration range. For the WLS curve, the interval was $\pm$ 8.6 ppt at the 12.5-ppt level and $\pm$12.1 ppt at 75 ppt. With the quadratic fit, the intervals were 12.5 and 8.2 ppt, respectively. Thus, the higher-order curve was not a worse fit than the full-range WLS plot.) These findings indicated that piecewise calibration might be appropriate. Therefore, another curve (Set 2–high) was generated, utilizing all replicates of the levels from 75 through 200 ppt. A straight-line model (with OLS) was found to be an adequate fit; the R^2_{adj} was 0.9746 and a LOF p value of 0.7428.

Because of the results for Set 2, the WLS curve for Set 1 was now considered suspect; insufficient data were available for detecting curvature in the low-concentration range. [This assumption was validated by testing a six-replicate version of Set 1 (i.e., the extra concentrations (12.5, 37.5 and 62.5 ppt) in Set 2 were eliminated). A WLS fit again was indicated: a quadratic term was not significant and standard deviation increased with concentration. As with the WLS in Set 1, the p value for the LOF test was high (0.2307), thereby implying that the WLS fit was adequate.] Therefore, a 75–200 ppt plot (Set 1–high) was generated. As with the Set 2 equivalent, a straight-line model (with OLS) was adequate; R^2_{adj} was 0.9761 and the LOF p value was 0.8677.

Finally, the term, g, was calculated for both high-range curves. This unitless term is easily calculated for OLS fits and is similar to a 'significance test' for the slope of the calibration line. The result should be less than 0.1 in order for the prediction intervals to be considered 'locally parallel' to the calibration line. Being 'locally parallel' means that the width of the (horizontal) uncertainty interval at a concentration, X, can be approximated by: $(1/b)$ times the height of the interval at X. The formula for g is:

$$g = [(\mathrm{RMSE})^2(t_{n-2,\,1-[\alpha/2]})^2] \div [(b)^2(S_{xx})]$$

Both calibrations met this criterion.

3.5. Calculation of DL via the Hubaux–Vos method

The statistically sound method of Hubaux–Vos [8] was used to calculate detection limits. This DL can be approximated from a graph obtained using JMP. Via the 'Fit Model' routine, let x = the concentration column, y = the response column and mass = the mass column. Then select 'Run Model', using 'Standard Least Squares'. Under the '$' pop-up menu, select 'Save Prediction Formula' and 'Save Indiv. Confid.'; columns will be created in the data table (for the formula, and for the upper and lower values for the 95% prediction intervals). From the data table's menu, next select 'Overlay Plots' under the 'Graph' option. Let x = the concentration column; let y = the three columns created above. Enlarge and print out the low-end portion of the graph, draw the appropriate lines and read the H–V DL off the x-axis. (See [8] for details on the construction.)

From this graphical approach, the low-end quadratic equation for Set 2 gave a DL of approximately 24 ppt, if $\alpha = \beta = 2.5\%$. (If the technique is applied to the suspect WLS equation for Set 1, a value of 15 ppt is obtained when $\alpha = \beta = 2.5\%$. The discrepancy further emphasizes the need for including the extra low-end concentrations. If these levels are not included, adequate modeling of the region cannot be accomplished and a deceptively low DL will result.)

3.6. Uncertainty intervals

The ±prediction intervals (±p.i.s) were determined at each concentration, using the graph generated in Section 3.5. Because $\alpha = \beta$, the width of the interval is symmetric about the calibration line in the vertical direction (the ±p.i.s are constructed using given values of y). However, the intervals may widen or shrink as concentration increases. Therefore, the width may be asymmetric in the x direction. In this work, the greater of the two values was used for each concentration, if a choice was necessary. Using these widths, %RMU and the number of significant digits were calculated at representative concentrations for each calibration curve. Results are given in Table 2.

3.7. Recalibration options

When recalibration is required, it would save both time and money if fewer levels and replicates could suffice. Since the behavior was quadratic at low concentrations and straight line at high ones, all 12 levels were considered necessary for adequate modeling. To see if fewer replicates would give reliable results, subgroups of the two data sets were analyzed. In each case, four repeats were included. Set 1 had only six replicates total, so only days 1 through 4 were chosen from it; Set 2 had eight replicates total, so two subdivisions (days 1–4 and days 5–8) were used.

Only Set 2 contained enough concentrations for adequate modeling of the low end. The quadratic behavior was detected correctly from the 5–8 set, but was missed with the 1–4 grouping. This finding suggested that the number of low-level replicates should not be reduced, if adequate modeling and a precise detection limit are desired.

With the high-concentration data, all three subgroups gave results (±u.i., %RMU, w) very comparable to those from the parent sets. Therefore, for higher levels, it should be sufficient to recalibrate with fewer than eight replicates.

For the 75-ppt standard (the one that is shared by both the low-end and the high-end curves), the ±prediction interval from the high curves were almost double (~14 vs. ~8) that from the low curves. Therefore, this ‘transition’ concentration cannot be reported out as definitively as can the others.

4. Conclusions

The behavior of the ICE borate/TBC-1 columns was found to be quadratic at low concentrations (i.e., below ~75 ppt), and straight-line with constant standard deviation at levels between ~75 and 200 ppt. These findings illustrate the importance of logical, stepwise evaluations of all proposed calibration curves. Only by invoking all of the available statistical tools (e.g., residuals plots, LOF tests, graphs of response standard deviation vs. concentration) and carefully considering the test results can one make an informed decision. In this case, such an

Table 2
Calculated±uncertainty intervals, %RMU, number of significant digits and degrees of freedom, using graphical approach

ppt	Curve[a]	±Uncertainty interval	%RMU	No. of sig. digits	Degrees of freedom
37.5	Set 2–low	10.0	26.7	0.70	53
75	Set 1–high	13.8	18.4	0.56	34
	Set 2–high	14.0	18.7	0.55	46
	Set 2–low	8.2	10.9	0.79	53
125	Set 1–high	14.4	11.5	1.54	34
	Set 2–high	14.0	11.2	1.55	46
200	Set 1–high	14.5	7.2	1.54	34
	Set 2–high	14.6	7.3	1.53	46

[a] Curve nomenclature: Low refers to the quadratic model found for the blank-through-75-ppt data; high refers to the straight-line model (with OLS fitting) found for the 75-ppt-through-200-ppt data. See text for details.

evaluation showed that the above piecewise calibration would be adequate, if sufficient levels and replicates were analyzed.

If α and β are each set at 2.5%, the following three statements hold. First, a H–V DL of approximately 24 ppt results from the low-level curve. Second, with the high-end curves, %RMU will be ~7% at 200 ppt (~1.5 significant digits). Third, this ion-chromatographic method is acceptable in situations where the DL is specified to be 50 ppt (a common standard in the semiconductor industry). It is up to the user to decide if the uncertainty interval, %RMU and/or significant digits are low enough for reporting a given concentration. Similarly, it is up to the user to decide if the values of the statistics mentioned in Section 3.4 are acceptable.

The column set exhibits a water dip whose width increases with time, so the analyst will have to monitor the situation to judge when the system should be recalibrated and/or when the concentrator should be replaced. Check standards (at least a high- and a low-level each time) should be run periodically and control charts constructed. Two plots (one for the average of the two concentrations and one for the difference between the two, plus optionally one plot for each individual concentration) will provide the tools needed to make replacement decisions. If these precautions are taken, the column set is capable of delivering low-ppt detection limits, where $\alpha = \beta =$ 2.5% and the calibration design is similar to the ones described above.

Acknowledgements

The authors would like to acknowledge Nancy Grams, Ray Maddalone and Robert Gibbons for their useful contributions in developing the interdisciplinary analytical–statistical methods.

Appendix A. Mathematical symbols

α: average probability of false positives.
b: slope of calibration curve.
β: average probability of false negatives.
g: similar to a 'significance test' for the slope of the calibration line. Equals $[(\text{RMSE})^2 \cdot (t_{n-2,\ 1-[\alpha/2]})^2] \div [(b)^2(S_{xx})]$ for OLS fits. Should be less than 0.1 for each curve.
MSD: measurement–standard deviation–degrees of freedom. The desired format for reporting results. The standard deviation is the ±uncertainty interval of interest.
R^2_{adj}: R^2, 'penalized' for each independent variable used in the regression. (R^2 measures the amount of total variation in the response 'explained' by the dependent variable.)
RMSE: root mean square error (often used for sample standard deviation).
%RMU: relative measurement uncertainty, expressed as a percent. Equals (±uncertainty interval ÷ concentration)·100.
w: number of significant digits in a measurement. Equals $[1 + k - log(2 \cdot \pm u.i.)]$.

Appendix B. Terms used

Confidence interval: a pair of limits (an 'upper' and a 'lower') used to bracket the uncertainty in a population parameter.
Confidence level: the likelihood that a hypothesis will not be rejected by a statistical test, when, in fact, the hypothesis is true.
Degrees of freedom: the number of observations in a study minus the number of parameters estimated using those observations.
DL: detection limit. The concentration below which the analytical method cannot reliably detect a response.
H–V: Hubaux–Vos.
Lack-of-fit (LOF) test: a test of the statistical significance of the residual variation that is above and beyond that attributable to pure error.
OLS: ordinary least squares – a fitting technique that minimizes the sum of squares of the residuals.
Prediction interval (p.i.): a pair of limits that bracket the uncertainty in one future measurement.
p value: the probability value associated with a statistical test, representing the likelihood that a test statistic would assume or exceed a certain value, if the null hypothesis is true.
Pure error (also called experimental error): unexplained variation that occurs when experimental

conditions are replicated and repeated experimental runs are performed.

Significant digits: the number of digits (possibly fractional) in a reported value that have information content. Equals $[1 + k - log(2 \cdot \pm u.i.)]$.

Statistical tolerance interval: a pair of limits that bracket the uncertainty in a chosen percentage of '*m*' (possibly infinite) future measurements.

Statistically significant: causing a null hypothesis to be rejected at some accepted confidence level.

Threshold (T): the measurement response below which nothing can be detected reliably. The response cut-off between detection and non-detection.

$\pm$Uncertainty interval ($\pm$u.i.): an interval within which the true measurement is believed to lie, at some chosen level of confidence.

WLS: weighted least squares – same methodology as OLS, except weights are incorporated to account for non-constant response variation.

References

[1] S. Malhotra, O. Chan, T. Chu, A. Fucsko, Ultrapure Water 13 (1996) 22.
[2] S. Malhotra, O. Chan, Pure Water Chem. Conf. 14 (1995) 69.
[3] Texas Instruments specification for deionized water.
[4] R.F. Nickerson, J. Inorg. Nucl. Chem. 32 (1970) 1400.
[5] In-house Dionex method.
[6] M. Oshima, S. Motomizu, Z. Jun, Anal. Sci. 6 (1990) 627.
[7] L.E. Vanatta, J. Chromatogr. 739 (1996) 199.
[8] L.E. Vanatta, D.E. Coleman, J. Chromatogr. 770 (1997) 105.