Anal. Chem. 1994, *66*, 841–849

Quality Self-Monitoring of Intelligent Analyzers and Sensors Based on an Extended Kalman Filter: An Application to Graphite Furnace Atomic Absorption Spectroscopy

Dietrich Wienke,* Theo Vijn,† and Lutgarde Buydens

Catholic University of Nijmegen, Department of Analytical Chemistry, Toernooiveld, 6525 ED Nijmegen (The Netherlands)

A method for on-line quality self-monitoring for automatical operating but drifting analytical sensors is presented. The method is based on an on-line state estimation by the Kalman filter extended by quality control (QC) sampling as known from process monitoring. A linear calibration model with linear drift parameters has been chosen. Compared to conventional approaches, the advantage of the proposed method is that it performs simultaneously calibration and recalibration, detection and correction of drift, and forecasting the expected drift situation, as well as outlier detection and repair. Compared to the existing Kalman filter algorithm, the presented one requires a minimal number of QC samples for updating its parameters. Thus, less recalibrations are necessary in variable time distances adapted to the actual situation in drift, analytical precision, and accuracy. The new procedure has been validated pseudo-on-line in a GF-AAS experiment with artifically enhanced drift. Approximately 1000 samples were analyzed using a continuously (45 h) running and independent working computer driven graphite furnace AAS/autosampler setup.

If the results of an analytical method are not stable within a given time interval, this is often related to the term "drift". Mclelland[1] defined drift as "a gradual change in a quantitative characteristic of a piece of equipment". A similar definition is given by Webster:[2] "a gradual change in the zero reading of an instrument or in any quantitative characteristic that is supposed to remain constant". There are several fields in analytical chemistry such as flow injection analysis[1,3] where drift can be a problem. In graphite furnace atomic absorption spectrometry (GF-AAS) drift can be caused, for example, by small changes of the optical and electrical parameters of the instrument caused by daily temperature changes in a laboratory. Drift can be introduced by the analyst via contamination of the instrument and of the samples. A large source for drift in GF-AAS is the corrosion of the inner surface of the graphite tube caused by reactions during each new excitation. In this way drift can affect the accuracy of an analytical method. Thus, especially in routine analysis, there is a need for a technique that can automatically detect and eliminate drift. One possibility is the use of equally spaced and identically concentrated quality control (QC) standards as used in chemical process monitoring. It is assumed that the absorbance of a QC standard drifts in the same way as that of the sample itself. The obtained time series of absorbances of the QC standards can be evaluated off-line or on-line with several drift detection techniques. Probably the easiest way for drift control is the well-known Shewhart control chart.[9,10] More advanced control charts are the Cusum chart[5] and Trigg's technique.[10,12] The sign tests of Wallis and Moore, where values higher and lower than the median are counted as a criterion for a trend, are alternatives. Further classical trend tests useful for drift detection are the following: Spearman's rank correlation, Cox and Stuart quick sign test, iteration test for randomness of the data, *t*-test of the slope of a fitted regression model, and the mean square successive difference test of Neumann and Moore.[4–10] An implementation of all these drift detection techniques in an user-friendly expert system INTERLAB made for analytical chemists is given in Danzer et al.[11] However, Trigg's technique is especially interesting because it is able to predict the next measurement. Implicitly, this method provides a forecast of the drift. This drift prediction is based on previous measurements and can be used to determine whether the future signal will be in control or not.

However, a disadvantage of the methods described thus far is that they can only detect the drift but they do not correct for it. This deficiency can be overcome using the Kalman filter.[13] This method is based on state space notation that allows recursive parameter estimation if ever the dynamics of the system under study causes the parameters to vary with time. Seelig and Blount first reported the use of the Kalman filter for applications of interest to analytical chemistry.[14] Poulisse and Jansen[15,16] applied the technique to variance

† Present address: IMAG-DLO, P.O. Box 43, 6700 AA Wageningen, The Netherlands.

(1) Mclelland, A. S.; Fleck, A. *Ann. Clin. Biochem.* **1978**, *15*, 281–290.
(2) *Webster's Seventh New Collegiate Dictionary*; G. & C. Merriam Co.: Springfield, MA, 1976.
(3) Thijssen, P. C.; Prop, L. T. M.; Kateman, G.; Smit, H. C. *Anal. Chim. Acta* **1985**, *174*, 27–40.
(4) Doerffel, K. *Statistik in der Analytischen Chemie*; VCH Verlagsgesellschaft: Weinheim, Germany, 1987.
(5) Sachs, L. *Angewandte Statistik—Statistische Methoden und ihre Anwendungen*; Springer Verlag: Berlin, 1976.
(6) Gottschalk, G. *Auswertung Quantitativer Analysenergebnisse (Analytiker Taschenbuch)*; Akademie-Verlag: Berlin, 1980.
(7) Weber, E. *Grundriss der biologischen Statistik*; VEB Gustav-Fischer-Verlag: Jena, Germany, 1972.
(8) Müller, H.; Neumann, P.; Storm, R. *Tafeln der mathematischen Statistik*; VEB Fachbuchverlag: Leipzig, 1973.
(9) Kateman, G.; Pijpers, F. W. *Quality Control in Analytical Chemistry*; Wiley: New York, 1981.
(10) Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. *Chemometrics: A Textbook*; Elsevier: Amsterdam, 1988.
(11) Danzer, K.; Wank, U.; Wienke, D. *Chemom. Intell. Lab. Syst.* **1991**, *12*, 69–79.
(12) Trigg, D. W. *Opt. Res.* **1964**, *15*, 27–33.
(13) Kalman, R. E. *J. Basic Eng.* **1960**, *4*, 35–45.
(14) Seelig, P. F.; Blount, H. N. *Anal. Chem.* **1976**, *48*, 252–259.

0003-2700/94/0366-0841$04.50/0

reduction in drifting processes in analytical clinical routine analysis. Thijssen et al.[3,17-19] proposed a model for on-line drift detection and drift correction. Furthermore, they showed that the calibration and recalibration can be simultaneously performed by the Kalman filter. A limited memory filter was developed by Poulisse and Jansen[20] to correct for the difficulties of abrupt changes in instrumental sensitivity. Reviews of the Kalman filter concerning further applications in analytical chemistry such as multivariate calibration and signal deconvolution have been published by Rutan[21] and Brown.[22]

The aim of the present work is to elucidate the Kalman filter method for the combined tasks of simultaneous (i) drift detection and drift correction, (ii) forecast of drift, and (iii) calibration and recalibration. In fully automated GF-AAS an additional difficulty arises under routine analysis conditions: Each measured sample or standard and the conditions of excitation cause a further erosion of the graphite tube's inner surface. This erosion determines duration and structure of the life cycle of a tube, causing a certain pattern of drift. However, in routine analysis it is desired to analyze a maximum number of samples before substituting the graphite tube by a new one. To achieve this goal, a special Kalman filter is needed that not only corrects on-line for drift but also minimizes the number of recalibration measurements. The present study proposes to use a few single QC standards in place of the entire set of recalibration standards. Further, it proposes to substitute the equidistant placement of recalibration standards by an optimal adapted nonequidistant placement of QC standards or recalibrations. This optimized placement of the standards in the space of time will be a task that has to be controlled and to decided by the Kalman filter. In contrast to the known Kalman filter versions, the proposed extended algorithm should be able to forecast on-line variable time intervals with as large as possible time distances between the QC standards. This information from the QC standards should be used to maximize the distance between the recalibrations. In other words: By means of the new algorithm we tried to maximize the number of analyzed samples between two recalibrations and/or two QC samples within the total life cycle of a graphite tube. The proposed extensions of the Kalman filter have been validated by a 45-h continuously running experiment with artificially enhanced drift. A fully automated computer-driven GF-AAS spectrometer/autosampler setup analyzed on-line Cd traces in aqueous HNO_3 in ~1000 standards and samples.

THEORETICAL SECTION

Detailed explanations of the classical Kalman filter algorithm in connection with special applications to quality control were already given in refs 17–19. For the full mathematics of the Kalman Filter quality control algorithm, the reader is referred to Table 1 of ref 3. Citing this notation, drift is coded in the state vector $\mathbf{x}_{(k)}$ of the Kalman Filter quality control algorithm by

$$\mathbf{x}_{(k)} = \begin{bmatrix} a_{(k)} \\ b_{(k)} \\ \alpha_{(k)} \\ \beta_{(k)} \end{bmatrix} = \mathbf{F}_{(k,k-1)} \begin{bmatrix} a_{(k-1)} \\ b_{(k-1)} \\ \alpha_{(k-1)} \\ \beta_{(k-1)} \end{bmatrix} + \mathbf{w}_{(k-1)} \qquad (1)$$

where α and β describe the drift in slope a and intercept b of a straight calibration line. $k-1$ and k refer to the $(k-1)$th and to the kth time point of measurement. This is different from Kalman filter applications in chemical multicomponent analysis where k refers to the wavelength index.

The transition matrix $\mathbf{F}_{(k,k-1)}$ is in this case

$$\mathbf{F}_{(k,k-1)} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

As structure for the covariance matrix[3] follows from model 1, then directly

$$\mathbf{Q}_{(k)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & q_a & 0 \\ 0 & 0 & 0 & q_b \end{bmatrix} \qquad (3)$$

For readers who are less familiar with the use of the Kalman filter in analytical drift control[3,17-20] but who are familiar with its use for analytical multicomponent analysis (MCA) of mixtures, an additional explanation is given. In terms of MCA, the drift is implemented in the state space vector, in principle, as an "additional component". The time domain, k, plays a role in the drift approach comparable to the wavelength domain in the MCA approach. Thus, in the drift approach, the Kalman filter tries continuously over time k to improve its estimation of the calibration parameters a and b for a single analyte and of the parameters α and β, describing the drift in a and b. In terms of MCA, this principle can be considered like a dynamical analysis of a "two-component" system.

The linear calibration model 1 with linear drift can be extended, for example, to a nonlinear calibration model with linear drift expressed by an extended state vector

$$\mathbf{x}_{(k)} \begin{bmatrix} a_{1,(k)} \\ a_{2,(k)} \\ b_{(k)} \\ \alpha_{1,(k)} \\ \alpha_{2,(k)} \\ \beta_{(k)} \end{bmatrix} = \mathbf{F}_{(k,k-1)} \begin{bmatrix} a_{1,(k-1)} \\ a_{2,(k-1)} \\ b_{(k-1)} \\ \alpha_{1,(k-1)} \\ \alpha_{2,(k-1)} \\ \beta_{(k-1)} \end{bmatrix} + \mathbf{w}_{(k-1)} \qquad (4)$$

where a_1 is the slope for the linear concentration term c and a_2 is the corresponding slope for a second-order concentration term c^2, for example. b corresponds again to the intercept. α_1, α_2, and β characterize the drift in the three regression parameters. However, the present study focuses on the case (1) of linear calibration and linear drift. This classical drift model forms a subpart of the present extended Kalman filter algorithm. Its flow chart (Figure 1) concerns seven main modules: (i) initialization, (ii) estimation of precision, (iii) check of precision, (iv) QC distance, (v) forecast of the calibration curve, (vi) QC sample measurement, and (vii) recalibration.

(15) Poulisse, H. N. J.; Jansen, R. T. P. *Anal. Chim. Acta* **1983**, *151*, 433–439.
(16) Jansen, R. T. P.; Poulisse, H. N. J. *Anal. Chim. Acta* **1983**, *151*, 441–446.
(17) Thijssen, P. C.; Wolfrum, S. M.; Kateman, G.; Smit, H. C. *Anal. Chim. Acta* **1984**, *156*, 87–96.
(18) Thijssen, P. C. *Anal. Chim. Acta* **1984**, *162*, 253–262.
(19) Thijssen, P. C.; Kateman, G.; Smit, H. C. *Anal. Chim. Acta* **1984**, *173*, 265–272.
(20) Poulisse, H. N. J.; Jansen, R. T. P. *Anal. Chim. Acta* **1983**, *151*, 433–441.
(21) Rutan, S. *J. Chemom.* **1987**, *1*, 7–18.
(22) Brown, S. D. *Anal. Chim. Acta* **1987**, *181*, 1–26.

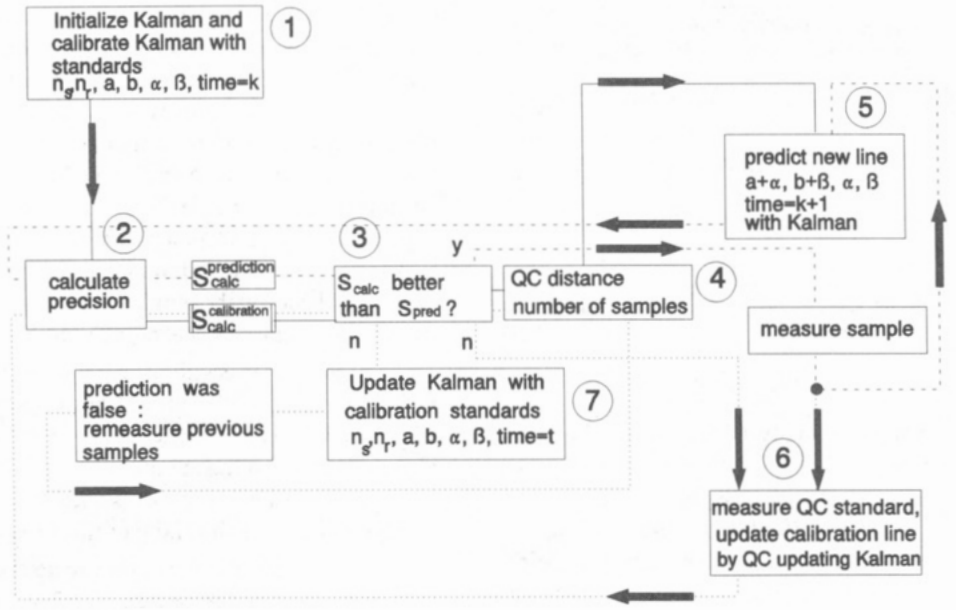

Figure 1. Flow chart for the algorithm of the extended Kalman filter (see theoretical section for detailed explanation).

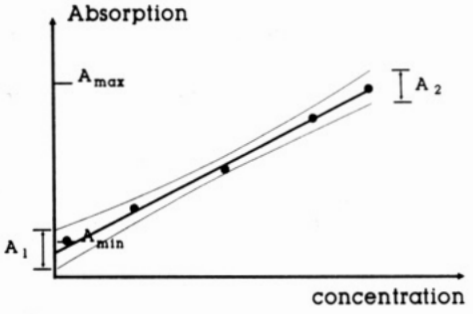

Figure 2. The chosen principle of predefined precision as a comparison of the Kalman filter calculated quantity S_{calc} for the actual considered regression line with a by the spectroscopically defined quantity S_{pred}. The quantity S_{calc} is estimated via eq 5 using the two widths at the end of confidence band for the calibration model.

The principles of the new algorithm will be described briefly. This introductory overview then will be followed by a detailed explanation of our approach.

The extended Kalman filter algorithm (Figure 1) has to be initialized with the measured absorbances of a set of calibration standards. Later on, the running algorithm is updated by the same set of standards (recalibration) that have been analyzed *or by* the measured absorbance of one QC standard. The algorithm itself decides which of both types of standards and at which time point k they have to be measured. Both decisions are based on the previously learned drifting signals of both types of standards and on the actual forecast of the expected drift situation one time unit $(k + 1)$ ahead. The criterion is a predefined precision, S_{pred} (Figure 2). It is used for a decision threshold of whether a new QC sampling or a recalibration should be started or whether the analysis of samples can be continued. A second precision criterion, S_{pred}^{LMS}, is used as a stop criterion in an appended outlier detection and repair module. This module uses the method of least median of squares regression (LMS). The extended Kalman filter (Figure 1) uses this modified LMS procedure after measurement of the set of new recalibration standards and before the updating step (point 7 in Figure 1) to prevent an adaptation of its parameters to the outliers. In the following items, the algorithm is outlined more in detail.

1. The first step in the new algorithm (point 1 in Figure 1) is the initialization of the Kalman filter. The measurement noise $\mathbf{v}_{(k)}$ (Table 1 in ref 3), the system noise $\mathbf{w}_{(k)}$ (Table 1 in ref 3), the number of possible repetitions for the measurement n_r, and the number of standards n_s have to be defined in the beginning. Initial estimations for $\mathbf{v}_{(k)}$ and $\mathbf{w}_{(k)}$ are available, for example, from previous experimental knowledge over the considered or similar analytical systems. The numerical choices for the predefined precision, S_{pred}, for the calibration graph, and for the outlier detection threshold, S_{pred}^{LMS}, are taken from realistic estimations over the reachable analytical overall precision for a given analytical method, including errors in intercept, slope, and model for standards as well as for samples. The number n_r of possible repetitions is determined by constraints in time, costs, and amount of sample that is disposable. Then the first calibration line is calculated, and an estimation of the intercept a and slope b with their drift parameters α and β is obtained (point 1 in Figure 1).

2. After this initial calibration, the actual precision of the obtained calibration line, S_{calc}, is estimated (solid line to point 2 in Figure 1). According to Figure 2, S_{calc} is based on a confidence band model with

$$S_{calc} = 100(A_1 + A_2)/2(A_{max} - A_{min}) \qquad (5)$$

expressing the overall precision of the calibration model as a percentage of the working range of the signal A. The widths A_1 and A_2 of the confidence band at its left- and right-hand end sites were calculated with the corresponding least squares estimator based formula for straight regression lines, as can be found, for example, in refs 4 and 5. After this estimation, S_{calc} is tested against its reference S_{pred} (solid line to point 3 in Figure 1).

3. If the calculated precision S_{calc} is better than the predefined reference value S_{pred}, the maximal number of samples of unknown concentration that can be measured sequentially is determined (solid line to point 4 in Figure 2). That means an estimation of the actually valid distance between two QC standards in the time space. This maximal

Table 1. Definition of the Experimentally Used Maximal Allowed QC Distance as a Function of the Actual Precision S_{calc} of the Kalman Filter Predicted Calibration Line[a]

size of S_{calc}/ %	QC distance/time units
$0 \le S_{calc} \le 1$	20
$1 \le S_{calc} \le 2$	12
$2 \le S_{calc} \le 3$	8
$3 \le S_{calc} \le 4$	4
$4 \le S_{calc} \le 5$	2
$S_{calc} > 5$	1

[a] Numbers of analytical samples to be analyzed between the measurement of two adjacent QC standards in the space of time.

number of samples that can be analyzed sequentially between two QC samples depends on S_{calc}. An example is given in Table 1. The number reflects, in principle, how far the Kalman filter can be trusted to predict α and β for future drift situations. To avoid too many samples of unknown concentration being measured, the maximum size of the QC standard distance was restricted (Table 1). The danger is that unexpected drift events can happen between two QC samples with time. Such events can be outliers or strong nonlinear drift that is unpredictable with the chosen linear drift model.

4. The system is ready now to analyze samples. Before each sample analysis, a prediction is made of the calibration line one time unit $(k + 1)$ ahead (solid line to point 5 in Figure 1) using the Kalman filter prediction formula (Table 1 in ref 3). That means that the regression parameters a and b are corrected with the drift parameters, α and β, that were predicted one time unit $(k + 1)$ ahead. By this operation, the calibration line is set for time $t_{(k+1),pred}$. Each sample will be analyzed in this way by its "own" calibration line adapted to the predicted drift at that time point. A new experimental estimation of the line parameters, a and b, and of the drift parameters, α and β, is received later (point 7 in Figure 1) with the recursive update equations as given, for example, in eqs 15–19 in ref 17 of the Kalman Filter algorithm.

5. The precision, S_{calc}, of each calibration line has been individually predicted in (point 5 in Figure 1) using the model (Figure 2, eq 5). Then S_{calc} is tested against S_{pred} (thick dotted line from point 5 to 2, 3 in Figure 1). If the precision S_{calc} of the predicted individual calibration line is out of control (point 3 in Figure 1), an updating of the Kalman filter with one QC standard is tried (thick dotted line point 3 to 6 in Figure 1). If this update provides a precision, S_{calc}, that is better than S_{pred} (thin dotted line point 6 via 2 to 3 in Figure 1), then the next sample can be analyzed (thick dotted line point 3 to 5 in Figure 1). If not, then refer to item 7.

6. After the analysis of all the samples according to the predicted QC distance, the next QC standard is measured (dotted line from thick dot to point in Figure 1). The Kalman filter is updated with the QC standard, and a new estimate of the parameters within the state vector is received. The precision, S_{calc}, of the experimentally corrected calibration line is compared (thin dotted line from point 6 via 2 to 3 in Figure 1) with the predefined precision, S_{pred}. If S_{calc} is in control, the next group of samples can be evaluated (solid line from point 3 via 4 to 5 in Figure 5). If not, refer to item 8.

7. However, it happens that S_{calc} is still out of control after the update of the Kalman filter by the scheduling QC standard (item 5). In practice, drift changes sometimes faster then predicted. The estimations of the old drift parameters and the QC distance are poor in such a situation and the last measured group of samples has to be analyzed again with a completely recalibrated Kalman filter (thin dotted line from point 3 via 7 to 4 and 5 in Figure 1). In this situation, the advantage of the variable QC distance becomes clear. In case of poor precision, S_{calc}, followed by poor predictions for sample concentration and drift development, the QC distance will be small. In this way, the number of samples for a reevaluation stays small. This saves accuracy but also time and costs.

8. A complete recalibration (thin dotted line from point 3 via 7 to 4 and 5 in Figure 1) has to be performed in item 6 if an update with a single QC standard (point 6 in Figure 1) provided no significant improved precision S_{calc}. The entire set of recalibration standards is used in place of a single QC standard to update the Kalman filter (eqs 15–19 in ref 17).

The main difference from the previous developed Kalman algorithms is that the present extended algorithm uses in addition to the calibration standards so-called quality control standards for the updating of its parameters. In a traditional process monitoring system, these QC standards would be distributed equidistant in time over the samples. In contrast to that, the extended Kalman filter tries to use (i) a minimum number of QC standards, (ii) a minimum number of recalibration steps between the samples, and (iii) an optimum position of both types of standards in the space of time. In other words: Additionally to the dynamic estimation of the calibration model, a and b, for a single analyte and to the dynamic estimation of the drift by α and β, the algorithm performs a self-control by placing a minimized number of QC standards at optimal positions between the recalibration standards and the samples in the space of time.

Outliers versus Drift. Outliers during process monitoring are undesired discontinuous events. The adaptive Kalman filter algorithm cannot have strong outliers, especially during the recalibration step. Estimation of the line parameters a and b must not be influenced by outlier recalibration standards. Otherwise the extended Kalman filter would try to adapt a, b, α, and β to such events. Therefore, an outlier detection and outlier repair procedure has been incorporated as a submodule into the recalibration step of the extended Kalman filter algorithm. This submodule is based on the work of Rousseeuw[23,24] in least median squares regression and experiences reported by Massart et al.,[25] Rutan and Carr,[26] and Hu et al.[27,28] about the use of this LMS for outlier detection and robust calibration. However, to find a suitable stop criterion of the LMS-based outlier rejection procedure, first the criteria proposed by Rutan and Carr (p 137 of ref 26) and Hu et al. (eqs 9 and 10 of ref 28) were applied. These criteria are statistically based, having the advantage of being independent from the individual calibration problem. Unfortu-

(23) Rousseeuw, P. J. *J. Am. Stat. Assoc.* **1984**, *79*, 871–879.
(24) Rousseeuw, P. J. *Robust Regression and Outlier Detection*; John Wiley: New York, 1987.
(25) Massart, D. L.; Kaufman, L.; Rousseeuw, P. J., Leroy, M. *Anal. Chim. Acta* **1986**, *187*, 171–179.
(26) Rutan, S. C.; Carr, P. W. *Anal. Chim. Acta* **1988**, *215*, 131–142.
(27) Hu, Y.; Smeyers-Verbeke, J.; Massart, D. L. *J. Anal. At. Spectrosc.* **1989**, *4*, 605–613.
(28) Hu, Y.; Smeyers-Verbeke, J.; Massart, D. L. *Chemom. Intell. Lab. Syst.* **1990**, *9*, 31–44.

nately, for most of the distribution of points along the calibration graph, the outlier selection procedure did not stop automatically as expected. For linear graphs with narrow confidence bands and for numerous experimental calibration graphs, this limitation has been observed also.[29] However, in practice no reasons exist to eliminate more outliers then necessary, providing eventually a practically not reachable and too optimistically estimated overall precision of the calibration graph. That is why, as an alternative stop criterion, in the previous section the already used principle of predefined precision (Figure 2 and eq 5) was studied too, for the outlier rejection submodule. The user has to predefine S_{pred}^{LMS} as a decision threshold before starting the drift control. The size of S_{pred}^{LMS} depends, like the size of S_{pred}, on knowledge over the overall precision of the analytical procedure used. That means basically the introduction of additional chemical knowledge about the monitored calibration graph in contrast to a pure statistically based stop criterion. The size of a second criterion, S_{pred}^{LMS}, of predefined precision as a vigilance parameter for LMS-based outlier rejection has to be chosen less strictly than for the previously described S_{pred}. In this way "strong" outliers can be distinguished from drift that is considered more as a "soft outlier event". Together with this new stop criterion, S_{pred}^{LMS}, another modification has been introduced into the LMS procedure. The pure outlier detection and rejection procedure has been extended by an outlier repair step. Outlier repair means a single rejected outlier initiates an automatical remeasurement of three new replicates, averages the three measurements, and adds this new point to the calibration graph. After this, the LMS outlier procedure is repeated until the calculated precision S_{calc}^{LMS} is better than the predefined precision S_{pred}^{LMS}.

EXPERIMENTAL SECTION

The decisions of the extended Kalman filter in points 1 and 5–7 (Figure 1) can be used in analytical practice to steer an autosampler or a laboratory robotic that introduces automatically samples or standards into the spectrometer or the chromatograph. To demonstrate the usefulness of the new algorithm for analytical practice, an experiment with enhanced drift has been set up. As an example, the determination of Cd traces in aqueous HNO_3 using an independent running GF-AAS/autosampler setup has been chosen. The aim of the experiment was to generate a strong drifting Cd absorbance versus time and to test the extended Kalman filter with these data.

To achieve artificial drift, the standard procedure of Cd determination in 1% HNO_3 has been modified. The HNO_3 concentration has been increased to 10% in order to provide an artificial enhancement of chemical drift by increased chemical corrosion of the inner surface of the graphite tube. In this way, the extended Kalman filter has been tested in a worst case situation. A very often repeated analysis of exactly the same Cd solutions allowed us to follow drift versus time in a quantitative way. The drifting Cd absorbances were then used to test the prediction and calibration behavior of the extended Kalman filter algorithm off-line in several computer studies.

(29) Vijn, T. Outlier Detection and Repair with an Extended Kalman Filter. Student Research Report, University of Nijmegen, Jan 1992.

Chemicals. Aqueous solution of $Cd(NO_3)_2$ (Merck) in 10% HNO_3 (Merck) was used. The calibration line consisted of a blank and four standards (0, 0.5, 1.0, 1.5, and 2.5 ppb Ca in 10% HNO_3) providing $n_s = 1 + 4 = 5$. For quality control a QC standard of 2.0 ppb Ca in 10% HNO_3 was prepared. A sample with known concentration (1.833 ppb in 10% HNO_3) was used to follow independently the accuracy and precision.

Instrumentation. The analytical data for a pseudo-on-line evaluation of the extended Kalman filter strategy were produced by using a computer-controlled graphite furnace atomic absorption spectrometer (Philips Pye Unicam PU-9200X) linked with a furnace autosampler (type 9380X). An uncoated normal-type graphite tube (No. 9423,393,90031) was applied in this corrosion experiment. The Cd hollow cathode lamp operated at 7 mA.

Procedure. The following standard temperature program was chosen: evaporation 110 °C/20 s, ash 300 °C/30 s, atomize 1200 °C/3 s, clean 2000 °C/3 s, cool 25 °C/10 s. The injection volume was 15 μL, and a deuterium baseline correction was used. The absorbance values were measured at 228.8 nm as the peak height for Cd. The peak area would be an alternative. However, the Cd peak shape did not change significantly in this drift experiment so that the height formed a representative choice. For each single sample the blank, the full set of the four calibration standards, and a QC standard were measured in a cyclic way. This was defined as one measurement cycle. Such a complete measurement cycle provides the best experimentally reachable drift compensation for the corresponding sample. In total, 47 samples, 235 calibration standards, and 47 QC standards were analyzed. Each sample and each standard were analyzed as three replicates $n_r = 3$ giving in total $3(47 + 47 + 235) = 987$ sequential measurements. According to the temperature program, the analysis of one sample or one standard took 66 s. The preparation and injection of it from a master solution took ~100 s. This required a total of 2.7 min per measurement over a total period of 45 h in which the spectrometer and autosampler worked continuously, fully automatically, and independently.

Computations and Software. The raw data were on-line saved to floppy disk by the computer of the running spectrometer/autosampler setup. Simultaneously the data were on-line transferred to a dedicated printer. Finally, the obtained experimental data set consisted of ~1000 Cd absorbances with their corresponding Cd concentration (alternatively 0.0, 0.5, 1.5, 2.5, 2.0, or 1.833 ppb Cd) and their corresponding time points of measurement (frequency 1/2.7 min). These data set formed the input for the pseudo on-line evaluation of the extended Kalman filter strategy at an IBM-PC. Pseudo-on-line means in this case that for the Kalman filter algorithm, running in stand-alone mode on a PC, the experimental measurements were received from floppy disk in the same sequence in place of a direct sequential input from the GF-AAS. It means further that the control commands of the algorithm to the autosampler were translated to commands for reading the corresponding absorbances for recalibration or QC standards of samples from floppy disk. A randomization of the order of recalibration measurements avoids additional drift that could be introduced by a fixed

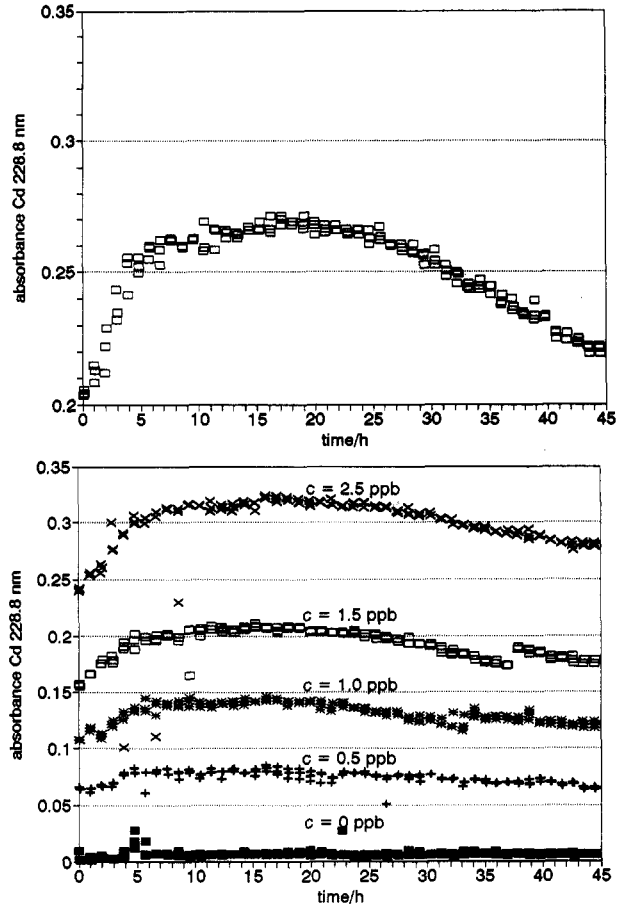

Figure 3. Drifting experimental absorbances of all analyzed (a, top) Cd quality control standards (2 ppb Cd) and (b, bottom) blanks and calibration standards versus time.

sample order. Unfortunately, the current software for the autosampler control did not allow such a randomized measurement. However, randomization has been realized afterward in the pseudo-on-line calculations by reading and processing the recalibration standards in random sequence from the file. The chosen pseudo-on-line validation of the new strategy does not influence future results. But it had some practical advantages for the research such as easier and faster data handling and independence from the slowly working spectrometer/autosampler setup. As predefined precision S_{pred} of the Cd analysis, a value of S_{pred} = 7% was chosen in all calculations, based on literature studies about Cd standard analytical methods. The vigilance threshold for outlier rejection chosen was ~2 times poorer with S_{pred}^{LMS} = 15%. The extended Kalman filter approach as given in Figures 1 and 2 was first tested with the experimental data using an in MATLAB programming language[30] written version. Later on, the program was transferred to a Turbo-C (Borland-International) version. An executable demonstration version for IBM-PC (MS-DOS operating system) with VGA graphics is available from the authors.

RESULTS AND DISCUSSION

Measured Data. Drift with a positive slope happens in a first stage of the graphite tube's life cycle. This can be seen from the drifting raw Cd absorbances (Figure 3a,b). During this first 5 h, the system is not stable. The reason for this positive drift is the thermal and chemical conditioning of the fresh graphite furnace.

The first stage is followed by a stable time interval (6–30 h, Figure 3a,b) where the absorbance of the QC standards does not vary much. Except in some few cases, outliers and discontinuities are observed. Examples are the absorbances for the standards of 2.5 and 1.5 ppb Cd (Figure 3b, time 5–10 h). The strong outliers were caused by the robotic arm of the autosampler, which had slightly touched the edge of the tube inlet. However, these outliers were successful detected by the implemented modified LMS procedure for outlier detection using the decision threshold of S_{pred}^{LMS} = 15%. The curve for the standard of 1.5 ppb (Figure 3b, time 37–43 h) contains an example of another type of discontinuity. For this steplike type of discontinuity, no obvious explanation was found up to now. The event occurred in the night, when nobody was present to observe the independently running spectrometer/autosampler setup. However, the parallel declination of the curve suggests also a deviating event comparable with that of the accident with the robotic arm and followed by a memory effect.

The last stage of the graphite furnace's life cycle (31–45 h, Figure 3a,b) is the aging phase with drift having a negative slope. This type of negative drift is caused by the rapid chemical erosion of the tube's inner surface. The end of the life cycle of the used graphite furnace can be seen from the deep trenches in the tube wall at the sample position (Figure 4a,b).

Selected calibration lines (Figure 5) taken within the time interval of 45 h were calculated from the raw data in Figure 3b. The slope of the lines reaches a stable maximum with a width of nearly 25 h and declines again in the tube's aging phase. This illustrates mathematically the chemically caused drift of the GF-AAS system as already seen in Figure 3a,b. In the caption of Figure 5 is given quantitative values, and the t-test results in Table 2 show that the drift is statistically significant. The largest differences were found for the slopes for the initial (0 h) and the final phase (40 h) of the tube's life compared to the middle phase (8–32 h). It is interesting that the intercept of the lines does not drift. The explanation for this can be found in Figure 3b. The height of the Cd absorbance peaks is also a function of the generated drift.

Quality Self-Monitoring with the Extended Kalman Filter Algorithm. The extended Kalman filter algorithm according Figures 1 and 2 has been sequentially and pseudo-on-line applied to the drifting 1000 sequential measurements as given in Figure 3a,b. The aim of the algorithm has been to predict whether a complete recalibration (rc) or a quality control sampling should be performed. In this way, the actual calibration line was adapted to any situation of drift with a minimum number of recalibrations and QC samples. Figure 6 and Figure 7 summarize all the obtained results by graphical visualization of the behavior of the extended Kalman filter algorithm during the 45-h corrosion experiment. In Figure 6, it can be seen that the algorithm recalibrated 4 times within the 45-h run after a first initial calibration. These four recalibration events happened with a flexible distance and at especially critical positions of the drift curve in Figure 6. That means that recalibrations were chosen by the algorithm automatically more frequently when the Cd absorbance showed

(30) PC-Matlab™, User's Guide, The Math-Works Inc., Natick MA, 1990.

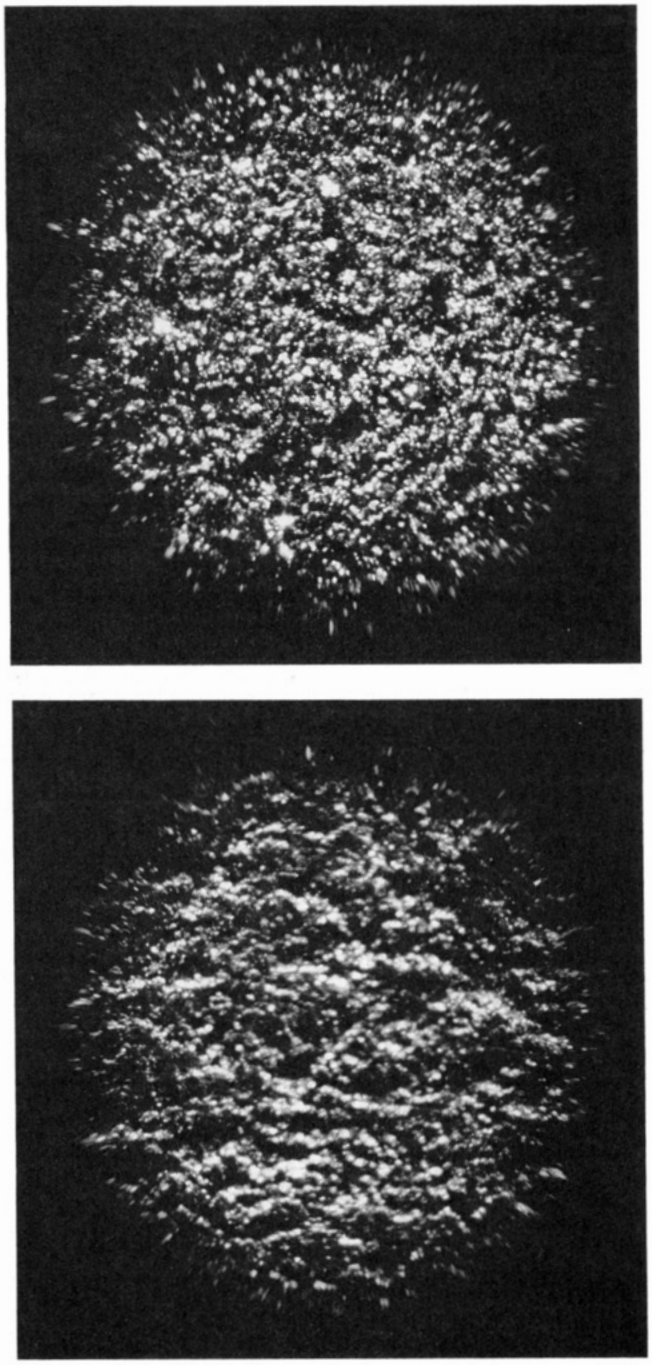

Figure 4. (a, top) Light microscopic photograph (enlargement 25×) of the sample site in the fresh graphite furnace tube before its first use in the corrosion experiment for the generation of artificially enhanced drift. (b, bottom) The same sample site (a) after the drift generation experiment (Cd in 10% HNO_3, ~1000 excitations within 45 h; also see text).

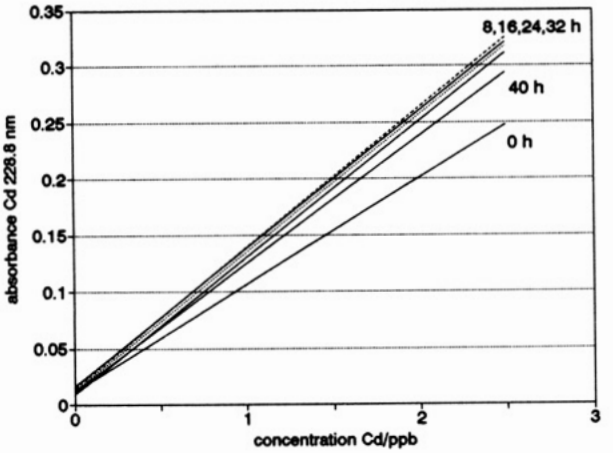

Figure 5. Drifting calibration lines each 8 h apart calculated from data in Figure 3b. Corresponding slopes a and their standard deviations, $s(a)$: $a_{00h} = 0.094\ 32$, $s(a_{00h}) = 0.004\ 13$; $a_{08h} = 0.119\ 37$, $s(a_{08h}) = 0.003\ 17$; $a_{16h} = 0.125\ 70$, $s(a_{16h}) = 0.003\ 36$; $a_{24h} = 0.121\ 78$, $s(a_{24h}) = 0.003\ 21$; $a_{32h} = 0.094\ 32$, $s(a_{32h}) = 0.002\ 50$; $a_{40h} = 0.109\ 89$, $s(a_{40h}) = 0.002\ 50$. Test results for significant differences given in Table 2.

Table 2. Calculated Pairwise *t*-Significance Tests for Differences between the Slopes a_i and a_j of Two Drifting Calibration Lines *i* and *j* Shown in Figure 5[a]

slope	a_{08h}	a_{16h}	a_{24h}	a_{32h}	a_{40h}
a_{00h}	4.32	4.78	4.11	3.86	3.27
a_{08h}		1.79	1.68	2.29	2.97
a_{16h}			2.32	2.47	3.14
a_{24h}				2.38	2.74
a_{32h}					2.68

[a] The calculated t values were obtained by $t_{calc} = \text{abs}(a_i - a_j)/s_d$ whereby s_d, the variance, is for the difference $\text{abs}(a_i - a_j)$. For the complete formula of s_d and more details of this t-test consult p 143 of ref 4. If t_{calc} exceeds the tabled t-value $t_{tabled,f,1-4} = 2.06$ with $f = (n_{r,i}n_{s,i}) + (n_{r,j}n_{s,j}) - 4 = (3)(5) + (3)(5) - 4 = 26$ degrees of freedom and statistical risk of $r = 0.05$, then both calibration lines i and j are significantly different in their slope.

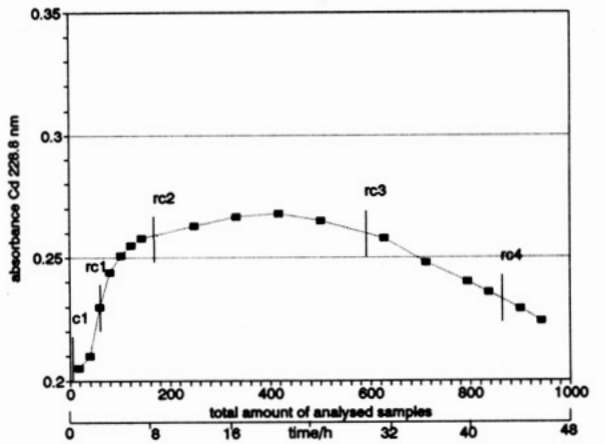

Figure 6. Minimized number of single QC standards and sets of calibration/recalibration standards measured sequentially and non-equidistant within 45 h according to the commands given by the extended Kalman filter algorithm in an optimal way over the total amount of ~1000 analyzed samples (c1, calibration; rc, recalibration).

stronger drift. Additionally it can be seen in Figure 6 that the algorithm sometimes decided to analyze only QC standards in place of a complete recalibration. The extended Kalman filter is able to handle the distance between two QC standards very flexibly according to the absence or the presence of strong drift. The correctness of the decisions, i.e., finally correct behavior of the extended Kalman filter in quality self-monitoring, can be seen in Figure 7. Good correction of the drifting absorbance and a high agreement with the experimentally best achievable curve were obtained.

For detailed understanding of how the algorithm made its different decisions we look more quantitatively to the different phases of the drift curve in Figure 6: After initialization with $R = 0.02$ and $Q = 10^{-6}$, the first calibration c1 was measured. R and Q as intermediate computational variables of the, by its nature, recursive algorithm were chosen on the basis of previous experimental and computational experiences. A drifting first-order calibration graph was used for state

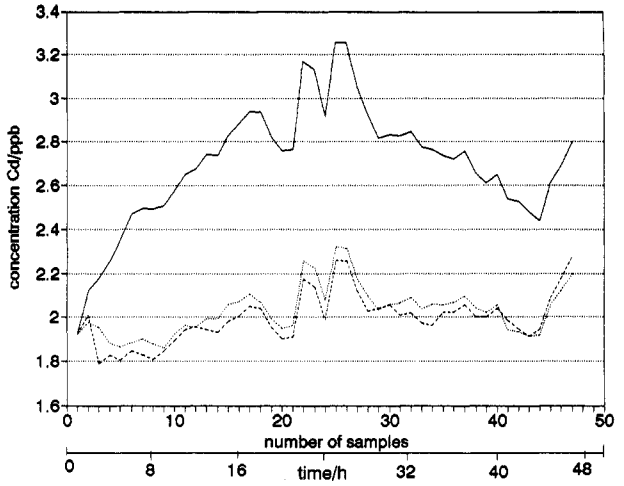

Figure 7. (—) Drift of the concentration of the "unknown" sample of constant 1.833 ppm Cd calculated from the absorbances based on the linear regression model obtained from the initial calibration c1 (not drift corrected). (- - -) Drift of the concentration of the same Cd sample corrected by the extended Kalman filter based on the few QC standards and the calibrations/recalibrations as shown in Figure 6. (– –) Best experimental drift corrected concentration of the same Cd sample reached by using the total set of 47 × 3 QC standards and 235 × 3 calibrations/recalibrations.

Table 3. Kalman Filter Calculated Actual Precision S_{calc} and State Vectors x(k) Containing the Actual Slope a and Intercept b of the Calibration Line as Well as the Corresponding Drift Parameters α and β[a]

		state vector $\mathbf{x}_{(k)}$			
action	S_{calc}/ %	a	b	α	β
rc1	8.306 05	0.111 60	0.010 90	0.000 22	−0.000 09
qc3	5.895 65	0.119 33	0.014 69	0.000 66	0.000 19
qc4	6.061 37	0.121 28	0.015 24	0.000 66	0.000 19
qc5	6.100 59	0.122 30	0.015 06	0.000 55	0.000 13
qc6	5.931 92	0.123 78	0.015 33	0.000 53	0.000 12
rc2	5.960 32	0.121 12	0.012 50	−0.000 22	−0.000 09
qc7	3.720 91	0.125 75	0.017 39	0.000 33	0.000 25
qc8	3.759 31	0.123 51	0.016 43	−0.000 09	0.000 03
qc9	3.720 87	0.125 60	0.018 36	0.000 12	0.000 13
qc10	3.641 46	0.122 74	0.017 93	−0.000 16	−0.000 01
qc11	3.360 36	0.120 85	0.017 88	−0.000 16	−0.000 01
rc3	3.176 62	0.118 66	0.006 15	−0.000 21	−0.000 17
qc12	2.590 89	0.116 67	0.004 40	−0.000 06	−0.000 08
qc13	3.533 85	0.109 86	0.000 49	−0.000 46	−0.000 28
qc14	5.259 38	0.109 84	0.000 24	−0.000 35	−0.000 23
rc4	6.019 66	0.108 78	−0.000 44	−0.000 35	−0.000 23

[a] The 15 data sets out of ~1000 sequential measured data sets were selected at the points of time when the Kalman filter itself initiated an autosampler action rc (measure recalibration standard) or an action qc (measure quality control standard). (Compare with Figure 6.)

estimation. Then the precision, S_{calc}, was determined (points 2 and 3 in Figure 1; point c1 in Figure 6). In this stage of the process, S_{calc} was significant poorer compared to S_{pred}. The algorithm decided on a QC distance of one sample using the user-defined limit as given in Table 1. This decision means that, within each measurement cycle, one QC standard (three replicates) had to be measured to update the Kalman filter. After three QC standards, the precision of the calibration line, S_{calc}, is still poor if compared with S_{pred}. Thus, the system needs to be recalibrated again (rc1 in Figure 6). The entire set of previous measured samples had to be remeasured because there was more drift than predicted by the Kalman filter. Fortunately, in this case, only one sample had to be reevaluated because the Kalman filter was not in stable drift and did predict only a QC distance of one. The internal behavior of the extended Kalman filter can be followed by looking to the intermediate calculation results (Table 3). It is obvious that after the first recalibration (rc1 in Figure 6; see also Table 3) the precision of the calibration line S_{calc} became better than the predefined threshold of S_{pred} = 7%. There is no more need for a recalibration. The system started to analyze samples. However, the QC distance has to still be one sample. According to the limits for the QC distance is based on the actual value for S_{calc} (Table 3 and Table 1) because the drift is strong. The Kalman filter monitored its own parameters by measuring one QC standard after each analyzed sample. In the first 5 h, the algorithm gives no confidence in the excitation stability of the fresh graphite tube. After measuring five QC standards and analyzing five samples, the system recalibrates again (rc2 in Figure 6). A limit of five samples maximum between two QC standards has been implemented as the upper limit to avoid an overweight of the increasing number of QC standards in the calibration line. The period of time before this maximum is dependent on the calculated precision S_{calc} and on user-defined limits. In a system with fast-changing drift characteristics, this time will be relatively small. This is in fact exactly what is required because in a system with poor predictable drift more recalibrations are needed.

If the drift stays linear, it can be seen how the Kalman filter increases the QC distance or how it tries to keep the QC distance at least constant. An example of this is the good S_{calc} values (Table 3) for all QC samplings between the actions rc2 and rc3 that stimulate the Kalman filter to extend carefully the distance between two QC standards in time (Figure 6). The Kalman filter relies more on the estimated parameters a, b, α, and β, based on up to the moment learned data. Thus, the Kalman filter relies on a higher excitation stability of the graphite tube. That means that relatively more samples but less QC standards were analyzed within the 6–30-h time interval (Figure 6) compared to the other time ranges. Recalibration rc3 (Figure 6) happened after the block of maximally allowed five QC standards. However, the precision S_{calc} of the calibration remained good enough (Table 3). Three QC standards later, it can be seen how the precision dramatically decreases. The graphite tube, which has been artificially etched by nitric acid, reaches the final phase of corrosion. After only three QC standards, the Kalman filter predicted this negative development and decided for a recalibration rc4 (Figure 6) to reach a better estimate for a, b, α, and β. After 34 h, the QC distance becomes smaller again. The deterioration of the tube causes fast-changing nonlinear negative drift. At this point more QC standards are needed. The graphite tube should be replaced, in general, when the absorbance starts to decrease dramatically, giving a poorer limit of detection. The size of the slope a of the regression line, provided by the Kalman filter, can be used as a test criterion as to whether the tube has lost its sensitivity or not. However, the extended Kalman filter provides additional information about the time point when the graphite

tube should be replaced. If the drift changes rapidly, then the QC distance decreases rapidly to a value of one. Thus, the Kalman filter tends then to call a QC standard measurement with the same frequency as an analysis. The tube should then be replaced because the drift became unpredictable and uncorrectable.

In total, the extended Kalman filter algorithm asked for 1 calibration, 17 QC standards, and 4 recalibrations. In total, ~1000 measurements in 45 h were performed. The algorithm continuously adapted the calibration parameters a and b and the drift parameters α and β. In this way, it was possible to evaluate each sample by an optimal adapted calibration line. However, from Figure 7 it can be seen that a limited amount of drift is still observed in the theoretical line and in the slightly different prediction line of the Kalman filter. It is assumed that nonlinear drift cannot be compensated for by a linear drift model. Abrupt changes also count like nonpredictable nonlinear events. It is assumed that small outliers that were not detected by S_{pred}^{LMS} (time 5–10 h) result 10 h (time 20–30 h) in the oscillating poor predictions. An indication for a similar delay in the prediction can be found later once more. The discontinuity (time 36–40 h) later (8 h) causes a sharp increasing predicted absorbance (time 44–48 h). An alternative possibility to overcome these problems could be a nonlinear model for drift.

CONCLUSIONS

The extension of the classical Kalman filter quality control algorithm by the possibility of adapting its parameters either by quality control standards or by recalibration standards provided a chemometrical technique for quality self-monitoring of independently working automatic analyzers and sensors.

The algorithm performs a continuous dynamic modeling of the parameters a and b of a calibration line as well as the drift α and β in both parameters. α and β allow one to take into account the future development of the drift. In this way, a prediction of the individual adapted calibration line becomes possible for each sample. The calculated precision, S_{calc}, for each individual calibration line is compared with S_{pred}, that is, by a user-predefined precision. Based on this test, the extended Kalman filter is able to forecast how many unknown samples can be measured, beginning with the time point k up to the next recalibration. The algorithm forecasts in this way the time point for the next necessary QC standard measurement and/or the next recalibration.

Additionally an automatical outlier treatment procedure based on a modified least median of squares method with new stop criterion, S_{pred}^{LMS}, was incorporated to overcome the influence of strong outliers and to repair them on-line.

The application of the extended Kalman filter in automated graphite furnace AAS makes it possible to use a graphite tube over a longer period of time. The first reason for that is that the tube can also be used in its drifting phases in the beginning and in the end of its life. The drift is compensated for. The second reason is that fewer recalibrations and fewer QC standards are necessary, providing space in time for the analysis of samples. More space for samples became available because QC standards and recalibrations were not equidistant, but optimal placed in the space of time adapted to the real drift situation. The self-monitoring of the system by QC standards happened more frequently in a drifting phase than in a phase with no drift.

Limitations of the extended Kalman filter algorithm are the mathematical complexity and some heuristics. The heuristics are the predefined precisions, S_{pred} and S_{pred}^{LMS}, the maximal allowed QC distances, and the choice of suitable initial parameters to force a fast and optimal adaptation of the algorithm in the starting phase of the long-term measurement. The choice of the predefined precision depends on the analytical technique. Especially for well-defined standard methods, this precision can simply be found in the related published method descriptions. The suitable choice of the QC distance (maximal number of samples allowed to be analyzed between two QC standards) and the maximal allowed number of QC standards between two recalibrations require more research. It depends strongly on the monitored sensor system. In our experimental case, it depended on the chosen excitation conditions, type of graphite tube, and strength of the acid. One way to overcome the initialization problem and to minimize the initial training phase of the Kalman filter is the use of starting parameters of previously used graphite tubes. Especially in routine analyses, the type of tubes and the analytical conditions stay more or less the same. Routine laboratories especially accumulate this kind of experience. Another disadvantage of the algorithm is that only a linear drift model is described, so only linear drift can be predicted. However, the choice of the model depends on the application and can be implemented easily as theoretically shown for a second-order calibration curve with linear drift (eq 5).

The general gain of an intelligent quality self-monitoring will be in this way a more economical use of the expensive graphite tubes or, in general, of the drifting sensor unit.

ACKNOWLEDGMENT

The authors are grateful to Gerrit Kateman (University of Nijmegen, NL), Hans v. Leeuwen (AKZO, Arnhem, NL), Anthony Mulick, Dave Norman, Steve Morton (Pye Unicam Analytical Instruments Cambridge, UK), Piet Rommers, and V. Vullings (Philips Research Laboratories, Eindhoven, NL) for discussions in the initial phase of the project. The authors thank Jelle Eygenstein and Rien v. d. Gaag (University of Nijmegen, Central Instrumental Facility of the Science Faculty, Nijmegen, NL) for their initial help in the drift experiments and students Marc Nulens and Margo Pulles for their partial contributions in the further surroundings of the project. Additionally, the authors thank both reviewers for their valuable hints and remarks.

Received for review August 11, 1993. Accepted December 15, 1993.®

® Abstract published in *Advance ACS Abstracts*, February 1, 1994.