Comparison of methods for accurate end-point detection of potentiometric titrations

R L A Villela[1,4], P P Borges[2] and L Vyskočil[3]

[1] Acoustics and Vibration Metrology Division, National Institute of Metrology, Quality and Technology - Inmetro. Av. Nossa Senhora das Graças, 50, Duque de Caxias, RJ, CEP 25250-020, Brazil.
[2] Chemistry Metrology Division, National Institute of Metrology, Quality and Technology - Inmetro.
[3] Slovak Institute of Metrology-SMU. Karloveská 63, SK-84255 Bratislava 4 Slovak Republic.

E-mail: rlvillela@inmetro.gov.br

Abstract. Detection of the end point in potentiometric titrations has wide application on experiments that demand very low measurement uncertainties mainly for certifying reference materials. Simulations of experimental coulometric titration data and consequential error analysis of the end-point values were conducted using a programming code. These simulations revealed that the Levenberg-Marquardt method is in general more accurate than the traditional second derivative technique used currently as end-point detection for potentiometric titrations. Performance of the methods will be compared and presented in this paper.

1. Introduction

Titration is one of the most utilized techniques for measuring the amount of an analyte by considering the stoichiometry relationship in a chemical reaction.

Potentiometric titration curve are represented by a plot of the indicator-electrode potential (y) versus the volume (or mass) of the titrant. For obtaining the end point of a titration, the first and second derivative plots of the potential as a function of the volume (or mass) of the titrant is normally used [1]. This method has no good accuracy due to its high susceptibility to measurement uncertainty of y and insensitivity to the number of data points around the end point.

In this paper is proposed a non-linear regression of measurement data from potentiometric titrations by the Levenberg-Marquardt algorithm [2] for end point determination. This non-linear regression can be used to determine the end point with better accuracy. For example, in potentiometry a 0.25 mV error represents a 1% relative error in the concentration of the analyte [1]. This can be a huge problem for characterization of reference materials, since it has influence in the accuracy of the property value measured due to experimental errors.

A code was developed in LabVIEW® (a proprietary development environment for programming) to determine the end point from acid-base coulometric titration with potentiometric detection. This paper aims to present the performance and comparison of the two methods, namely the traditional second derivative (SD) technique and the Levenberg-Marquardt (LM) approach.

[4] To whom any correspondence should be addressed.

2. Implementation of the methods

2.1. Potentiometric titration data

The coulometric titration data were obtained by using the primary system of coulometry from Inmetro [3]. The experimental data are related both to the initial and to the final acid-base coulometric titration. The end points were calculated through the software developed using both the SD method and the LM algorithm.

2.2. Second derivative method

This method consists of the estimation of the end point by the second derivative that is obtained numerically. It was implemented in the software that reads the n experimental data points, time (t_i), in seconds, and pH (y_i), where i = 1, 2, ..., n. The n–2 points of the second derivative (y_j'') are used to estimate the value of the end point (t_e), that is computed considering that $y''(t_e) = 0$ and using a linear interpolation of t_j and t_{j+1}, where $t_j \le t_e \le t_{j+1}$ and $y''(t_{j+1}) \le y''(t_e) \le y''(t_j)$. The respective value of pH, $y(t_e)$, was calculated from the linear interpolation of $y(t_i) \le y(t_e) \le y(t_{i+1})$ and $t_i \le t_e \le t_{i+1}$ (figure 1).

Unfortunately, in many cases this technique cannot provide a measurement result with low uncertainty because numerical derivatives are very prone to measurement noise (random error). Furthermore, this method requires two successive differentiations and uses only two points to estimate t_e from the second derivative data. Therefore a better approach to reduce the uncertainty of end-point estimations is reasonably expected.

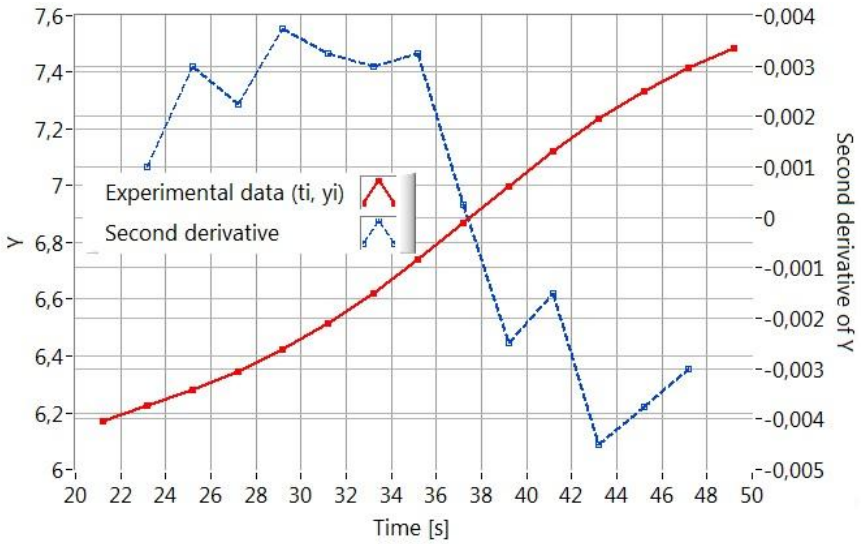

Figure 1. Experimental data and second derivative both linearly interpolated.

2.3. The Levenberg-Marquardt Method

An empirical equation (1) presented at the first time by Vyskočyl and Mathiasová [4] was used in the LabVIEW® code to model the electrochemical reaction occurring in coulometric titrations when the detection of the end point is obtained by potentiometry.

$$y = \frac{a}{1 + e^{k(t-q)}} + b(t - q) + c \qquad (1)$$

Where: y is the electrode potential; t, time for the titration reaction (may be volume or mass); c, position of the titration curve; b, slope of the titration curve; q, position of the end point; k, width of the titration curve and a amplitude of the titration curve.

An initial guess of the q value can be obtained from the maximum value of the first numerical derivative. It can be easily demonstrated, according to equation (2), that $t_e = q$ is the end-point value.

$$\frac{\partial^2 y(t_e)}{\partial t^2} = \frac{ak^2 e^{k(t_e+q)}\left(e^{kt_e} - e^{kq}\right)}{(e^{kt_e} + e^{kq})^3} = 0 \quad (2)$$

The program implemented reads the n experimental data points, time (t_i) and pH (y_i), where i = 1, 2, ..., n, computes numerically the first derivative to specify the initial guess for the parameter q (maximum value) and uses the LM algorithm to determine the set of parameters that best fit the set of experimental data points. The best fit parameters minimize the weighted mean squared error between the observations y_i and the best nonlinear fit. From the parameters a, c and q, the end point (t_e) and the respective value of pH (y_e) are obtained from equations (3) and (4).

$$t_e = q \quad (3)$$

$$y_e = a/2 + c \quad (4)$$

The stopping conditions for the fitting process were set to 5×10^5 for the maximum number of iterations of the fitting routine and the value of 10^{-10} for the tolerance that specifies the relative change in the weighted distance between y_i and the fit curve. The best fit curve and the experimental data are shown in figure 2.

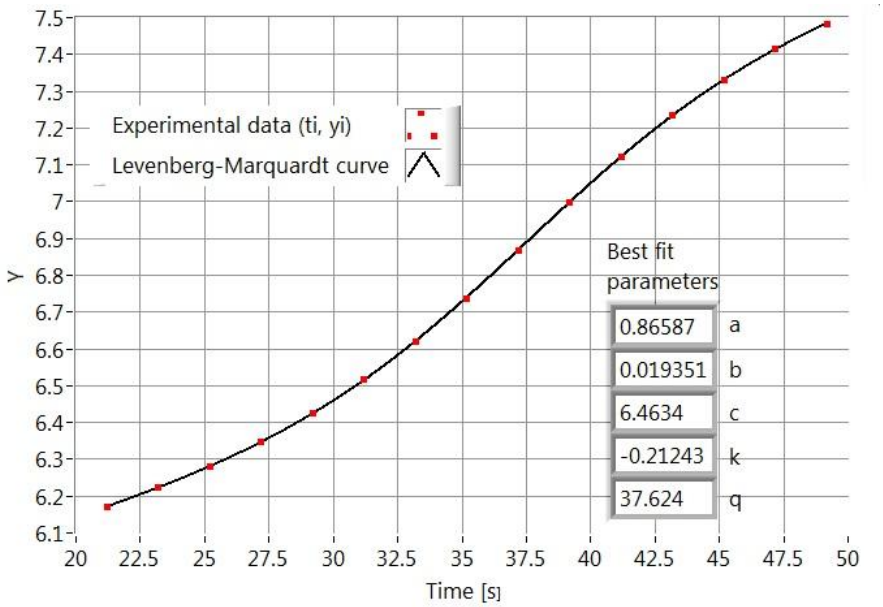

Figure 2. Experimental data and best fit curve using the Levenberg-Marquardt algorithm.

3. Comparison of the methods

3.1. General procedure

The comparison between the SD technique with the LM method was carried out taking into account their performance. This comparison was undertaken through the evaluation of result deviations obtained from simulated data points t_i and y_i.

It was considered that the equation (1) describes the electrochemical reaction and the measurement data y_i can be mathematically modeled by equation (5).

$$y_i = \frac{a}{1 + e^{k(t_i - q)}} + b(t_i - q) + c + e_i \quad (5)$$

Where e_i is the measurement error of y_i and a normally distributed random variable with zero mean and standard deviation σ_e, where $i = 1, 2, ..., n$.

When the σ_e and all parameters of the equation (5) are specified, the value of each y_i can be computed (simulated data) and the true values of t_e and y_e can be calculated from equations (3) and (4) for a given set of data points t_i and y_i generated.

Accordingly, the generated values of t_i and y_i can be used by the SD and LM methods for estimation of t_e and y_e and finally to compare with the true values of t_e and y_e. The assessment of method accuracy for a specific data set t_i , y_i where $i = 1, 2, ..., n$, can be carried out computing the method errors $e(t_{eSD})$, $e(y_{eSD})$, $e(t_{eLM})$ and $e(y_{eLM})$ presented at the flow diagram in figure 3.

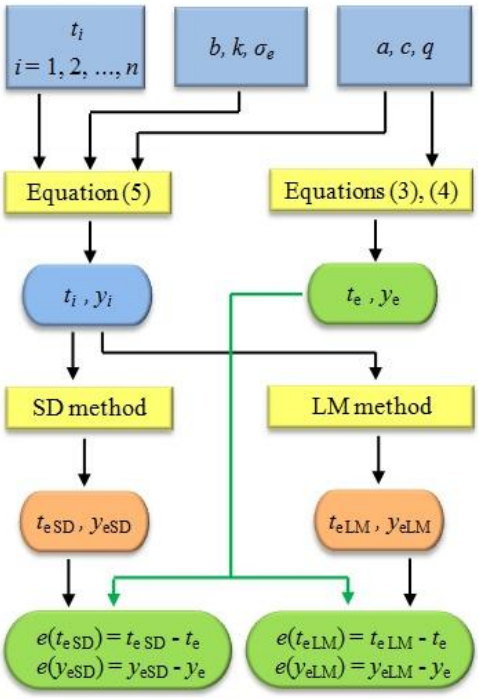

Figure 3. Flow diagram for calculation of the resultant errors of t_e and y_e due to the errors e_i.

In summary, a simulation of experimental data consisting of n points t_i , y_i , where $i = 1, 2, ..., n$, provides the errors of t_e and y_e for both methods: $e(t_{eSD})$, $e(y_{eSD})$, $e(t_{eLM})$ and $e(y_{eLM})$. Consequently, if a large number of simulations are performed the probability distribution, mean and standard deviation of the errors can be computed for an error analysis.

3.2. Parameter specifications

If a simulation for an error analysis is conducted using only a specific set of values of *a, b, c, k, q,* the conclusion of the analysis (the best method for end-point detection) will be very particular to the case tested. In order to achieve some generality, 1×10^5 simulations were performed, where each simulation s has its own parameters a_s, b_s, c_s, k_s, q_s, t_{is} (n is constant) and resultant errors $e_s(t_{eSD})$, $e_s(y_{eSD})$, $e_s(t_{eLM})$,

$e_s(y_{eLM})$. Moreover, to obtain representative values of a_s, b_s, c_s, k_s, q_s, t_{is} and σ_e for all simulations (s = 1, 2, ..., 1×10^5) many data sets of titration experiments were used to determined their minimum and maximum values. This was computed through the LM method and was crucial to avoid unrealistic simulations. In each simulation, the parameters a_s, b_s, c_s, k_s, q_s were generated as random variables according to a uniform distribution with their respective minimum and maximum limits (table 1).

Table 1. Parameter limits used for error analyses.

Parameter	Min value	Max value
a	0.76	0.99
b	0.015	0.028
c	6.4	6.5
k	-0.3	-0.21
q	27	49

3.3. Error analyses

Each error analysis consisted in 1×10^5 simulations of n experimental data points to compute the errors $e_s(t_{eSD})$, $e_s(y_{eSD})$, $e_s(t_{eLM})$ and $e_s(y_{eLM})$ 1×10^5 times and the corresponding means and standard deviations. It was performed four error analyses to reveal the behavior of the methods in different conditions of σ_e and n (table 2).

After some examination of the analysis results the following conclusions were reached for the cases tested:

- For both methods, an increase in the value of σ_e causes an increase in the standard deviations of the errors of t_e and y_e.
- The means of the errors of t_e and y_e are negligible compared to their respective standard deviations. Thus, an unbiased estimation of the end-point value can be obtained by application of any of the two methods.
- The performance of the LM method was markedly superior than the SD method for most of the simulations performed. For all error analyses of t_e using fourteen data points, the absolute value of LM error was smaller than the respective absolute value of SD error in approximately 95% of the simulations.
- The LM method exhibited improved performance with an increase of the number of data points (n) while the SD method did not show this behavior.

Table 2. Error analyses of t_e for different conditions.

Error analysis		1	2	3	4
Standard deviation of e_i (σ_e)		0.0008	0.0016	0.0024	0.0008
Number of data points (n)		14	14	14	10
LM method	Mean of $e(t_e)$	6.48×10^{-6}	8.60×10^{-5}	9.30×10^{-5}	1.09×10^{-4}
	Standard deviation of $e(t_e)$	0.0145	0.02896	0.0429	0.02523
SD method	Mean of $e(t_e)$	2.30×10^{-4}	1.87×10^{-4}	-2.14×10^{-4}	-6.66×10^{-5}
	Standard deviation of $e(t_e)$	0.2163	0.3909	0.5088	0.2161
Percent of simulations $\lvert e_s(t_{eLM})\rvert \le \lvert e_s(t_{eSD})\rvert$		95.0	94.8	94.5	91.9

4. Conclusions

The error analyses revealed that in most cases the Levenberg-Marquardt method compared to the second derivative method reduces drastically the error of the end-point detection from potentiometric titrations. In consequence, the application of the Levenberg-Marquardt method can increase the accuracy of the end-point value from potentiometric titration data and contribute to a better characterization of reference materials by using coulometric titration with potentiometric end-point detection.

References

[1] Sawyer D T, Sobkowiak A and Roberts Jr J L 1995 *Electrochemistry for Chemists* 2nd edition (New York:John Wiley & Sons).
[2] Marquardt D W 1963 *J. Soc. Indl. Appl. Mathematics* **11** 431.
[3] Borges P P, Sobral S P, Silva L, Araujo T O and Silva R S 2011 *J. Braz. Chem. Soc.* **10** 1931.
[4] Vyskočyl L and Mathiasová A 1986 *8. Medzinárodné Sympózium Metrológie*. INSYMET'86.