

Rounding error, an unexpected fault in the output from a recording spectrophotometer: implications for model discrimination

María Luz CÁRDENAS and Athel CORNISH-BOWDEN*

Laboratoire de Chimie Bactérienne, Centre National de la Recherche Scientifique, B.P. 71, 31 chemin Joseph-Aiguier, 13402 Marseille Cédex 20, France

Although commonly ignored in discussions of experimental error, rounding may sometimes be the major source of error, especially with modern precision instruments: some recording spectrophotometers are optically and photometrically capable of making absorbance measurements with errors less than 0.0003, but provide no numerical information more precise than ± 0.001 .

The problem may be diagnosed by a characteristic arrangement of points in a residual plot, which resembles the result of cutting a stroboscopic picture of a bouncing ball into several strips and modifying it by sliding the strips relative to one another to bring the points closer to the axis. Harmful effects of rounding error can be critical in experiments designed for model discrimination.

INTRODUCTION

A major objective of experimental design is to decrease error to the point where the predictions of different models can be clearly distinguished. In this context, discussions of sources of error in statistical analysis (e.g. Ellis and Duggleby, 1978; Draper and Smith, 1981, pp. 141–192; Nimmo and Atkins, 1981; Henderson, 1992) usually restrict attention to systematic error, which is the error caused by fitting data to an incorrect model, and random error, which is what is left when all errors with assignable causes are allowed for. Rounding error, which arises from introducing arithmetic approximations too early in the calculations, is a third fundamental kind of error that is often ignored because it can always be made trivial if one has full control over the calculations. For an example of its potentially serious effects, see Draper and Smith (1981, pp. 257–258).

When spectrophotometers could not measure absorbances of about 1 more precisely than ± 0.01 , relatively simple arithmetic was sufficient to extract all of the available numerical information from the data. However, improvements in the optical quality of recording spectrophotometers have not always been matched by improvements in the software built into such instruments. In the past this would have been no more than a minor irritation, because one could always analyse the data by hand. With increasing automation one no longer necessarily has direct access to the primary data, however, and the output values may be rounded to the point where they are an order of magnitude less precise than the measurements.

We show that rounding error may be the major source of experimental error; we discuss how to recognize this condition and how to minimize its effects if the obvious solution of correcting the underlying design fault is not available. For simplicity we assume that only the absorbances are subject to error, i.e. that times are known exactly; although departure from this assumption can greatly complicate the treatment of random error, we believe that it is insignificant in relation to systematic and rounding errors.

RESIDUAL PLOTS

In a residual plot, an effective tool for studying error types (Draper and Smith, 1981, pp. 141–192), the weighted difference between observed and calculated values of a measured quantity

is plotted against an independent variable or the calculated value of the measured quantity. (We shall not discuss weighting here, as it is not directly related to the recognition of rounding error, apart from noting that the proper weight is the square root of the weight appropriate for least-squares fitting.) When the measurements constitute a time series it is convenient to use time as the abscissa variable, and this is done in all the examples in the present paper.

Figure 1 shows the kinds of residual plots generated by the three kinds of error. The examples are idealized in that they assume only one kind of error in each case, although in reality they will normally occur together. The random scatter shown in Figure 1(a) fills the available space more uniformly than one usually observes in practice: when the number of points is small they may often present a spuriously systematic appearance, and before interpreting a scatter as non-random one should check that the kind of non-randomness is consistent in repeated experiments.

With random error (Figure 1a), no rule predicts how any point will lie in relation to its neighbours; systematic error (Figure 1b) generates points on or close to an obvious line; rounding error (Figure 1c) results in a systematic arrangement at a local level, but with the whole plot broken into disconnected strips.

MATERIALS AND METHODS

Glyceraldehyde-3-phosphate dehydrogenase [glyceraldehyde-3-phosphate dehydrogenase (NADP⁺) (phosphorylating), EC 1.2.1.13] was isolated from spinach (*Spinacia oleracea*) chloroplasts prepared by the method of Jensen and Bassham (1966).

Phosphoenolpyruvate, 3-phosphoglycerate, ATP, NADPH, pyruvate kinase and phosphoglycerate kinase were from Boehringer Mannheim.

Reduction of 1,3-bisphosphoglycerate (generated *in situ* with phosphoglycerate kinase) by NADPH was monitored at pH 7.9 at 25 °C in a 0.2 cm cuvette containing 50 mM glycylglycine, 0.5 mM EDTA, 50 mM KCl, 9.3 mM Mg²⁺, 2 mM ATP, 2 mM phosphoenolpyruvate, 3 mM 3-phosphoglycerate, 2 units/ml pyruvate kinase, 5 units/ml phosphoglycerate kinase and 0.7 mM NADPH in a total volume of 0.4 ml. The reaction was started by addition of NADPH after incubation of the enzyme for 5 min at

* To whom correspondence should be sent.

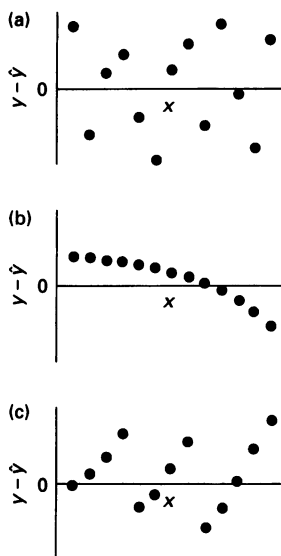


Figure 1 Residual plots expected from three types of error

(a) Random error; (b) systematic error; (c) rounding error. The difference between the observed (y) and calculated (\hat{y}) value of a measured quantity is plotted against an abscissa variable x . In the examples discussed in the present paper, x is normally the time, but it can be any variable.

25 °C with the assay mixture without NADPH. Absorbances were measured at 340 nm in a Uvikon 860 recording spectrophotometer from Kontron Instruments AG, Zürich, Switzerland. As progress curves for glyceraldehyde-3-phosphate dehydrogenase deviate from linearity over the whole range (see below), one cannot adequately estimate the initial rate by using the built-in software to fit a straight line to all or part of the data. Absorbances taken at intervals of 4 or 5 s, with integration periods of 2 or 3 s respectively, were therefore printed and used for subsequent curve-fitting. Although integration over a period comparable with the interval between measurements could itself be a source of error, calculations indicate that for the curves considered such errors would be less than 0.0001% and thus negligible.

Readers may not easily believe that the manufacturer of a precision instrument could recommend applying straight-line regression to data that do not fall on a straight line. However, the Instruction Manual for the Uvikon 860 spectrophotometer (Kontron Instruments, 1986) illustrates the recommended calculation with a Figure showing 30 absorbances from 0.948 to 0.374 measured at intervals of 3 s. The points deviate from a straight line much more obviously than any of the examples we use in the present paper, but there is no suggestion in the text surrounding the Figure that the calculation is inappropriate.

RESULTS

Experimental example

Figure 2 shows a 2 min trace for the glyceraldehyde-3-phosphate dehydrogenase reaction. As the rate varies during the time of measurement, it is inappropriate to estimate the initial rate by fitting a straight line, the only possibility offered by the software supplied with the instrument: one can restrict the range of points to be fitted, but there is no option to fit a curve of any kind. Selecting points may be valid if there is no deviation from linearity during a significant period, but in Figure 2 there is no identifiable moment at which there is a change from a straight

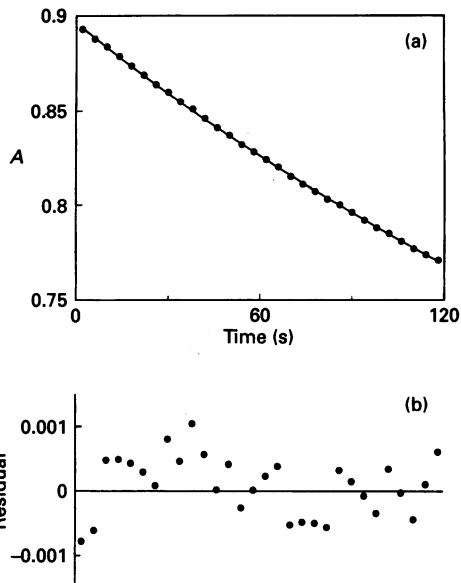


Figure 2. Reduction of 1,3-bisphosphoglycerate by NADPH

(a) The absorbance at 340 nm was measured as described in the text. The line was calculated from the equation $A = 0.896398 - 0.001312t/(515.5 + t)$. (b) The difference between observed and calculated absorbance is plotted against time.

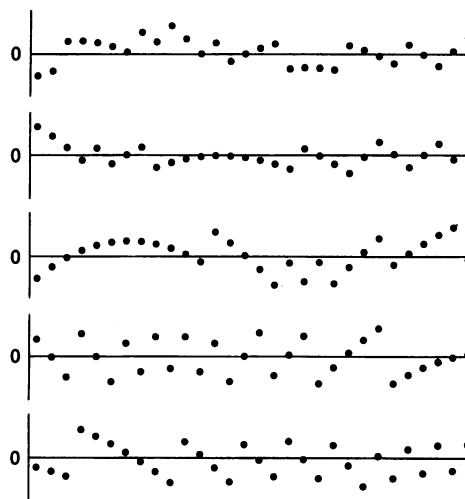


Figure 3 Five residual plots

Residual plots are drawn for five consecutive experiments carried out in the same way as in Figure 2. Units have no relevance to the interpretation of the plot and are omitted. For the top three plots the points were taken at intervals of 4 s, with an integration period of 2 s; for the other two they were taken at intervals of 5 s, with an integration period of 3 s.

line to a curve: the rate begins to decrease at zero time. Even if one were to assume that there was no diminution of rate during the first 15 s, fitting only the first five points would imply discarding any information in the rest of the trace.

It was therefore essential to work with the numerical values of the absorbance given by the spectrophotometer. A curve is shown in Figure 2 that is calculated from the following equation:

$$A = A_0 + v_0 t / (1 + t/\tau) \quad (1)$$

where A is the absorbance at time t , and the fitted parameters are as follows: A_0 (= 0.896398) is the absorbance at $t = 0$, v_0 (= -0.001312 s^{-1}) is the initial rate and τ (= 515.5 s) determines the degree of curvature. This type of equation generates a closer fit

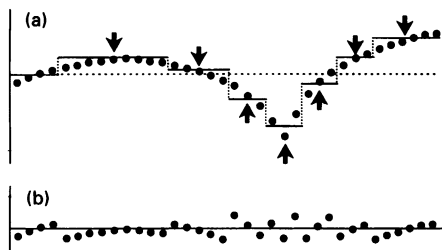


Figure 4 Rationalization of the characteristic appearance of a residual plot when rounding error is dominant

In (a) a series of points are drawn similar to a stroboscopic picture of a bouncing ball, and a step function is drawn through these as a baseline. Each strip, corresponding to one step of this line, is displaced in the direction marked by an arrow so as to eliminate the steps. The resulting distribution of points (b) then resembles a residual plot generated by rounding error.

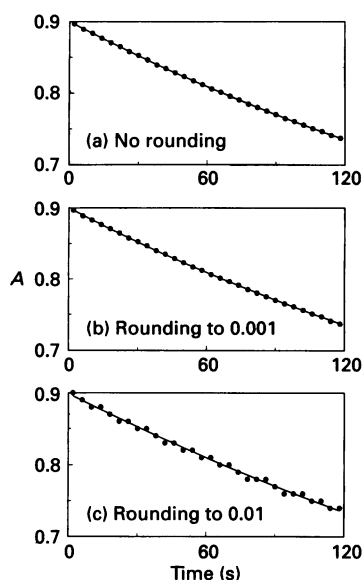


Figure 5 Simulated progress curves with three degrees of rounding

Absorbances were calculated from an integrated rate equation as described in the text, and normally distributed random errors with a mean of zero and an S.D. of 0.0003 were added. The resulting simulated observations were then fitted to eqn. (1) without rounding (a), or with rounding to 0.001 (b) or 0.01 (c). The curve shown in (b) represents approximately our interpretation of the real experiment in Figure 2.

to a more realistic equation than one can obtain with a quadratic, and it can be fitted much faster (potentially in real time without needing a fast arithmetic processor) than a realistic equation: it contains neither logarithms nor exponentials, and it expresses A in terms of t , whereas integrated rate equations for enzymes are normally written in inverse form, expressing t in terms of A , so that determining A requires solution of a non-linear equation at each value of t .

Although the line in Figure 2(a) apparently fits the data precisely, one cannot immediately judge whether there are systematic deviations. In the residual plot (Figure 2b) the deviations are magnified by a factor of more than 30; systematic effects are slight and random scatter is negligible, but there is an obvious pattern that cannot be adequately explained in terms of either Figure 1(a) or 1(b): four sets of four or five neighbouring points (3–7, 14–17, 18–21, 22–25) appear to lie on smooth curves.

As the apparent non-randomness seen in Figure 2(b) could be a freak phenomenon, we have examined numerous other residual

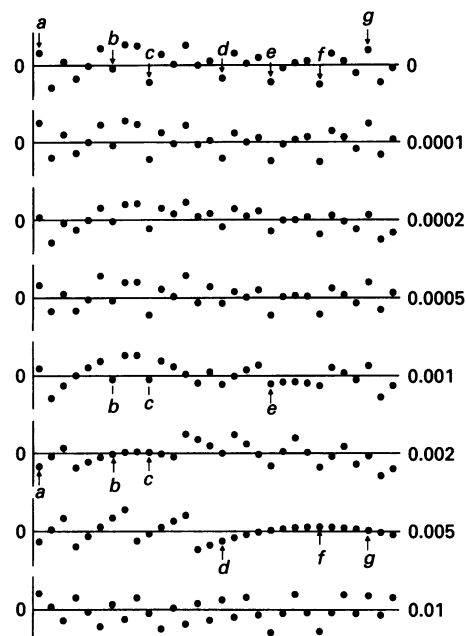


Figure 6 Residual plots from the simulation shown in Figure 5

Each plot is labelled at the right with the level of rounding. This Figure shows five additional cases that were omitted from Figure 5 as they would not have been distinguishable at the resolution of that Figure from the ones shown.

plots produced in similar experiments, always with essentially the same results. In none of the five examples shown in Figure 3 can the distribution of points be explained by either systematic or random error. The top three plots derive from measurements made with a period of integration of 2 s, not much greater than the minimum of 1 s recommended by the manufacturers. To check that this did not cause the patterns observed, the lowest two plots were done with an integration time of 3 s, but the patterns are still evident with this longer period of integration.

The patterns in Figure 3 may not all seem to be identical, but all can be rationalized by the sort of transformation illustrated in Figure 4. Starting with a pattern resembling a stroboscopic picture of a bouncing ball (Figure 4a), one can draw a step function through the points as a baseline and then displace the steps (with their associated points), so as to flatten the baseline (Figure 4b).

To demonstrate that this sort of residual pattern is characteristic of rounding error, one should ideally show that the pattern disappears when the calculations are done with more precise values of the absorbance, but the spectrophotometer does not permit this: there is no way to print, display or otherwise determine absorbances more precise than ± 0.001 . We therefore show that one can generate the same behaviour in simulated experiments.

Simulated example

Traces were simulated by solving for the concentration of product, p , at increasing times, t , using the integrated Michaelis–Menten equation:

$$V^{\text{app}} \cdot t = p - K_m^{\text{app}} \cdot \ln(1 - p/p_\infty) \quad (2)$$

in which V^{app} and K_m^{app} are the apparent limiting rate and Michaelis constant respectively, and p_∞ is the limiting value of p approached as t becomes large. The following parameters gave results similar to those in Figure 2: $V^{\text{app}} =$

-0.5 s^{-1} ; $K_m^{\text{app}} = -150$; $p_\infty = 0.5$, the absorbance A being calculated as $0.9 - p$. [Unlike the parent parameters V and K_m , V^{app} and K_m^{app} can be negative if the reaction product binds more tightly to the enzyme than the substrate (Henri, 1903; Huang and Niemann, 1951; Cornish-Bowden, 1976).] Random errors from a normal distribution with mean zero and S.D. 0.0003 were added to the calculated absorbances, which were then used as calculated, or rounded to various extents over the range 0.0001–0.01. Three representative progress curves are shown in Figure 5, and the whole set of eight residual plots is shown in Figure 6.

Although the observations were generated by means of eqn. (2), a theoretically realistic equation, the progress curves were calculated by fitting eqn. (1). One might expect to see some evidence of systematic error, therefore, but none is obvious, even for the results calculated without rounding. This confirms that eqn. (1) is capable of giving an accurate approximation to a real progress curve, at least for the sort of numerical values used.

Figures 5(a) and 5(b) are barely distinguishable to the eye, but the corresponding residual plots (those labelled 0 and 0.001 in Figure 6) are recognizably different, and the plot labelled 0.001 is qualitatively quite similar to the experimental plots shown in Figures 2(b) and 3 (all of which used data rounded to ± 0.001 , with unknown levels of random error). Thus the form of these plots was caused primarily by rounding error, and the true optical precision of the spectrophotometer was better than ± 0.001 .

The true precision was probably better than ± 0.0003 , because the simulated examples in Figure 6 suggest that a residual plot does not lose all obvious random character until the level of rounding is well above the level of random error: the seven points a to g in plot 0, each of which is well separated from its immediate neighbours, are all recognizable in plot 0.0005, and three (d , f and g) do not disappear from view until a rounding level of 0.005. Indeed, even without knowledge of the true random arrangement one should recognize points b and c as anomalous in plot 0.001, as they fail to lie on the same line as their eleven neighbours: no corresponding anomalies can be seen in any of the experimental examples (Figure 3), suggesting that the random error in these examples was less than ± 0.0003 .

With very coarse rounding, as in Figure 5(c), its effects are obvious even in the progress curve, and the peculiar arrangement of points ought to be noticed even without plotting residuals. However, a residual plot allows rounding effects to become visible at a level of rounding about an order of magnitude finer.

DISCUSSION

If the main interest in analysing progress curves is to estimate initial rates, the consequences of rounding are minor, as one may see from Table 1: the true initial slope (calculated by differentiating eqn. 2 with the parameter values given and setting t at zero) was $-1.672 \times 10^{-3} \text{ s}^{-1}$, and the estimates show very little tendency to get worse as the rounding becomes coarser. Moreover, even the worst value is much better than the value of about $-1.4 \times 10^{-3} \text{ s}^{-1}$ if one relied on the software supplied with the spectrophotometer and fitted a straight line to the entire data set.

Unavoidable rounding error becomes more than a minor irritation if one is not merely estimating an initial slope (in which case a precision instrument is hardly needed), but trying to

Table 1 Initial-slope estimates for different levels of rounding for the data of Figure 5

The data shown in Figure 5(a) were rounded to the precision indicated in each line of the Table, and then fitted to eqn. (1) by the non-linear least-squares method.

Rounding level	$10^3 \times v_0 \text{ (s}^{-1}\text{)}$
0	-1.687
0.0001	-1.677
0.0002	-1.709
0.0005	-1.688
0.001	-1.691
0.002	-1.714
0.005	-1.657
0.01	-1.606

establish the true form of the equation, a crucial stage in discrimination between theoretical models. Eqns. (1) and (2), although very different in form, generate curves that are virtually indistinguishable when compared at a precision of ± 0.0003 . To distinguish between them, one must work at the limits of experimental precision, an impossible task if there is an unnecessary and arbitrary precision limit to the data display that is much coarser than the true precision of the measurements.

At the point of manufacture inaccuracies derived from rounding error are almost ludicrously simple to eliminate, as it is vastly simpler to increase the precision of the computation by an order of magnitude than it would be to achieve a comparable improvement in optical and photometric precision. Thus they would be overcome from one day to the next if instruments were supplied with software capable of matching the accuracy of the measurements; until this is done a large part of the investment in obtaining an instrument of high accuracy is wasted. However, appropriate pressure will only be brought to bear on manufacturers if users can easily recognize when the software provided is inadequate. Significant rounding error should never be tolerated, as it can in principle be made as small as desired.

It would be interesting to know whether the problem is specific to the particular spectrophotometer used, or whether it is more general. If it arose out of a perception that improving the technical quality of an instrument is a difficult task worthy of major investment of time and effort, whereas writing the accompanying software is almost trivial, one may fear that similar problems may be widespread, but at present there is no information about this.

REFERENCES

- Cornish-Bowden, A. (1976) Principles of Enzyme Kinetics, pp. 142–152, Butterworths, London
- Draper, N. R. and Smith, H. (1981) Applied Regression Analysis (2nd edn.), Wiley, New York
- Ellis, K. J. and Duggleby, R. G. (1978) Biochem. J. **171**, 513–517
- Henderson, P. J. F. (1992) in Enzyme Assays: a Practical Approach (Eisenthal, R. and Danson, M. J., eds.), pp. 277–316, IRL Press, Oxford
- Henri, V. (1903) Lois Générales de l'Action des Diastases, Hermann, Paris
- Huang, H. T. and Niemann, C. (1951) J. Am. Chem. Soc. **73**, 1541–1548
- Jensen, R. G. and Bassham, J. A. (1966) Proc. Natl. Acad. Sci. U.S.A. **56**, 1093–1101
- Kontron Instruments (1986) Uvikon 860 Instruction Manual (version 05.86), section 3.3.6, p. 44, Kontron Instruments AG, Zürich
- Nimmo, I. A. and Atkins, G. L. (1981) in Kinetic Data Analysis (Endrenyi, L., ed.), pp. 309–315, Plenum, New York