EVALUATION OF UNCERTAINTY IN DIGITAL PROCESSING OF QUANTIZED DATA

Nicola Locci, Carlo Muscas, Emilio Ghiani

Dipartimento di Ingegneria Elettrica ed Elettronica, Università di Cagliari, Italy
Phone (39) 0706755860 Fax (39) 0706755900 e-mail: carlo@diee.unica.it

Abstract – *The paper deals with the uncertainty in measurement based on digital signal processing algorithms, like those achievable with the virtual instruments. The correct estimation of bias and uncertainty is discussed with reference to a simple case study. Three possible approaches to this question are examined and compared.*
It is shown how a Monte Carlo method, based on numerical simulations and implemented with commercial software packages, can allow virtual instruments to perform an auto-evaluation of both bias and uncertainty affecting their results. Some theoretical considerations, computer simulations and experimental tests are shown to support the proposed technique.

Keywords - measurement uncertainty, digital signal processing, virtual instruments, sampled data.

1 INTRODUCTION

Modern instrumentation requires in many cases the digital processing of sampled data. This is the case, for instance, of the virtual instruments. As a consequence, the widespread use of measurement algorithms on digitized data generates the need to analyze the practical usability of these data, with reference to the reliability of the final result of the measurement. In other words, the quality of the measurement must be assessed as a function of the uncertainties that affect the input data. Formally speaking, it is necessary to state the uncertainty that can be associated to the measurement result, according to the recommendations of the ISO "Guide to the expression of uncertainty in measurement" (GUM) [1].

The GUM describes quite clearly the main goals to be reached in the statement of the measurement uncertainties and the main directions that should be followed to pursue this goals. However, as far as practical problems are considered, owing to the variety of specific situations, several different approaches, more or less complicated, can be required. For instance, in the ambit of accredited laboratories, the EA (*European co-operation for Accreditation*) published the document [2], whose purpose is to state the specific demands in reporting uncertainty of measurement on calibration certificates, in compliance with the recommendations of [1].

This paper aims to provide practical indications to evaluate the uncertainty in measurements performed by applying computational algorithms to sampled and digitized data. The contribution to the uncertainty given by the acquisition and A/D conversion system and its propagation through the measurement algorithm will be considered. The specifications provided by the manufacturer of the acquisition unit will be used to extract information about the statistical parameters (class B uncertainties, according to the GUM) to be used for the estimation of the combined uncertainty.

First, three possible approaches to this problem are presented: the strict application of the GUM, an analytical procedure and a numerical method based on computer simulations. Their advantages and limitations are discussed, also on the basis of a numerical comparison. Since the numerical technique seems to be the more flexible, it will be chosen for the final practical applications, which consists of virtual instruments able to perform both the correction of the bias and the statement of the uncertainty besides the measurement result.

2 EVALUATION OF THE UNCERTAINTY

2.1 *The international guide*

Several aspects should be taken into account while applying the GUM's recommendations, and its limitations should be considered carefully in order to achieve reliable uncertainty evaluations.

First, as clearly stated in the GUM, the rule for the calculation of the standard uncertainty of the output estimate (combined uncertainty) arises from a first order approximation of the Taylor's series. Indeed, the estimation of the combined uncertainty $u_c(y)$ of an indirect measurement $y = f(x_1,x_2,..,x_N)$, as a function of both the standard uncertainty $u(x_i)$ affecting each input variable and the covariance $u(x_i,x_j)$ associated with them, can be performed by applying the well-known expression:

$$u_c^2(y) = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i,x_j). \quad (1)$$

The second term of the left hand expression in (1) is null in the presence of totally uncorrelated input quantities.

An approximate expression is also provided to consider higher order terms, but it is valid only under some restrictive hypothesis (e.g. the input quantities must be

ISBN 972-98115-4-7

uncorrelated and have distributions symmetrical with respect to their mean value). Therefore, special care should be used in the presence of nonlinear functions, for which the approximation implicit in (1) might introduce significant errors.

Furthermore, the GUM supposes the measurement result to be corrected from any significant systematic effects (*bias*), and that every possible effort aimed at identifying such effects has been done. But in many practical cases, especially when the measurand is evaluated by means of computational algorithms, it is difficult to understand *a priori* whether an uncertainty source (even with null mean value) introduces a bias in the final result or not. Even more difficult is determining a quantitative value of this bias.

2.2 A case study

A simple example may help to clarify the problems mentioned above. Let us consider a periodic signal $x(t)$, of which N samples x_k (k= 1..N) per period are acquired, and define an output function y as the mean square value of x:

$$y = \frac{1}{N}\sum_{k=1}^{N} x_k^2 \tag{2}$$

This function has been chosen in virtue of its simplicity and can be considered as the fundamental expression for the digital calculation of other widely used quantities, such as root mean square values and power related terms (e.g. power losses due to the Joule effect).

For the sake of simplicity, we make the hypothesis that each sample is only affected by the quantization noise, expressed through the random variable δx_k. This source of uncertainty is always present in analog to digital conversion systems, and the corresponding random variable can be represented with a uniform distribution in the range [$-q/2$, $+q/2$], where q is the quantization step. Thus the mean value of this variable is zero and its variance is $q^2/12$.

The input estimates, i.e. the estimated values for each sample (as defined in [2]), are $\tilde{x}_k = x_k + \delta x_k$, and the output estimate is:

$$\tilde{y} = \frac{1}{N}\sum_{k=1}^{N} \tilde{x}_k^2 = \frac{1}{N}\sum_{k=1}^{N} \left(x_k + \delta x_k\right)^2 \tag{3}$$

The random variables δx_k can be considered to be statistically independent of each other. Therefore, in order to calculate the standard uncertainty of the measurement u_y, i.e. the combined uncertainty expressed in terms of standard deviation, the application of (1) provides the following value:

$$\begin{aligned} u_y &= \sqrt{\sum_{k=1}^{N}\left(\frac{\partial y}{\partial x_k}\right)^2 u_{xk}^2} = \sqrt{\sum_{k=1}^{N}\left(\frac{1}{N}2x_k\right)^2 u_{xk}^2} \\ &= \sqrt{\frac{4}{N}\left[\frac{1}{N}\sum_{k=1}^{N} x_k^2\right]\frac{q^2}{12}} = \sqrt{\frac{1}{3}y\cdot\frac{q^2}{N}} \end{aligned} \tag{4}$$

This standard deviation is proportional to y, whereas decreases when the number of samples N increases. It is absolutely singular to observe that in the presence of a null signal (y=0), the corresponding uncertainty is null. This evident incongruence is due to the nonlinearity of the function y.

2.3 The analytical approach

A rigorous and complete analytical approach can be used sometimes to evaluate the uncertainty of measurement. This procedure is possible in this case, since the function is simple and we considered only the uncertainties arising from quantization. The standard deviation of the output estimate $\tilde{u}_y$ can be therefore assessed by means of suitable analytical developments (see Appendix):

$$\tilde{u}_y = \sqrt{\frac{1}{3}y\frac{q^2}{N} + \frac{1}{180}\frac{q^4}{N}}\,, \tag{5}$$

The additive term in (5), which is independent of y, allows us to evaluate correctly the uncertainty even in the presence of very low values of the measured quantity y.

In addition, by considering the expressions (A4) of the appendix, it can be shown that the expected value $E\{\tilde{y}\}$ of the output estimate is affected by a bias, equal to the variance of the quantization noise [3]:

$$E\{\tilde{y}\} = y + \frac{1}{N}\sum_{k=1}^{N} E\left\{\delta x_k^2\right\} + \frac{1}{N}\sum_{k=1}^{N} E\{2x_k\delta x_k\} = y + \frac{q^2}{12} \tag{6}$$

Unfortunately, despite its formal correctness, the analytical method is difficult to apply when the measurement algorithms become more complex, and/or when different causes of uncertainty must be considered at the same time.

2.4 The numerical simulation

A possible solution to this problem can be represented by methods based on numerical simulations, such as those recently proposed in [4-7]. These methods exploit the capability of several commercial software packages to generate sequences of random numbers characterized by prefixed statistical parameters (distribution, mean, variance, etc.).

This feature allows the statistical parameters of the output quantity to be evaluated by means of a Monte-Carlo procedure [7]. First, the specifications of the acquisition system, provided by the manufacturer, can be used as uncertainties evaluated by means of the Type B method, according to the GUM. A suitable distribution is then assigned to these uncertainty terms, which can be numerically represented by sequences of random variables defined by the software package. Then, a large number of simulated tests can be performed by applying to the input data the known measurement function or algorithm. In each test the samples are corrupted by the different contributions, whose values are extracted from the above populations. Finally, a sequence of output values is obtained, whose standard deviation represents the standard uncertainty of the measurement result. In addition, the bias can also be calculated as the difference between the mean value of the output sequence and the output value obtained when all the uncertainty contributions are null.

The accuracy of the numerical approach depends on the good quality of the random number generator. Preliminary tests have confirmed that many commercial software packages are suitable for this application, provided that a sufficient number of extractions is performed. In the following examples the method has been implemented by performing one thousand extractions of random values with the software package LabView.

2.5 Numerical comparison between the three methods

To verify and emphasize the above considerations, the three methods (combined uncertainty assessed according to the GUM, analytical method and numerical simulations) have been applied to evaluate the uncertainty of the output function y previously defined in (2). In this example the input data are the samples x_k of a sinusoidal waveform, acquired through an ideal 8 bit data acquisition board with 10 V full scale range. In the tests, the amplitude of the signal has been changed from very low values up to the full scale.

While the GUM does not provide indications on how to evaluate the bias in this situations, according to the analytical method the bias can be calculated by means of (6): $q^2/12 = 5\cdot10^{-4}$ V^2. The values achieved with the numerical simulations were in all tests very close to the theoretical one, with differences always lower than $5\cdot10^{-6}$ V^2, and therefore definitely negligible.

As for the uncertainty, Fig. 1 shows the difference between the standard deviations evaluated by means of the three methods as a function of the signal amplitude. The curves relevant to the analytical and numerical methods are practically superimposed, while the error introduced by (4) in the GUM's method is significant for very low values of y.

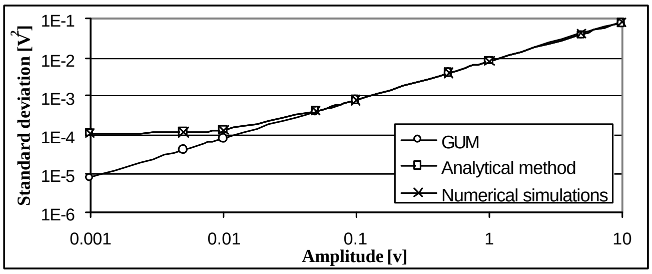


Fig.1 – Comparison between three methods to evaluate the uncertainty.

2.6 Other sources of uncertainty

In addition to quantization, other causes of uncertainty have to be always taken into account when analog quantities are converted into digital data. The most important of these sources are noise, offset and gain error, whose values can be again found in the specifications of the acquisition system.

The noise contribution can be dealt with exactly in the same way of quantization. On the contrary, both gain error and offset should be assumed to be totally correlated variables, i.e. the random variables that represents these contributions affect each sample of the acquired signal to the same extent. This assumption is realistic if short-term acquisitions are performed.

By considering each contribution separately and referring to (3), the offset introduces a constant additive term (δx_k constant), whereas the gain error causes δx_k to be proportional to x_k, with constant multiplicative factor.

In this way, the analytical approach can be used to evaluate bias and uncertainty introduced by each oh the above contributions in the calculation of the mean square value (2). The mathematical procedure is similar to the one presented in the appendix for the quantization. The results are shown in Table 1, where all the distributions have been assumed rectangular, according to the GUM's indications on Type B uncertainties. Given that the input signal is usually a voltage, the ranges of noise and offset have been expressed with their measurement unit (volt), whereas, the gain error is provided as a relative value (*per unit*). X_m is the mean value of the acquired signal x.

Table 1 – Calculated bias and uncertainty for different sources

	Range	**Bias**	**Standard deviation**
Noise	$\pm n$ [V]	$\frac{n^2}{3}$	$\sqrt{\frac{4}{3}y\frac{n^2}{N}+\frac{4}{45}\frac{n^4}{N}}$
Offset	$\pm o$ [V]	$\frac{o^2}{3}$	$\sqrt{\frac{4}{3}X_m{}^2 o^2+\frac{4}{45}o^4}$
Gain error	$\pm g$ [p.u.]	$y\cdot\frac{g^2}{3}$	$\sqrt{y^2\left(\frac{4}{3}g^2+\frac{4}{45}g^4\right)}$

The numerical simulations have been applied also to the single sources of uncertainty described in this subsection. The results are again in excellent agreement with those given by the expressions of Table 1.

When probability distributions other than the uniform have to be considered, the complexity of the analytical calculations may increase significantly. On the other hand, the numerical approach shows insensitive to this problem, given that the software packages used to generate the populations of random numbers are usually able to deal with the most common distributions (e.g. the Gaussian one). Furthermore, if several independent observations are made for one of the input quantities to apply the Type A evaluation of the uncertainty [1], the set of the results constitutes a population of values that can be directly used in the numerical simulations.

In addition, when two or more contributions are present at the same time, as happens in real situations, it becomes critical understanding how the corresponding biases and uncertainties act in combination with each other.

3 APPLICATION TO VIRTUAL INSTRUMENTS

The numerical method is especially suitable to evaluate bias and uncertainty in measurements performed by means of virtual instruments. In this case the complexity of the implemented algorithms is not a major problem, and all the most important sources of uncertainty can be considered, provided that their distributions have been defined by means of either Type A or Type B evaluations.

In addition, when measurements are performed off-line, the virtual instrument itself can be designed purposely to follow two distinct procedures: the first one to achieve the measurement result, the second one to implement the numerical approach in order to calculate the bias (and eventually correct the result) and to provide directly in the front panel also the uncertainty associated to this result.

The same goal can be achieved also for on-line measurements, when the observation frequency is so low as to allow the numerical procedure to run between two consecutive evaluation of the measurand. This situation could be verified, for instance, when very slow events have to be monitored (e.g. environmental measurements), or when *single shot* measurements have to be performed at prefixed time intervals.

It is clear that in this kind of applications the input data of the numerical procedure are the acquired samples, i.e. quantities which are already affected by measurement uncertainty. Some doubts could therefore arise concerning the correctness of using these data as reference quantities of the method. In the opinion of the authors, this operation should be considered formally correct, since it complies with the sense of the international guide [1], which often uses a very similar concept.

In fact, in order to estimate the combined uncertainty of an indirect measurement $y = f(x_1,x_2,..,x_N)$ by means of (1), it is necessary to evaluate the so-called sensitivity coefficient, i.e. the partial derivative $\frac{\partial f}{\partial x_i}$ of the model function f with respect to the input variables. According to the strict rule, these derivatives should be evaluated in correspondence of the expected values of the input quantities. Nevertheless, they are usually calculated at the input estimates x_i, as the GUM itself suggests. In the same way, when the numerical procedure is applied to evaluate the uncertainty in virtual instruments, the reference data should be theoretically the expected values of the acquired samples. But, owing to the impossibility of an exact knowledge of these values, from a practical point of view the procedure can be applied by considering as initial data the measured values, without losing the conceptual validity of the method.

4 EXPERIMENTAL VALIDATION

In order to verify the effectiveness of the numerical method for the evaluation of the uncertainty in virtual instruments, experimental tests have been performed on a low-cost data acquisition board (AT-MIO16E-10 by National Instruments, max sampling rate 100 kS/s, 12 bit resolution) plugged in a PC and controlled by means of the LabView 5.1 software package. The board acquires and digitizes the signal supplied by a precision calibrator (Fluke 5720) and the virtual instrument calculates its RMS value.

Only low frequency sinusoidal signals have been considered (50 Hz and 100 Hz), so that the main causes of uncertainty were offset, gain error, nonlinearity, quantization and noise, whereas problems like aliasing, leakage, slew rate, etc., could be neglected. The actual values used to define these Type B uncertainties have been extrapolated from the specifications provided by the board's manufacturer (relevant to the selected range of ± 10 V) and are summarized in Table 2.

Table 2 – Data acquisition board specifications

Offset	*Gain error*	*Nonlinearity*	*Noise*
6.4 mV	0.072 %	2.4 mV max	3.5 mV

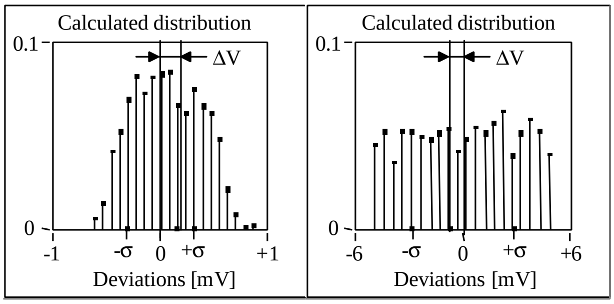
Three different sample rates (256, 512 and 1024 samples per period) and three different amplitudes of the input signal, corresponding to 10 %, 50 % and 100 % of the full scale, were considered.

In all tests the bias of the measurement was found to be negligible compared to the corresponding uncertainty. The difference $\Delta V = V_{out} - V_{ref}$ between the output values of the measurement algorithms applied to the acquired data and the reference values of the calibrator have been compared with the uncertainties (expressed in terms of the standard deviation σ) achieved with the calculation method, implemented in the same virtual instrument. In all tests the ratio $|\Delta V|/\sigma$ was found to be always less than one.

Another useful feature of the numerical method is that not only the standard deviation can be determined but also the distribution of the output quantity can be graphically represented, thus providing further information. As an example, fig. 2 shows the histogram numerically determined in two tests at 100 Hz, with amplitude equal to 10 % and 100 % of the FSR, respectively. It can be observed that in the presence of low signals different sources of uncertainty (having uniform distribution) contribute to a similar extent to the overall uncertainty, thus leading to a nearly gaussian output distribution (fig. 2a). On the contrary, when the input signal approaches the full scale, the distribution tends to become uniform (fig. 2b), since the term proportional to the reading (gain error) prevails over the others. In both cases ΔV is included in the range $\pm\sigma$.

5 CONCLUSIONS

Evaluating the bias and the uncertainty of measurements performed by digitally processing sampled data can be a difficult task, owing to both the complexity of the algorithms used and the number of different causes that affect the accuracy of the result. In this connection, three



a) b)

Fig.2 – Histograms of the output deviations for different amplitudes of the input signal: a) 10 % of FSR, b) 100% of FSR

possible approaches have been analyzed to determine how the uncertainty on the input data propagates in the measurement function: the first is based on the application of the international guide to the expression of uncertainty in measurement; the second involves rigorous analytical procedures; the third consists of numerical simulations. The latter method seems to be the most suitable for virtual instrumentation, since its applicability is not influenced by the complexity of the measurement algorithm and by the number of uncertainty sources affecting the input samples. A simple case study has been used to emphasize these considerations.

As an example of application, a virtual instruments capable of compensating the eventual bias and stating the uncertainty of the measurement result has been realized. A first series of experimental tests, in which the nominal specifications of the acquisition system have been used as type B uncertainty affecting the input data, confirm the practical effectiveness of the proposed method.

REFERENCES

[1] "Guide to the expression of uncertainty in measurement", ISO, 1995.

[2] Publication EA-4/02: "Expression of the Uncertainty of Measurement in Calibration", European co-operation for Accreditation, December 1999.

[3] I. Kollar: "Bias of mean value and mean square value measurements based on quantized data", IEEE Transactions on Instrumentation and Measurement , Volume 43, Issue 5 , Oct. 1994.

[4] A. Passavanti, F. Francisci: "Modelization and propagation of uncertainty through digital signal processing algorithms" (in Italian), Tutto_Misure, N. 2, 2000, pp. 60-65.

[5] S. Caldara, S. Nuccio, C. Spataro: "Measurement uncertainty estimation of a virtual instrument", IEEE IMTC/2000, Baltimore (USA), May 2000.

[6] K. Chrzanowski, R. Matyszkiel, M. Szulim: "Calculations of uncertainty in measurements", XVI IMEKO World Congress, Wien, September 2000.

[7] N. Locci, C. Muscas, L. Peretto, R. Sasdelli: "A numerical approach to the evaluation of uncertainty in nonconventional measurements on power systems", IEEE IMTC/2001, Budapest (Hungary), May 21-23 2001, in press.

[8] G. Betta, C. Liguori, A. Pietrosanto: "Propagation of uncertainty in a discrete Fourier transform algorithm", Measurement, 27, (2000), pp. 231-239.

APPENDIX

Let us consider the variance of the output estimate (3) of the function (2):

$$\tilde{u}_y^{\,2} = E\left\{\left[\tilde{y} - E\{\tilde{y}\}\right]^2\right\} = E\left\{\tilde{y}^2\right\} - \left[E\{\tilde{y}\}\right]^2 \tag{A1}$$

If each sample is only affected by the random variable δx_k associated with the quantization, the first term of the last expression of (A1) is given by:

$$E\left\{\tilde{y}^2\right\} = E\left\{\frac{1}{N}\sum_{k=1}^{N}(x_k+\delta x_k)^2 \cdot \frac{1}{N}\sum_{j=1}^{N}\left(x_j+\delta x_j\right)^2\right\}$$

$$= E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[x_k^2+2x_k\delta x_k+(\delta x_k)^2\right]\cdot\left[x_j^2+2x_j\delta x_j+\left(\delta x_j\right)^2\right]\right\}$$

$$= E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\begin{bmatrix} x_k^2x_j^2+2x_k^2x_j\delta x_j+x_k^2\left(\delta x_j\right)^2+ \\ 2x_k\delta x_kx_j^2+4x_k\delta x_kx_j\delta x_j+2x_k\delta x_k\left(\delta x_j\right)^2+ \\ (\delta x_k)^2x_j^2+2(\delta x_k)^2x_j\delta x_j+(\delta x_k)^2\left(\delta x_j\right)^2 \end{bmatrix}\right\}$$

(A2)

The quantization noise has a uniform distribution in the range $[-q/2, q/2]$ and its central moment of h-th order is defined as:

$$E\left\{\delta x_k^{\,h}\right\} = \frac{1}{q}\int_{-\frac{q}{2}}^{\frac{q}{2}}\delta x_k^{\,h}\,d\delta x_k\ . \tag{A3}$$

It follows that the odd order moments are null, whereas:

$$E\left\{\delta x_k^{\,2}\right\} = \frac{q^2}{12}\quad ;\quad E\{\delta x_k 4\} = \frac{q^4}{80} \tag{A4}$$

Each term including random variables δx in the last expression of (A2) can be therefore evaluated separately:

$$E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[2x_k^2x_j\delta x_j\right]\right\} = E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[2x_j^2x_k\delta x_k\right]\right\} = 0$$

$$E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[2x_j\delta x_j\delta x_k^2\right]\right\} = E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[2x_k\delta x_k\delta x_j^2\right]\right\} = 0$$

$$E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[x_k^2\delta x_j^2\right]\right\} = E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[x_j^2\delta x_k^2\right]\right\} = y\frac{q^2}{12}$$

$$E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\left[4x_kx_j\delta x_k\delta x_j\right]\right\} = \frac{1}{3}\cdot\frac{yq^2}{N}$$

$$E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\delta x_k^2\delta x_j^2\right\} = \left(\frac{q^2}{12}\right)^2 + \frac{1}{N}\frac{q^4}{180} \tag{A5}$$

In addition, by recalling (6), we get:

$$\left[E\{\tilde{y}\}\right]^2 = \left(y+\frac{q^2}{12}\right)^2 = y^2 + 2\frac{q^2}{12}y + \left(\frac{q^2}{12}\right)^2 \tag{A6}$$

Finally, the variance (A2) can be calculated as:

$$\tilde{u}_y^{\,2} = \frac{1}{3}y\frac{q^2}{N} + \frac{1}{180}\frac{q^4}{N} \tag{A7}$$