



# A procedure for evaluating uncertainty associated with averaging or totalisation calculations on time-sampled data in gas energy measurements<sup>☆</sup>

Jože Kutin<sup>a,\*</sup>, Gregor Bobovnik<sup>a</sup>, Federica Gugole<sup>b</sup>, Adriaan M.H. van der Veen<sup>b</sup>

<sup>a</sup> University of Ljubljana, Faculty of Mechanical Engineering, Laboratory for Measurements in Process Engineering, Aškerčeva 6, SI-1000 Ljubljana, Slovenia

<sup>b</sup> VSL National Metrology Institute, Department of Data Science and Modelling, Thijssseweg 11, NL-2629 JA Delft, the Netherlands

## ARTICLE INFO

### Keywords:

Gas energy measurement  
Time-sampled data  
Deterministic variations  
Random variations  
Time-domain filtering  
Measurement uncertainty

## ABSTRACT

The measurement uncertainty attributed to gas energy measurements includes, among others, a contribution for the calculation of the average/total from time-sampled data. The aim of this study is to establish and verify a procedure for the evaluation of this component of uncertainty, which considers contributions related to the deterministic and random variations of the measured data. The first step of the proposed procedure is a separation of deterministic and random components via time-domain filtering, followed by the evaluation of the uncertainty associated to the deterministic component by means of a decimation method and of the uncertainty associated to the random component by statistical analysis. The procedure concludes with the calculation of the combined uncertainty due to these effects. Basic guidelines for the optimal selection of procedure parameters related to filtering and decimation are discussed. Demonstration and verification of the proposed procedure on experimental time-sampled data for the superior calorific value and on artificially generated data confirms that the proposed procedure enables reliable and robust evaluation of the uncertainty.

## 1. Introduction

Managing consumption of energy gases requires an appropriate metrological infrastructure along the entire supply chain, from production to storage and end use, with special challenges arising from the incorporation of new renewable gases, such as hydrogen and biogas. Typical measurands for fiscal metering are volume, mass, energy and composition [1–3]. Calculation units convert data from measuring instruments (e.g., flow meters, temperature and pressure transmitters, gas chromatographs) and calculate their average/total over certain periods of time. In addition to the uncertainties associated with the measuring instruments and the semi-empirical correlations for gas thermophysical properties, the overall measurement uncertainty also includes contributions due to the calculation of the average or total from the time-sampled data [4–8]. In this paper, this component of measurement uncertainty is referred to as the calculation uncertainty.

The calculation uncertainty shows different dependencies on the number of samples for random and deterministic components in time-sampled data. For random variations, this uncertainty is inversely

proportional to the square root of the number of samples, like the standard uncertainty of an arithmetic mean computed from a series of indications [9]. On the other hand, the error due to the numerical integration (quadrature) of deterministic variations is inversely proportional to the number of samples in the case of using the rectangle rule approximation [10]. Consequently, random and deterministic components require different methods for uncertainty evaluation. Time-sampled data in gas quantity and energy measurements generally comprises both random and deterministic variations. In the procedure proposed by the authors of this paper, Savitzky-Golay (S-G) time-domain filtering [11] is used to separate the two components, where the part of the signal that passes the filtering is treated as the deterministic component, whereas the removed part is considered as the random component. S-G filtering represents a generalized moving average method based on least squares polynomial fitting across a moving window. Unlike the classic moving average, the S-G filter can better preserve the local shape of the signal, including trends and derivatives, so this makes it more suitable for applications where signal characteristics like peaks, valleys, and slopes are important.

<sup>☆</sup> This article is part of a special issue entitled: 'MEASUR\_XXIV IMEKO World Congress' published in Measurement.

\* Corresponding author.

E-mail address: [joze.kutin@fs.uni-lj.si](mailto:joze.kutin@fs.uni-lj.si) (J. Kutin).

The uncertainty associated with the deterministic component is evaluated by the decimation method. Decimation is the process of down sampling time-sampled data by an integer decimation factor [12]. The decimation method implemented in this paper is based on the analysis of changes in calculated values for different decimation factors and through this predict the reference value, i.e. the value assigned to a continuous signal with an infinite number of samples. The uncertainty associated with the random component is evaluated by statistical analysis, both by considering uncorrelated and correlated data [13,14]. The autocorrelation is partially introduced into the deterministic component because of employing time-domain filtering for its determination. To account for the correlation effects in the evaluation of uncertainty, there are various methods to determine significant coefficients in the autocorrelation function (ACF): based on the first transit through zero (FTZ method) [13], based on the last significant nonzero element defined by the confidence interval (LSN method) [13,14] or by the partial ACF [15], etc. In this paper, the FTZ method is used in the uncertainty analysis when assessing correlation effects.

The research area in this paper is connected to the field of application of statistical models for time series in metrology, see JCGM GUM-6, Section 11.7 [16]. The procedure that we present in the paper aims at evaluating the whole calculation uncertainty associated with the estimation of the average (or any other aggregated quantity, for what matters) in fiscal metering. It does not make any assumption on the stationarity of the data, and it can be applied also to non-stationary time series. ARIMA (auto-regressive, integrated, moving average) models mentioned in [16] are only applicable to (statistically) weakly stationary time series, so they are useful, e.g., for determining the correlations and thus the resulting uncertainty contribution from the separated random component. Other parts of the presented procedure go beyond the use of ARIMA models, especially the proposed decimation method with subsequent linear regression applied to the deterministic component.

The current paper represents a significantly extended version of the IMEKO 2024 conference paper by the same authors [17]. A completely new study was conducted, verifying the developed procedure through the analysis of artificially generated data sets with different frequency contents and different magnitude ratios between deterministic and random components. Furthermore, this paper presents an illustrative test case analysis of an experimental data set, obtained from a different practical scenario than in the conference paper, which further confirms applicability of the proposed procedure. Taking into account the findings from the analysis of experimental and artificial data sets, guidelines for the selection of optimal procedure parameters have been established.

The paper is structured as follows. Section 2 outlines the basic steps of the proposed procedure for evaluating uncertainty associated with averaging or totalisation calculations on time-sampled data. Section 3 illustrates the procedure functionality on experimental data for the superior calorific value from actual gas energy measurements. Section 4 is focused on verifying the procedure using different artificial data sets and establishing guidelines for optimal selection of procedure parameters. The conclusions are summarized in Section 5.

## 2. Procedure for evaluation of the calculation uncertainty

An outline of the proposed procedure for evaluating uncertainty associated with calculation of average/total from time-sampled data is presented in Fig. 1. The procedure is applied to the input data in the form of a time series, i.e., a series of indications of a measuring instrument over a specified period of time. The order of the indications matters, i.e., the indications are not assumed to be independent or exchangeable. Although the uncertainty components due to the measuring instruments and their calibration are not discussed in this paper, it is assumed that the said time series is corrected for any known systematic effects, such as those from instrument calibration.

Signal processing and other evaluations within the procedure are performed in Mathematica, ver. 13.2. In the first step, the S-G moving average filter (Mathematica, built-in smoothing kernel SavitzkyGolay-Matrix) is used to separate the deterministic and random components of the time-sampled input data,  $q_i, i = 1 \dots N$ . The setting parameters of the S-G filter are the order of the smoothing polynomial (set to 2 in this work) and the size of the smoothing window in terms of the number of samples  $N_{win}$  on either side of centre point (smoothing window contains  $(2N_{win} + 1)$  samples). The part of the signal that passes through is considered as the deterministic component,  $q_{det,i}, i = 1 + N_{win}, \dots, N - N_{win}$ , and the removed part of the signal is considered as the random component  $q_{ran,i} = q_i - q_{det,i}, i = 1 + N_{win}, \dots, N - N_{win}$ .

The contribution to the calculation uncertainty, which is associated with the numerical integration of the deterministic component  $q_{det,i}$ , is estimated using the decimation method. Decimation or downsampling with the decimation factor  $n_{dec}$  means that only every  $n_{dec}^{th}$  sample is taken from the observed data  $q_{det,i}$ . The sought total or average  $Q$  is calculated for different decimation factors,  $Q(n_{dec}), n_{dec} = 1 \dots N_{dec}$ , and its dependence on the decimation factor is approximated by fitting the function  $Q_{fit}(n_{dec}) = a n_{dec} + b$  using ordinary least squares (OLS) (Mathematica, built-in function LinearModel Fit). Using a linear regression presumes that the rectangle-rule integration error is inversely proportional to the number of samples. The intercept  $b = Q_{fit}(n_{dec} = 0)$  is considered as a prediction of the reference value for the case of an infinite sample rate. Thus, the error of the value  $Q(n_{dec} = 1)$  obtained by integration of the original, undecimated deterministic component is estimated as:

$$e_{det} = Q(1) - Q_{fit}(0), \quad (1)$$

with its standard uncertainty equal to the standard deviation of the parameter  $b = Q_{fit}(0)$  of the OLS:

$$u(e_{det}) = s(Q_{fit}(0)) \quad (2)$$

The standard uncertainty of calculation associated with the deterministic component is determined as:

$$u_{det} = \sqrt{\left(\frac{e_{det}}{\sqrt{3}}\right)^2 + u^2(e_{det})}, \quad (3)$$

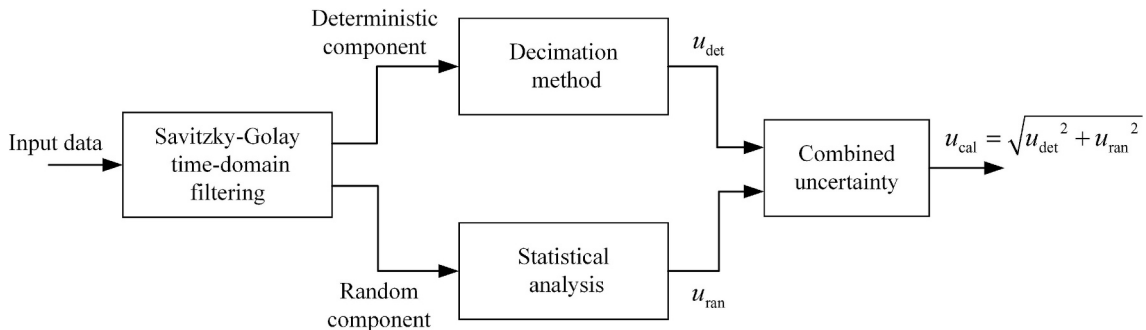


Fig. 1. Schematic diagram of the procedure for evaluating the calculation uncertainty.

where  $e_{\text{det}}$  is considered as the half-width of a rectangular distribution of the integration error that leads to standard uncertainty of  $e_{\text{det}}/\sqrt{3}$ , to which the standard uncertainty of its estimate  $u(e_{\text{det}})$  is added.

The contribution to the calculation uncertainty, which is related to the averaging of the random component, is estimated by statistical analysis. Without taking correlations into account, it is determined as:

$$u_{\text{ran}}^{(\text{uncor})} = \frac{s(q_{\text{ran},i})}{\sqrt{N}} \quad (4)$$

and with consideration of correlation effects as:

$$u_{\text{ran}}^{(\text{cor})} = \frac{s(q_{\text{ran},i})}{\sqrt{N}} \sqrt{1 + \frac{2 \sum_{k=1}^{N_{\text{cor}}} (N_{\text{ran}} - k) \rho(k)}{N_{\text{ran}}}} \quad (5)$$

where  $\rho(k)$  is the estimated  $k^{\text{th}}$  autocorrelation coefficient (Mathematica, built-in function CorrelationFunction),  $N_{\text{cor}}$  is the number of autocorrelation coefficients considered and  $N_{\text{ran}}$  is the number of samples of the separated random component. Using the FTZ method [13],  $N_{\text{cor}}$  is determined by identifying the smallest  $k$  for which  $\rho(k) > 0$  and  $\rho(k+1) < 0$ .

Combined calculation standard uncertainty, which considers the contributions related to deterministic and random components, is determined as:

$$u_{\text{cal}} = \sqrt{u_{\text{det}}^2 + u_{\text{ran}}^2} \quad (6)$$

### 3. Illustrative test case using experimental time-sampled data

The evaluation of the calculation uncertainty is carried out for the experimental time-sampled data of the superior calorific value  $H_{s,i}$  under reference conditions with the sampling period of  $t_{\text{samp}} = 15$  min (see data in Fig. 2 under label “Input”). The purpose is to evaluate the uncertainty associated with the calculation of an average value of these data. The average value is determined using the rectangle-rule numerical integration, written as follows for the observed data with the constant sampling period,  $t_{\text{samp},i} = t_{\text{samp}}$ ,  $i = 1 \dots N$ :

$$\bar{H}_s = \frac{\sum_{i=1}^N H_{s,i} t_{\text{samp},i}}{\sum_{i=1}^N t_{\text{samp},i}} = \frac{1}{N} \sum_{i=1}^N H_{s,i} = 40.8768 \text{ MJ/m}^3 \quad (7)$$

Fig. 2 also shows the deterministic component, which is separated from the input data using the Savitzky-Golay filter of the 2nd order with the

window length set to  $N_{\text{win}} = 5$ . The random component, which is determined as the difference between the input data and the deterministic component, is shown in Fig. 3. The deterministic component captures the changes in the superior calorific value due to deterministic changes in the properties of the gas metered. The random component describes the noise in the data set.

Fig. 4 shows the approximation of the average values of the decimated deterministic component of the superior calorific value with  $n_{\text{dec}}$  up to  $N_{\text{dec}} = 4$ . The given case leads to the following value for the standard uncertainty of calculation for the deterministic component,  $u_{\text{det}}$ :

$$e_{\text{det}} = \bar{H}_{s,\text{det}}(1) - \bar{H}_{s,\text{det,fit}}(0) = 0.19 \text{ kJ/m}^3 \quad (8)$$

$$u(e_{\text{det}}) = s(\bar{H}_{s,\text{det,fit}}(0)) = 0.035 \text{ kJ/m}^3 \quad (9)$$

$$u_{\text{det}} = \sqrt{\left(\frac{e_{\text{det}}}{\sqrt{3}}\right)^2 + u(e_{\text{det}})^2} = 0.12 \text{ kJ/m}^3 \quad (10)$$

Fig. 5 shows the ACF for the random component of the superior calorific value. It is important to note that the random component is not what is commonly understood as random. The given ACF plot displays a clear structure with negative correlation values regularly alternating

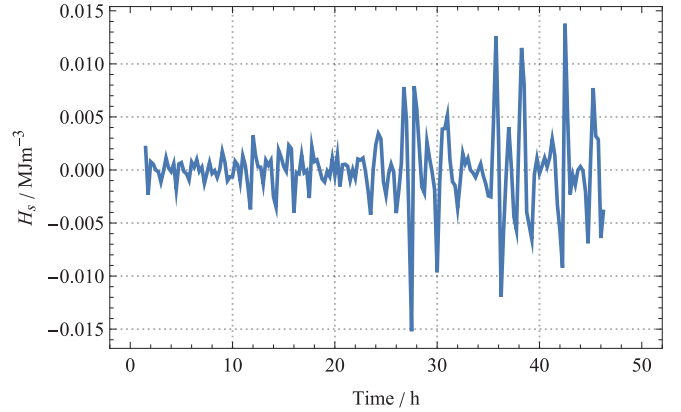


Fig. 3. Random component of the superior calorific value determined as the difference between the input data and the deterministic component ( $N_{\text{win}} = 5$ ).

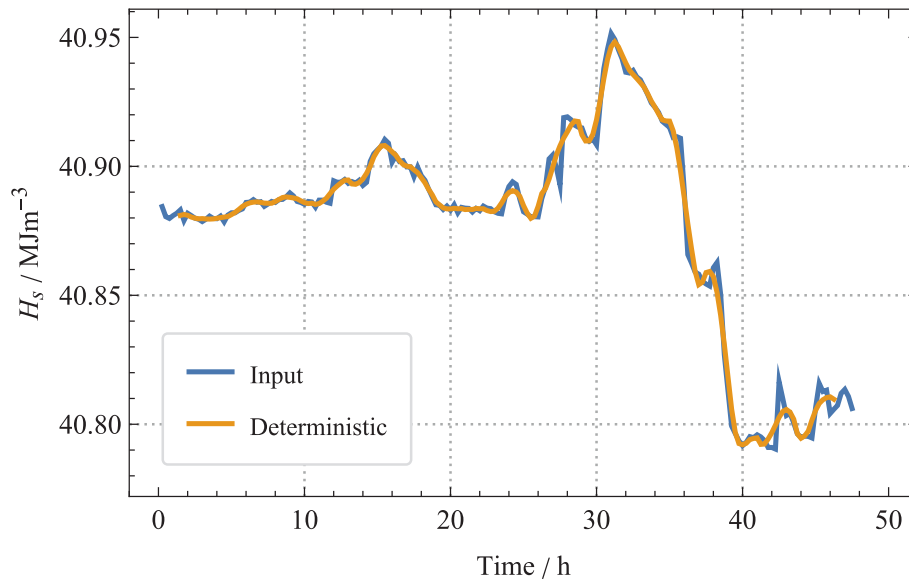


Fig. 2. Input data for the superior calorific value and its deterministic component determined by time-domain filtering ( $N_{\text{win}} = 5$ ).

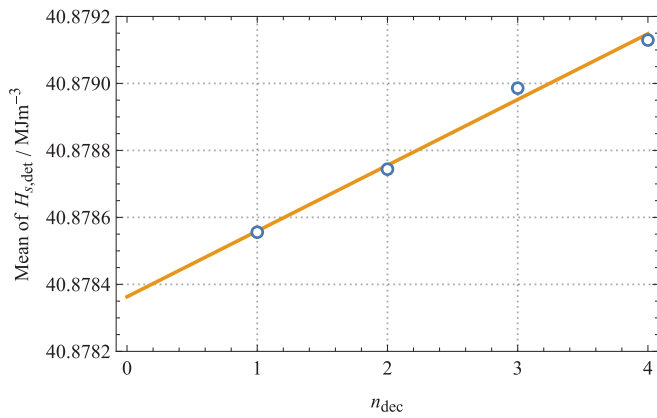


Fig. 4. Approximation of the average values of the decimated deterministic component ( $N_{win} = 5$ ,  $N_{dec} = 4$ ).

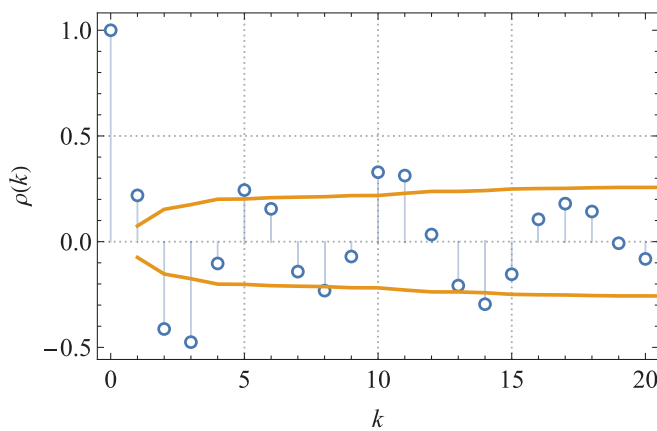


Fig. 5. Autocorrelation coefficients of the random component of the superior calorific value ( $N_{win} = 5$ ); leads to  $N_{cor} = 1$ .

with positive values. Part of the correlation is due to the filtering method, however there might be also other factors causing correlation. As a consequence, we consider correlation effects in the evaluation of the uncertainty. The given case leads to the following estimate of the standard uncertainty of calculation associated with the random component,  $u_{ran}$ , without and with consideration of the correlation between the samples:

$$u_{ran}^{(uncor)} = \frac{s(H_{s,ran,i})}{\sqrt{N}} = 0.25 \text{ kJ/m}^3 \quad (11)$$

$$u_{ran}^{(cor)} = \frac{s(H_{s,ran,i})}{\sqrt{N}} \sqrt{1 + \frac{2 \sum_{k=1}^{N_{cor}} (N_{ran} - k) \rho(k)}{N_{ran}}} = 0.30 \text{ kJ/m}^3. \quad (12)$$

By taking both contributions into account, the combined standard calculation uncertainty can be determined:

$$u_{cal}^{(uncor)} = \sqrt{u_{det}^2 + u_{ran}^{(uncor)2}} = 0.28 \text{ kJ/m}^3 \quad (13)$$

$$u_{cal}^{(cor)} = \sqrt{u_{det}^2 + u_{ran}^{(cor)2}} = 0.32 \text{ kJ/m}^3 \quad (14)$$

To show the influence of the window length of the S-G filter ( $N_{win}$ ) and the number of decimation steps considered ( $N_{dec}$ ) on the estimated measurement uncertainty, we performed calculations for  $N_{win} = (2 \dots 10)$  and  $N_{dec} = (3 \dots 10)$  (where, e.g.,  $N_{dec} = 3$  means that the uncertainty associated with the deterministic component is estimated by the OLS fit on the datapoints obtained setting  $n_{dec} = (1 \dots 3)$ ), and the corresponding results are presented in Figs. 6 and 7, respectively. The

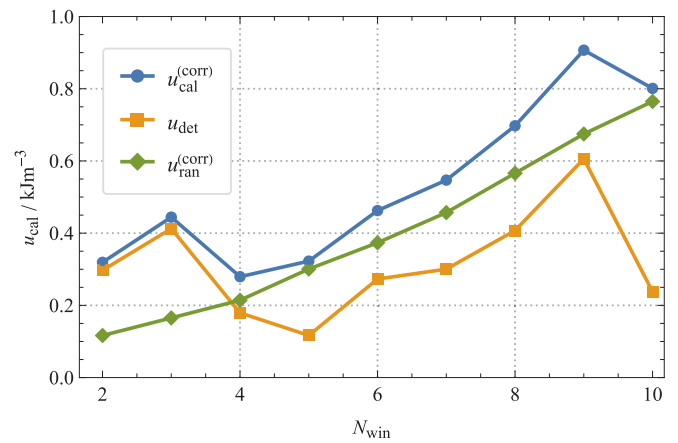


Fig. 6. Standard uncertainty of calculation for different values of the window length  $N_{win}$  ( $N_{dec} = 4$ , considering autocorrelation).

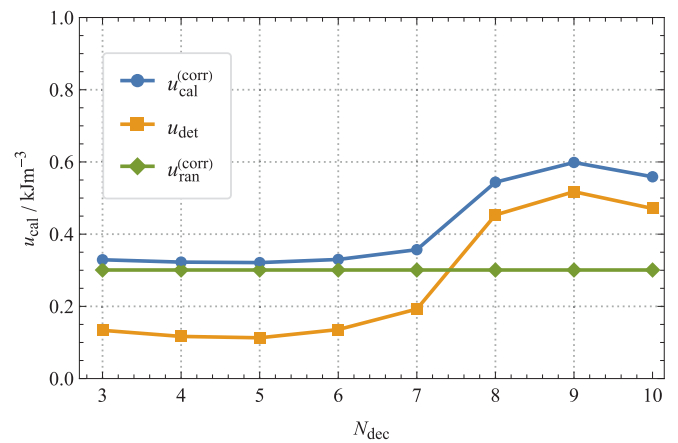


Fig. 7. Standard uncertainty of calculation for different values of the decimation range  $N_{dec}$  ( $N_{win} = 5$ , considering autocorrelation).

results are presented only for calculations considering the autocorrelations effects, as both trends and conclusions are not significantly different for the uncorrelated analysis.

Fig. 6 shows that with the increase of the smoothing window length  $N_{win}$ , the estimated value of the calculation uncertainty associated with the random component increases significantly. Such result can be expected, as the observed data also contain rapid variations (see Fig. 2 between 30 and 40 h) and therefore some of the deterministic variations are included in the estimated random component for a wider smoothing window. As further confirmed by the analysis of artificial data sets in the next section, more realistic estimates of the calculation uncertainty for the studied case are at lower  $N_{win}$  values.

Fig. 7 shows that with the increase of the decimation range  $N_{dec}$ , the estimated value of the calculation uncertainty associated with the deterministic component also shows a certain increasing trend. The analysis of artificial data sets in the next section shows that the estimates of the calculation uncertainty at lower  $N_{dec}$  values are more appropriate.

To conclude this analysis, let us look at another case that shows the advantages of the proposed procedure in terms of more realistic estimates of the calculation uncertainty, because the effects of the deterministic and random components are treated separately. Let us assume that someone (inappropriately) decided to evaluate the standard uncertainty under discussion as the standard deviation of the mean of the whole input data, i.e., without separate evaluation of the deterministic and random components:

$$u_A = \frac{s(H_{s,i})}{\sqrt{N}} = 2.94 \text{ kJ/m}^3. \quad (15)$$

Such an uncertainty estimate is more than 3 times larger than all the results presented in Figs. 6 and 7, and when considering the more optimal values of  $N_{\text{win}}$  and  $N_{\text{dec}}$ , which are in the left half of these graphs, this ratio increases up to 6 to 8. The analysis of artificial data sets in the next section confirms that the lower uncertainty obtained by the proposed procedure is a realistic estimate of the calculation uncertainty.

#### 4. Procedure verification using artificial time-sampled data

In this section, the proposed procedure for evaluating the calculation uncertainty is applied to selected artificial input data containing both deterministic and random variations. Since the defined artificial signals have a known true value of the calculated quantity, it is possible to assess the validity of the prediction of the calculation uncertainty and define guidelines for the optimal selection of the procedure parameters. For the study in this paper, we chose the deterministic component as an asymmetric sinusoidal signal (absolute value of sine function) and the random component as normally distributed data. The chosen form of artificially generated time-sampled input data set  $q_i$ ,  $i = 1 \dots N$ , is expressed as:

$$q_i = \bar{q} + A_q \left| \sin(2\pi f_q t_{\text{samp}} i + \phi_q) \right| + \mathcal{N}(0, \sigma_q), \quad (16)$$

where  $\bar{q}$ ,  $A_q$ ,  $f_q$  and  $\phi_q$  are the mean value, the amplitude, the frequency and the initial phase of the deterministic component, respectively, and  $\mathcal{N}(0, \sigma_q)$  is the random normally distributed data set with the mean of zero and the standard deviation of  $\sigma_q$ . For each generated data set the initial phase is determined as a random value between zero and  $\pi$ ,  $\phi_q = \pi \text{rand}(1)$ . Table 1 presents the values of the input parameters for Eq. (16) used in simulations in this paper. An example of a data set generated in this manner is shown in Fig. 8.

The error in determining the mean value, which is associated with the calculation uncertainty, is:

$$e = Q - Q_{\text{ref}} \quad (17)$$

where  $Q$  is the mean value estimated using the rectangle-rule numerical integration:

$$Q = \frac{1}{N} \sum_{i=1}^N q_i \quad (18)$$

and  $Q_{\text{ref}}$  is the reference mean value determined by integration of the deterministic component:

$$Q_{\text{ref}} = \bar{q} + A_q \frac{1}{t_{\text{signal}}} \int_0^{t_{\text{signal}}} \left| \sin(2\pi f_q t + \phi_q) \right| dt \quad (19)$$

Eq. (19) is also analytically solvable by separating regions where  $\sin()$  is positive or negative, however, we used high-precision numerical integration (Mathematica, built-in function NIntegrate) to estimate this

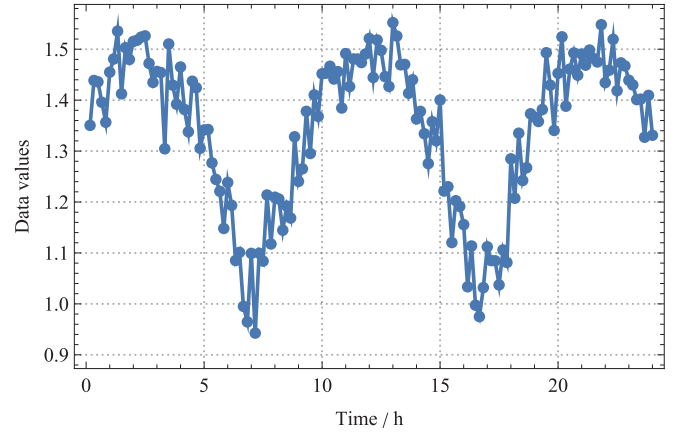


Fig. 8. An example of the generated artificial time-sampled data set based on Eq. (16) and Table 1 ( $A_q = 0.5$ ,  $f_q = 1.25/T_{\text{signal}}$ ).

reference value.

For further comparisons between the actual calculation errors  $e$  determined by Eq. (17) and the range of the estimated calculation uncertainty, the standard uncertainty determined by Eq. (6) is converted into the expanded uncertainty with coverage probability of 95.45 % (coverage factor of  $k = 2$ ):

$$U_{\text{cal}} = 2u_{\text{cal}} \quad (20)$$

Figs. 9 and 10 present the comparison results for the actual calculation errors and the estimated calculation expanded uncertainties for different values of the window length  $N_{\text{win}}$  and the decimation range  $N_{\text{dec}}$ , respectively. The analysis is performed for 3 types of artificial time-sampled data sets: (a) considering the nominal amplitude and frequency of the deterministic component, (b) for the tripled frequency of the deterministic component, and (c) for the tripled amplitude of the deterministic component. For each selected signal type and set of the procedure parameters, 10 simulations are performed, thus showing the effect of changes in random signal part.

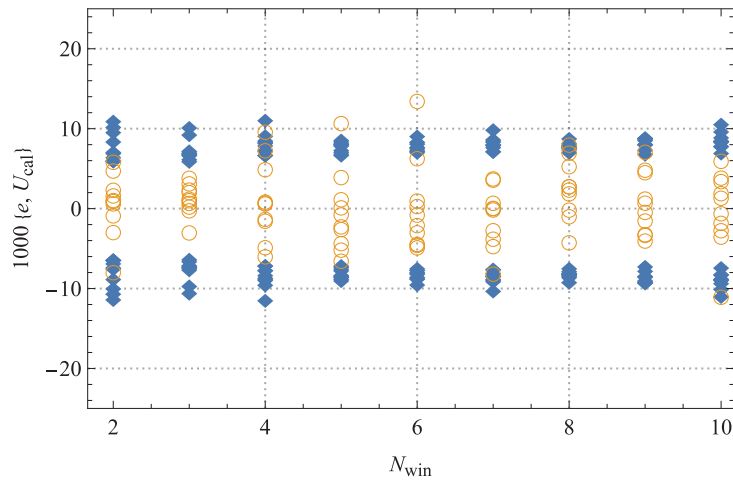
Based on the presented results in Figs. 9 and 10, it can be concluded that the proposed procedure for estimating the calculation uncertainty works reliably and robustly. Regardless of the large range in the different forms of the analysed signals and varying procedure parameters, the calculation errors are within the estimated expanded uncertainties in a larger proportion than the 95 % coverage probability. There are certain ranges of the procedure parameters where the estimated uncertainties are noticeably larger than the error scatter. The analysis of the signal with the higher-frequency deterministic component in Fig. 9(b) shows a systematic shift in the estimated uncertainties with increasing the smoothing window length  $N_{\text{win}}$ . Too large window lengths  $N_{\text{win}}$  cause over-filtering, which results in an excessively larger estimated random component and associated measurement uncertainty. There is therefore a certain limit value of the smoothing window length, which depends on the frequency of deterministic changes in the analysed signals. The results in Fig. 9(b) show that this limit is approximately at  $N_{\text{win}} = 5$  for the signal with 19 samples in one half-period of the deterministic component, which means that their ratio can be at most approximately 1:4.

Certain ranges of the procedure parameters show increased random scatter of the estimated calculation uncertainty. This effect is characteristic for too small values of the smoothing window length  $N_{\text{win}}$ , especially for the analysed signals with a higher-amplitude deterministic component (Fig. 9(c)). If the value of  $N_{\text{win}}$  is too small, the deterministic component removed after filtering still retains a significant portion of the random signal, which causes a larger scatter of the uncertainty estimated by the decimation method. An increase of the decimation range  $N_{\text{dec}}$  does not improve the quality of uncertainty estimate, as

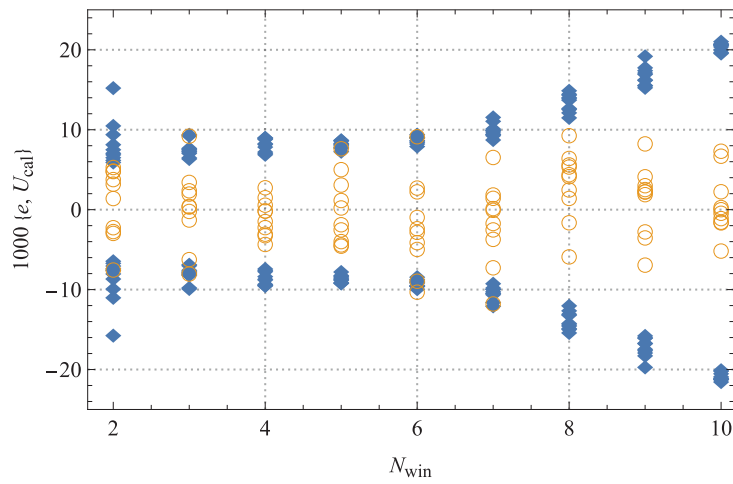
Table 1

Input parameters for generating the artificial time-sampled data sets based on Eq. (16).

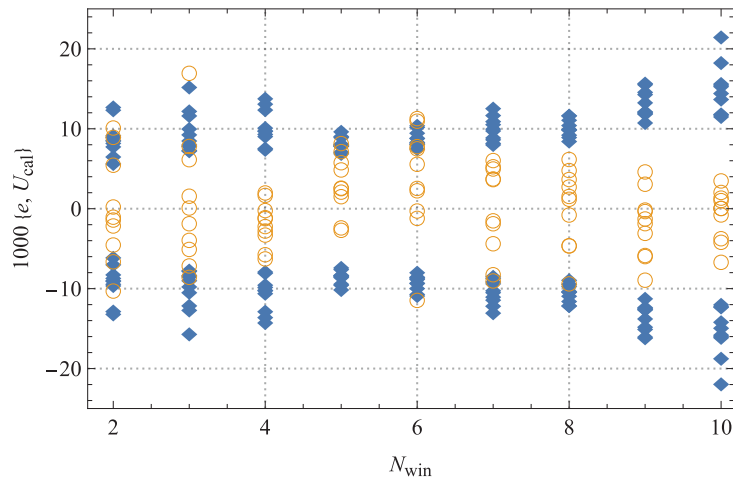
$t_{\text{samp}}$	10 min
$N$	144
$T_{\text{signal}} = N t_{\text{samp}}$	24 h
$\bar{q}$	1
$A_q$	0.5; 1.5
$f_q$	$1.25/T_{\text{signal}}$ ; $3.75/T_{\text{signal}}$
$\phi_q$	$\pi \text{rand}(1)$
$\sigma_q$	0.05



(a) For nominal parameters of the deterministic component ( $A_q = 0.5, f_q = 1.25/T_{\text{signal}}$ ).



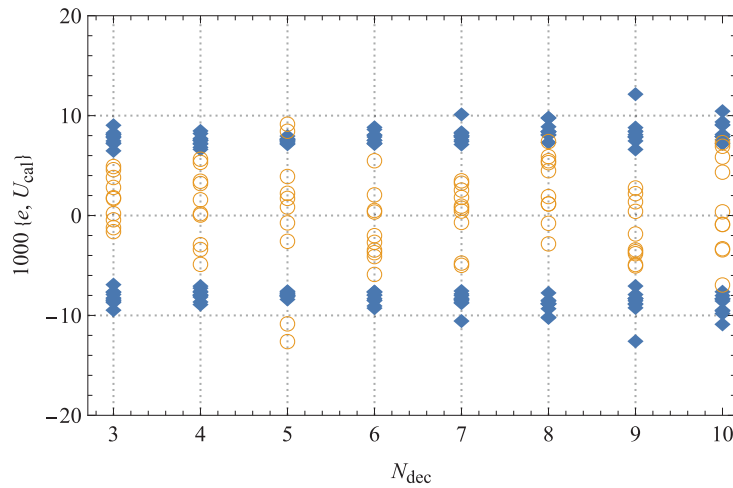
(b) For the increased frequency of the deterministic component ( $A_q = 0.5, f_q = 3.75/T_{\text{signal}}$ ).



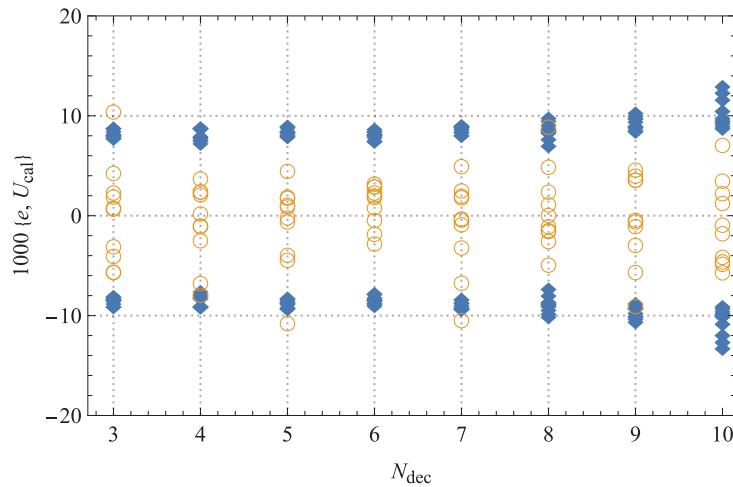
(c) For the increased amplitude of the deterministic component ( $A_q = 1.5, f_q = 1.25/T_{\text{signal}}$ ).

**Fig. 9.** Errors and predicted expanded uncertainties of calculation of the mean value of the artificial time-sampled data sets for different values of the window length  $N_{\text{win}}$  ( $N_{\text{dec}} = 4$ , considering autocorrelation;  $\circ$  – error,  $\blacklozenge$  – expanded uncertainty).

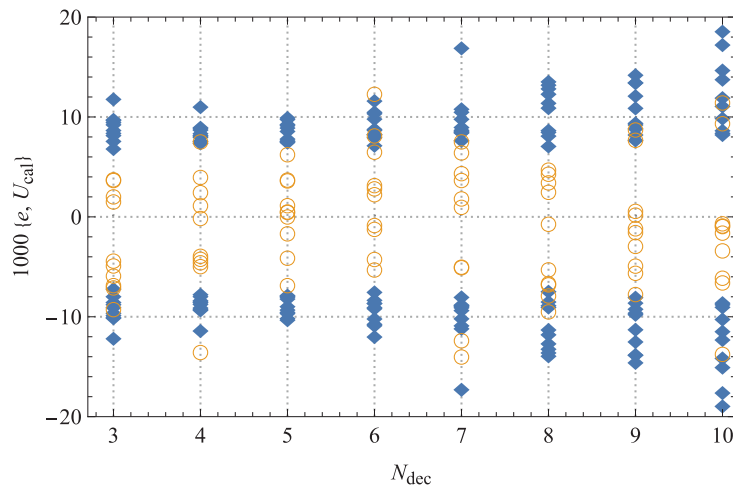




(a) For nominal parameters of the deterministic component ( $A_q = 0.5, f_q = 1.25/T_{\text{signal}}$ ).



(b) For the increased frequency of the deterministic component ( $A_q = 0.5, f_q = 3.75/T_{\text{signal}}$ ).



(c) For the increased amplitude of the deterministic component ( $A_q = 1.5, f_q = 1.25/T_{\text{signal}}$ ).

**Fig. 10.** Errors and predicted expanded uncertainties of calculation of the mean value of the artificial time-sampled data sets for different values of the decimation range  $N_{\text{dec}}$  ( $N_{\text{win}} = 5$ , considering autocorrelation;  $\circ$  – error,  $\blacklozenge$  – expanded uncertainty).

shown in Fig. 10. Too large  $N_{\text{dec}}$  can even lead to an increased scatter of the estimated uncertainty, as is clearly seen in Fig. 10(c) for the signals with a higher-amplitude deterministic component; the results show that the optimal choice of the decimation range is up to  $N_{\text{dec}} = 5$ .

## 5. Conclusions

In this paper, we presented and verified a procedure for evaluating the uncertainty associated with calculations of average/total values

from time-sampled data in gas energy measurements. The main findings of the study can be summarized as follows:

- The developed procedure enables separate treatment of contributions to the calculation uncertainty that are related to deterministic and random variations of the observed data.
- Demonstration of this procedure on the experimental data from actual gas energy measurements and on artificially generated signals confirms its strong potential for reliable and robust uncertainty estimation.
- The results show a certain sensitivity to the procedure parameters, i. e., the window length  $N_{\text{win}}$  and the decimation range  $N_{\text{dec}}$ , but the estimated uncertainty in the case of non-optimal selection of these parameters is mostly on the safe side, with in a bit higher value.
- The choice of the smoothing window length  $N_{\text{win}}$  can be optimized according to the estimated dynamics of deterministic changes of the observed data. The results show that the total width of the smoothing window of the time-domain filter should not exceed one quarter of the time period of data variations.
- The optimal choice of the decimation range  $N_{\text{dec}}$ , which defines the number of decimation steps in estimating the reference value in averaging of the deterministic component, has been shown to lie between three and five.

While the findings of this study are promising for applicability of the developed procedure, future research should focus on additional verifications and validations through analyses of different experimental and artificial data sets. A challenge for future development could be an automated, adaptive adjustment of the procedure parameters according to the changing properties of the observed data. Another potential challenge relates to deciding in which cases it is necessary to take into account the contributions from both the random and deterministic components, as presented in this paper. For example, in the final totalisation of the gas quantity readings that are previously internally integrated in the flow meter, it makes sense to consider only the random-component uncertainty. A certain open question remains regarding how to properly combine different components of the calculation uncertainty for, e.g., the gas energy, which is calculated as the product of time-sampled enthalpy values and internally integrated gas quantity readings.

#### CRedit authorship contribution statement

**Joze Kutin:** Writing – original draft, Validation, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Gregor Bobovnik:** Writing – original draft, Validation, Software, Methodology, Funding acquisition, Formal analysis. **Federica Gugole:** Writing – review & editing, Validation. **Adriaan M.H. van der Veen:** Writing – review & editing, Validation, Funding acquisition.

#### Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This work was supported by the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States (joint research project No. 21GRD05, Met4H2). This work was also supported by the Slovenian Research and Innovation Agency (research core funding No. P2-0223) and by the Ministry of Economic Affairs of the Netherlands.

#### Data availability

Data will be made available on request.

#### References

- [1] ISO 15112:2018, Natural Gas – Energy Determination, International Organization for Standardization (ISO), 2018.
- [2] OIML R 140:2007, Measuring Systems for Gaseous Fuel, International Organization of Legal Metrology (OIML), 2007.
- [3] EN 1776:2015, Gas Infrastructure – Gas Measuring Systems – Functional Requirements, European Standard (EN), 2015.
- [4] G. Ficco, M. Dell'Isola, P. Vigo, L. Celenza, Uncertainty analysis of energy measurements in natural gas transmission networks, *Flow Meas. Instrum.* 42 (2015) 58–68.
- [5] A. S. Ribeiro, D. Loureiro, M. C. Almeida, M. G. Cox, J. A. Sousa, M. A. Silva, L. Martins, R. Brito, A. C. Soares, Uncertainty evaluation of totalization of flow and volume measurements in drinking water supply networks, *FLOMEKO 18<sup>th</sup> Int. Flow Measurement Conference* (June 26–28, 2019, Lisbon, Portugal), 2019.
- [6] G. Ficco, L. Canale, G. Cortellessa, F. Zuena, M. Dell'Isola, Effect of flow-rate measurement accuracy on unaccounted for gas in transmission networks, *Flow Meas. Instrum.* 90 (2023) 102336.
- [7] P.G. Spazzini, W. Bich, Uncertainty in discrete-time integration – the case of static gas meters, *Measurement* 224 (2024) 113821.
- [8] A.M.H. van der Veen, K. Folgerø, F. Gugole, Measurement uncertainty in the totalisation of quantity and energy measurement in gas grids, *Gases* 5 (2025) 7.
- [9] JCGM 100:2008, Evaluation of Measurement Data – Guide to the Expression of uncertainty in Measurement, Joint Committee for Guides in Metrology (JCGM), 2008.
- [10] P.J. Davis, P. Rabinowitz, *Methods of numerical integration*, Second edition, Academic Press, 1984.
- [11] B. Boashash, *Time-Frequency Signal Analysis and Processing – a Comprehensive Reference*, Academic Press, 2015.
- [12] M. Parker, *Digital signal processing 101 – everything you need to know to get started*, Elsevier Science, 2017.
- [13] A. Zięba, P. Ramza, Standard deviation of the mean of autocorrelated observations estimated with the use of the autocorrelation function estimated from the data, *Metrol. Measur. Syst.* 18 (2011) 529–542.
- [14] N.F. Zhang, Calculation of the uncertainty of the mean of autocorrelated measurements, *Metrologia* 43 (2006) S276–S281.
- [15] F. Gugole, M. Li, A.M.H. van der Veen, On the autocorrelation of measurement results for gas volume and calorific value in fiscal metering in gas grids, *EPJ Web of Conferences* 323 (2025) 09003.
- [16] JCGM GUM-6:2020, Guide to the expression of uncertainty in measurement – Part 6: developing and using measurement models, Joint Committee for Guides in Metrology (JCGM), 2020.
- [17] J. Kutin, G. Bobovnik, F. Gugole, A.M.H. van der Veen, Evaluation of uncertainty associated with totalisation of time-sampled data in gas quantity and energy measurements, *Measur. Sens.* (2024) 101572.