



LINEAR REGRESSION ANALYSIS AND THE GUM: EXAMPLE OF TEMPERATURE INFLUENCE ON FORCE TRANSFER TRANSDUCERS

D. Röske¹

¹ Physikalisch Technische Bundesanstalt (PTB), Braunschweig, Germany, dirk.roeske@ptb.de

Abstract:

The deflection of strain-gauge force and torque transducers (the zero-reduced output signal for a given mechanical load) is dependent on the ambient temperature. This is also true of high-precision force transfer transducers used to compare force standard machines. To estimate the extent to which temperature deviations during measurements on different machines affect comparison results and – if necessary – to correct for such deviations, it is important to know the influence of the temperature on the deflection. This effect is usually investigated in special temperature chambers in which the transducer is exposed to various temperatures within a given temperature range while the machine is operated under unchanged laboratory conditions. The regression analysis of the results allows the temperature coefficient to be determined, including an uncertainty analysis. This was done for five force transfer transducers used for a bilateral comparison between NPL (UK) and PTB (Germany).

Keywords: comparison; temperature influence; regression; uncertainty

1. INTRODUCTION

The manufacturers of strain-gauge force and torque transducers devote significant effort to minimising the influence of the ambient temperature on the measurement results of their devices. The best transducers from this group are very well compensated for temperature influences. However, no measurements of the transducer's temperature behaviour are taken and only the upper limit of the absolute value of the sensitivity's temperature coefficient (the relative change of the deflection with temperature) is given in the transducer's specifications. For example, a range between $-0.02 \text{ \%}/(10 \text{ K})$ and $+0.02 \text{ \%}/(10 \text{ K})$ is given for the HBM Z30A transducer [1], while a range between $-0.01 \text{ \%}/(10 \text{ K})$ and $+0.01 \text{ \%}/(10 \text{ K})$ is given for the GTM KTN transducer [2]. This means that the absolute value of the temperature coefficient should be below $1 \dots 2 \cdot 10^{-5} \text{ K}^{-1}$. For a 1 K temperature

change, this is more than the standard uncertainty of the best force standard machines.

2. MEASUREMENT RESULTS

For a comparison measurement between PTB's new 200 kN force standard machine [3] and the relevant standard machines of the NPL, five high-precision compression force transfer transducers (one Z30A, four KTNs) were selected. These transducers were investigated with respect to their temperature behaviour in a temperature chamber inside the 200 kN force standard machine.

The temperature range was defined as $20 \text{ °C} \dots 25 \text{ °C}$ with additional measurement points beyond these threshold values at 21 °C , 22 °C and 23 °C . The resulting deflections d_i in mV/V for the 50 kN KTN transducer are given in Table 1 for the five temperatures T_i and for two load steps (20 kN and 50 kN).

Table 1: Deflections d_i in mV/V for the 50 kN transducer at different temperatures T_i and for two load steps

T_i in °C	d_i at 20 kN in mV/V	d_i at 50 kN in mV/V
20.03	0.800 996	2.003 036
21.00	0.800 992	2.003 028
21.99	0.800 988	2.003 011
23.06	0.800 984	2.003 002
24.90	0.800 978	2.002 981

3. PRELIMINARY EVALUATION

In the evaluation, a linear dependency between the temperature T and the deflection d is assumed to exist within a sufficiently small temperature range (and possibly beyond this range).

In the first approach, the data can be evaluated using the linear fit functions of programming languages such as Python and R or standard software programs such as Excel, all of which are based on the least-squares method and require a very small number of code lines, see Figure 1. Excel even contains a function wizard that can find the right function and defines the necessary arguments. The results agree well at a relative level of 10^{-12} and below. The limitation is caused by the floating-point

accuracy of the processor, operating system and software used; however, this is not a problem for the uncertainty in question in this investigation.

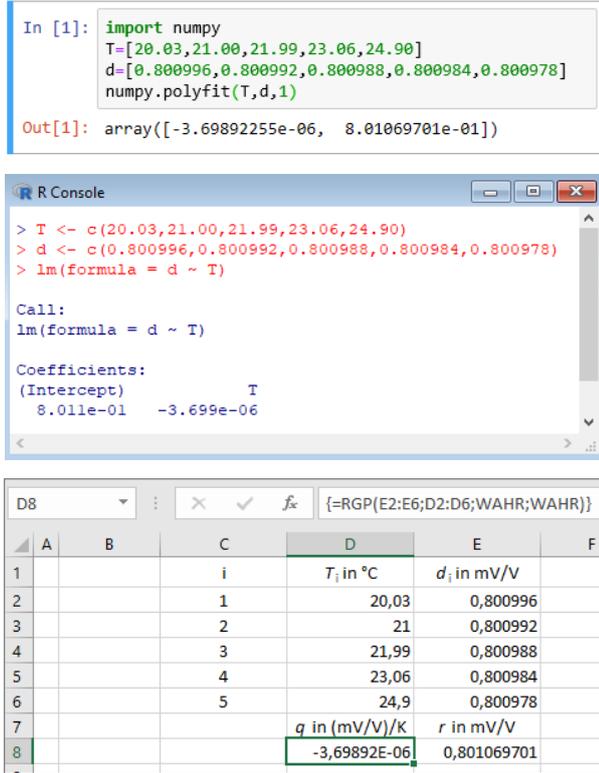


Figure 1: Parameters (for the 20 kN force in Table 1) of a linear fit in Python (top), R (middle) and Excel (bottom, German user interface)

Most linear-fit packages and functions contain additional information about the result such as the residual sum of squares (RSS) of the least-squares fit in Python and Excel and the average variation of points around the fitted regression line, the residual standard error (RSE), in R. These values can be used to check how well the model fits the data; this step is necessary to ensure that the right model is used for the data.

The residual sum of squares of the least-squares fit in Excel yields a value of $6.07 \cdot 10^{-13} \text{ (mV/V)}^2$ for the data in Table 1 (for 20 kN). To improve comparability, this value is linked to another value – namely, the sum of squared deviations (SSD) from the mean. The result is $1.95 \cdot 10^{-10} \text{ (mV/V)}^2$ and the relative deviation $(SSD - RSS) / SSD = 0.997$ is a measure of the model quality: the closer this value is to 1, the better the model is.

Nevertheless, this calculation does not consider the uncertainty of the input data. It is expected that the resulting parameters will also have an associated uncertainty and that this value will depend on the uncertainty of the input data.

To obtain the uncertainty of the results of the fitting procedure, the following regression analysis was carried out using a linear regression model and a least-squares approach in combination with the

GUM, the Guide to the expression of uncertainty in measurement [4].

4. REGRESSION ANALYSIS

We consider five pairs of input (T_i) and output (d_i) values with their associated uncertainties (u) $((T_i; u(T_i)), (d_i; u(d_i)))$ $i = 1, 2, \dots 5$.

Here, the aim is to find a linear approximation function $\tilde{d}(T)$ that best describes the data

$$\tilde{d}(T) = q \cdot T + r \quad (2)$$

with the coefficients q (slope) and r (intercept). The values of these coefficients should be determined in such a way that a minimum condition is met. In the context of the model considered here, we require the sum of the squared deviations between the measured outputs and the corresponding outputs calculated in accordance with (2) to be a minimum (“least squares”)

$$\sum_{i=1}^N (d_i - \tilde{d}(T_i))^2 \rightarrow \min . \quad (3)$$

With (2), equation (3) can be also written as

$$\sum_{i=1}^N (d_i - q \cdot T_i - r)^2 \rightarrow \min . \quad (4)$$

The condition necessary for the minimum is that the partial derivatives of the given term to the unknown values q and r be zero. This yields

$$2 \sum_{i=1}^N (d_i - q \cdot T_i - r)(-T_i) = 0 \quad (5)$$

$$-2 \sum_{i=1}^N (d_i - q \cdot T_i - r) = 0 ,$$

a system of two equations via which the values of the coefficients can be determined. A more compact representation of (5) is

$$q \cdot \overline{T^2} + r \cdot \overline{T} = \overline{T \cdot d} \quad (6)$$

$$q \cdot \overline{T} + r = \overline{d}$$

with

$$\overline{T^k} = \frac{1}{N} \cdot \sum_{i=1}^N T_i^k , \quad \overline{d} = \frac{1}{N} \cdot \sum_{i=1}^N d_i \quad (7)$$

and

$$\overline{T \cdot d} = \frac{1}{N} \cdot \sum_{i=1}^N T_i \cdot d_i . \quad (8)$$

The condition sufficient for the minimum value is that the second derivative be positive. Equation (4) yields

$$\frac{\partial^2}{\partial q^2} \sum_{i=1}^N (d_i - q \cdot T_i + r)^2 = 2 \sum_{i=1}^N T_i^2 > 0 , \quad (9)$$

$$\frac{\partial^2}{\partial r^2} \sum_{i=1}^N (d_i - q \cdot T_i + r)^2 = 2N > 0$$

showing that the condition sufficient for the minimum value is fulfilled independently of the parameter values.

From (6), the coefficients q and r can be calculated. They are functions of the \bar{T}^k ($k \in (1, 2)$) and $\bar{T} \cdot \bar{d}$ and, due to (7) and (8), functions of the T_i and d_i . This means that their uncertainties can, in principle, be calculated by applying the standard methods of the GUM.

The solution of (6) can be written as

$$q = \frac{\bar{T} \cdot \bar{d} - \bar{d} \cdot \bar{T}}{\bar{T}^2 - \bar{T}^2}, \quad r = \bar{d} - q \cdot \bar{T}. \quad (10)$$

It must be noted that the slope of the linear function is the same over the whole temperature range, whereas the intercept depends on the value taken for zero. If, instead of the °C scale used in Table 1, the Kelvin scale is used, the value of the intercept calculated will change, see Figure 2.

A	B	C	D	E	F
1	i	T_i in °C	T_i in K	d_i in mV/V	
2	1	20,03	293,18	0,800996	
3	2	21	294,15	0,800992	
4	3	21,99	295,14	0,800988	
5	4	23,06	296,21	0,800984	
6	5	24,9	298,05	0,800978	
7			q in (mV/V)/K	r in mV/V	
8			-3,69892E-06	0,802080062	

Figure 2: Parameters (for the 20 kN force in Table 1) of a linear fit in Excel with Kelvin temperatures

Force transfer standards are usually stored and operated (and, if possible, transported) in a narrow temperature range from 18 °C to 28 °C (for key comparison measurements, a much smaller interval such as 20 °C ± 0.5 K may be required). This means that the behaviour of the force transducer near 0 °C is not of interest; it is therefore not important that it be known for temperatures close to 0 K. It should be sufficient to describe the transducer's behaviour in the narrow temperature interval where it was investigated.

Usually, the reference temperature T_{ref} for comparison measurements is agreed in advance when the technical protocol is compiled. Then, the calculation can be carried out with the new temperatures T'_i defined as $T'_i = T_i - T_{\text{ref}}$. Then, (10) will be written as

$$q = \frac{\bar{T}' \cdot \bar{d} - \bar{d} \cdot \bar{T}'}{\bar{T}'^2 - \bar{T}'^2}, \quad r = \bar{d} - q \cdot \bar{T}'. \quad (11)$$

In a special case in which the reference temperature T_{ref} equals the mean temperature \bar{T} , the mean of the new temperatures becomes zero:

$$\bar{T}' = \frac{1}{N} \cdot \sum_{i=1}^N T_i - \bar{T} = \bar{T} - \bar{T} = 0. \quad (12)$$

In this case, (11) is simplified to

$$q = \frac{\bar{T}' \cdot \bar{d}}{\bar{T}'^2}, \quad r = \bar{d}. \quad (13)$$

It must be noted that the single temperature points may change in different measurement campaigns; for example, a new value of 24 °C may be defined instead of or in addition to 23 °C. For better comparability of results, it could be beneficial to define the number of temperature points for all such investigations in advance. The single points may then be chosen in such a way that their mean equals the reference temperature agreed. Although it may be difficult to reproduce all single temperatures very accurately, the remaining deviations of 0.1 ... 0.2 K should be small enough to be neglected.

With (7) and (8), equations (13) can now be re-written as

$$q = \sum_{i=1}^N T'_i \cdot d_i / \sum_{i=1}^N T'^2_i, \quad r = \frac{1}{N} \cdot \sum_{i=1}^N d_i. \quad (14)$$

Depending on the reference temperature chosen, (11), respectively (14), are the model functions to which the GUM should be applied to find the uncertainties $u(q)$ and $u(r)$. Here, the aim is to find the regression function $\bar{d}(T)$ and the uncertainties of the fitted values, preferably also given by a function $u(\bar{d})$.

In this work, the more common approach (11) was used because the reference temperature T_{ref} did not match the mean of the temperatures T_i .

5. RESULTS

The calculations were carried out under the assumption that no correlations existed between the input values T'_i, d_i , which were treated as uncorrelated quantities. The equations (11) for the determination of q and r are quite simple, whereas those for the determination of $u(q)$ and $u(r)$ are rather complex. Following the GUM procedure, partial derivatives to each of the ten input variables must be calculated. Maxima, a computer algebra system [5], was used to carry out this calculation analytically, yielding terms with several hundred parts as the result for the uncertainties. Although this software could have been used to calculate the complete result $q, u(q), r, u(r)$ for all the different transducers, this would have been time-consuming. A better way was to obtain the analytical result in Maxima and to use this formula in Excel. The known formula was then applied as a spreadsheet function in all subsequent calculations with the same input data scheme (five pairs of values as in Table 1).

The results show that, if a temperature outside the interval of interest is taken as zero, the uncertainty of the intercept r will be larger than it would be if a value from the given interval were taken as a reference temperature T_{ref} , see Table 2. Moreover, the uncertainty of the intercept is minimal if the reference temperature chosen is the arithmetic mean of the temperature values T_i .

Table 2: Values and associated standard uncertainties of the slope and intercept of a linear fit calculated for different reference temperatures

T_{ref} in °C	$q, u(q)$ in (mV/V)/K	$r, u(r)$ in mV/V
-273.15	$-3.699 \cdot 10^{-6}$, $0.798 \cdot 10^{-6}$	0.802 080, 0.000 236
0.00	$-3.699 \cdot 10^{-6}$, $0.798 \cdot 10^{-6}$	0.801 070, 0.000 017 8
20.50	$-3.699 \cdot 10^{-6}$, $0.798 \cdot 10^{-6}$	0.800 994, 0.000 001 9
22.20	$-3.699 \cdot 10^{-6}$, $0.798 \cdot 10^{-6}$	0.800 988, 0.000 001 4
24.00	$-3.699 \cdot 10^{-6}$, $0.798 \cdot 10^{-6}$	0.800 981, 0.000 002 0

The uncertainty of the slope q is not affected by the reference temperature selected; due to the linear function (1), its contribution to the uncertainty of the fitted value $\tilde{d}(T'_i)$ is calculated with the temperature T'_i as a sensitivity coefficient. This means that the uncertainty contribution of the slope will be lower the closer the reference temperature T_{ref} and the temperatures T_i are to each other.

By means of the known coefficients, the regression function $\tilde{d}(T)$ can be calculated. The known standard uncertainties of the coefficients also allow other functions to be determined – namely, functions defining a $1\text{-}\sigma$ band along the approximation function. Expanded uncertainties with $k = 2$ yield a $2\text{-}\sigma$ band. Figure 3 and Figure 4 show these results for the given 50 kN transducer at the 20 kN load step with the measured values (blue symbols) and with standard uncertainty bars for T and d , the fit function (red line) and the $2\text{-}\sigma$ bands (dashed red lines). The standard uncertainty of the temperature was calculated from a rectangular distribution with a half-width of 0.1 K; the standard uncertainty of the deflection was 0.000 003 mV/V. The reference temperature was 20.5 °C.

Figure 5 shows how the results change if another reference temperature is chosen. If the reference temperature T_{ref} is the arithmetic mean of the temperatures T_i at which the investigation is carried out, the uncertainties become lower and their minimum value is at the reference temperature.

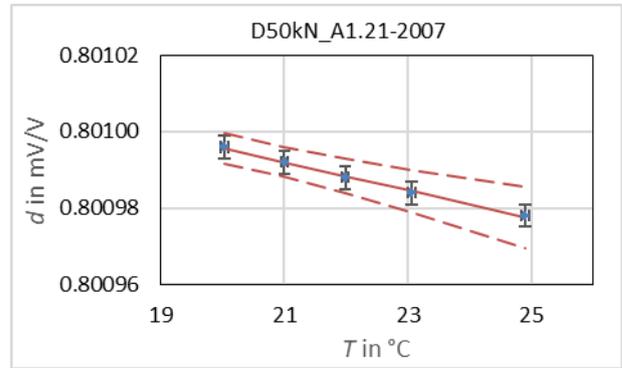


Figure 3: Result of the regression analysis for the 20 kN force step of the 50 kN transducer (details see text above)

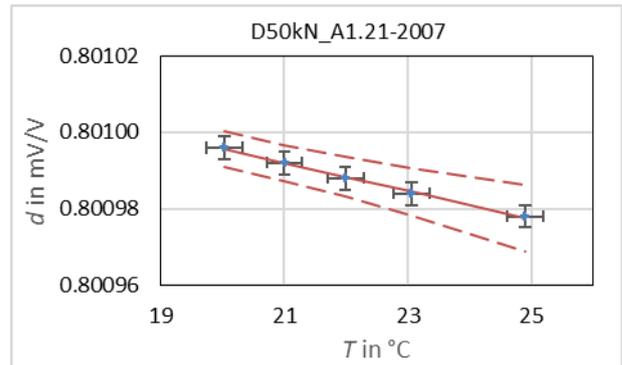
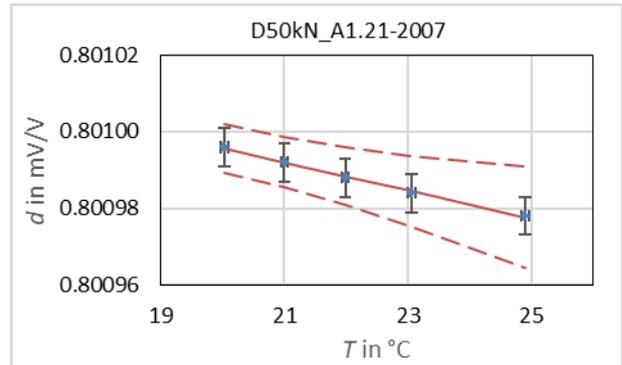


Figure 4: Result of Figure 3 when the uncertainty of the deflection is increased to 0.000 005 mV/V (top) and when the half-width of the temperature distribution is increased to 0.5 K (bottom)

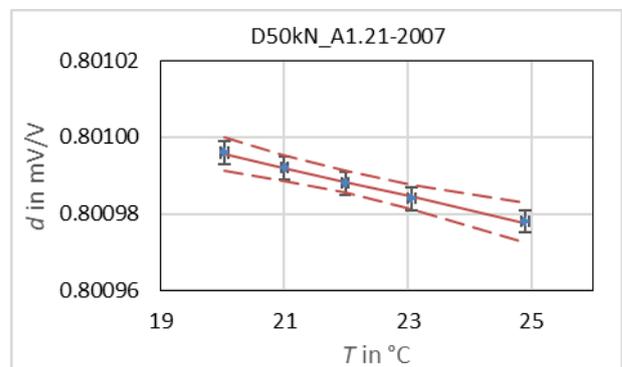


Figure 5: Result of Figure 3 when the arithmetic mean of the temperature values T_i (22.196 K) is taken as the reference temperature T_{ref}

The results and figures shown above are an example of a transducer with a very linear deflection dependency on the temperature; the relative deviation between RSS and SSD is 0.997. In addition, the slope value of $-3.7 \cdot 10^{-6}$ (mV/V)/K at a 0.8 mV/V signal is very low and a good value for a transfer transducer. The results for the other transducers are shown as examples in Figure 6 to Figure 8.

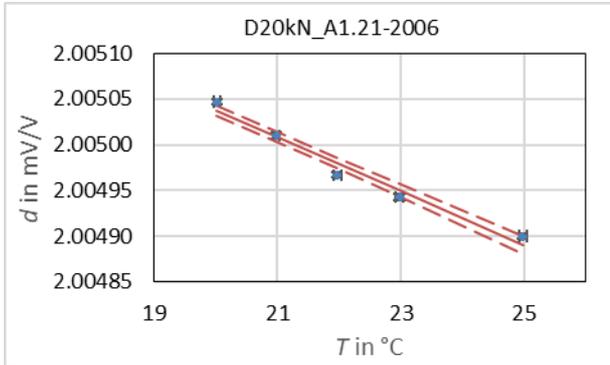


Figure 6: Result of the 20 kN transducer at a force of 20 kN and a reference temperature of 20.5 °C (further details as in Figure 3)

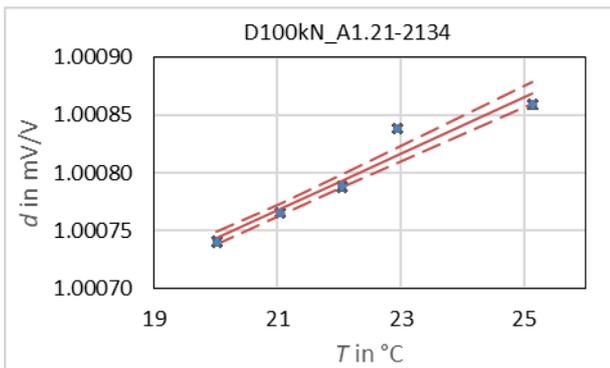


Figure 7: Result of the 100 kN transducer at a force of 50 kN and a reference temperature of 20.5 °C (further details as in Figure 3)

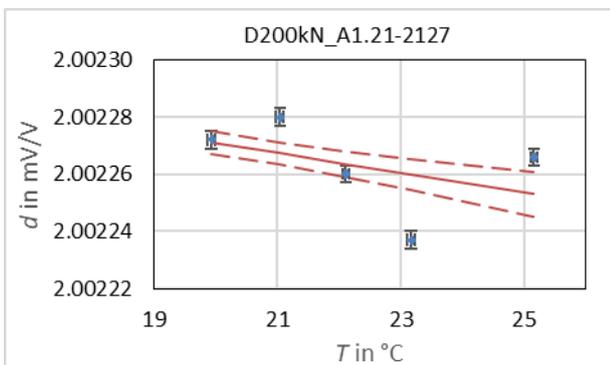


Figure 8: Result of the 200 kN transducer at a force of 200 kN and a reference temperature of 20.5 °C (further details as in Figure 3)

The 20 kN transducer (Figure 6) has a slope of $-3 \cdot 10^{-5}$ (mV/V)/K at 2 mV/V (20 kN force). The absolute value is larger than that of the 50 kN transducer. The linearity of the measured values is

not as perfect as that of the 50 kN transducer in Figure 3; the relative deviation between RSS and SSD is 0.970.

A special feature of the 100 kN transducer (Figure 7) is the positive slope of the deflection/temperature function, whereas the 20 kN and 50 kN transducers have a negative slope. The absolute value of the slope is slightly larger than that of the 20 kN transducer and amounts to $2.5 \cdot 10^{-5}$ (mV/V)/K at 1 mV/V (50 kN force) and $3.9 \cdot 10^{-5}$ (mV/V)/K at 2 mV/V (100 kN force). Apart from the 23 °C result, the values are very linear. Although the reason for the deviation of the 23 °C result has not yet been found, this single value had no significant effect on the fit function, as can be seen in Figure 7. On the other hand, this result indicates that temperature coefficients should not be determined from measurements at only two different temperatures.

Finally, the deflection of the 200 kN transducer (Figure 8) did not show good temperature sensitivity. The variations of the deflection appear to be random. On the other hand, the overall relative span of the deflection values in the temperature range measured is 21 ... 24 ppm; this is comparable with the corresponding value of the 50 kN transducer, which showed the lowest temperature sensitivity. Nevertheless, the behaviour of the 200 kN transducer should be investigated further. The results obtained to date would not be sufficient to calculate the corrections.

The last step in these calculations is the determination of the function $u(\tilde{d})$, which describes the standard uncertainties associated with the fit function \tilde{d} . This can be realised by applying higher-order regression methods such as cubic regression to the uncertainty values calculated, see Figure 9.

6. APPLICATION

The temperature influence measured on a single transducer from different measurement campaigns or even on different transducers can be compared by determining (as the final step) the regression function $\tilde{d}(T)$ and the associated uncertainty $u(\tilde{d}) = u(\tilde{d}(T)) = u(T)$. For the example in Figure 3 with $T_{\text{ref}} = 20.5$ °C, we obtained the following functions (T in °C)

$$\begin{aligned} \frac{\tilde{d}(T)}{\text{mV/V}} &= -3.7 \cdot 10^{-6} \cdot \frac{T}{\text{°C}} + 0.801070, \\ \frac{u(T)}{\text{mV/V}} &= -1.45 \cdot 10^{-8} \cdot \left(\frac{T}{\text{°C}}\right)^3 + 1.061 \cdot 10^{-6} \cdot \left(\frac{T}{\text{°C}}\right)^2 - 2.524 \cdot 10^{-5} \cdot \frac{T}{\text{°C}} + 0.0001981. \end{aligned} \quad (15)$$

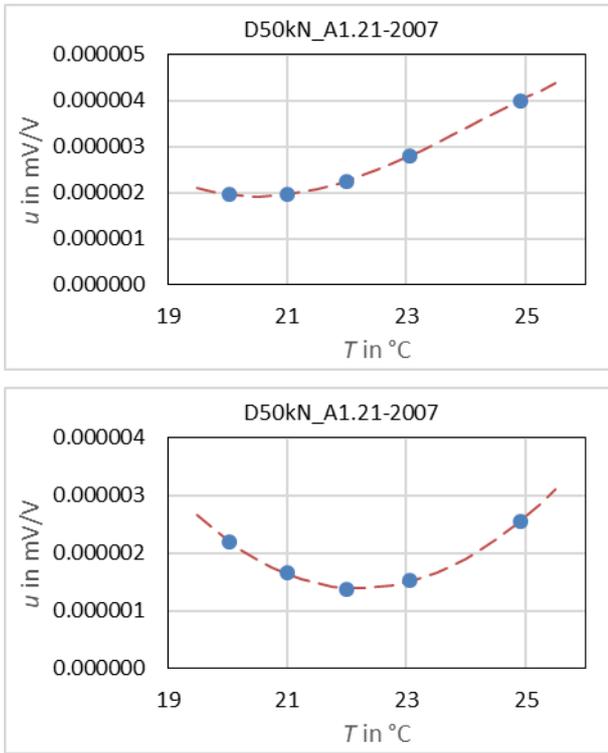


Figure 9: Standard uncertainty (1σ) part of the result from Figure 3, third-order regression function $u(\tilde{d})$ (dashed line) based on the calculated values $u(\tilde{d}_i) = u(\tilde{d}(T_i))$ (blue dots) for reference temperatures of 20.5 °C (top) and 22.196 °C (bottom)

However, in most cases, the results will be applied in order to correct comparison measurement results $d, u(d)$ obtained at deviating temperatures T . Usually, the corrected values \hat{d} are calculated in accordance with

$$\hat{d}(T_{\text{ref}}) = d(T) + q \cdot \Delta T, \quad \Delta T = T_{\text{ref}} - T \quad (16)$$

where the intercept r does not appear. When a measurement result is corrected, the regression function is defined in such a way that it runs through the uncorrected result, meaning that the intercept is no longer a free parameter. Nevertheless, the regression function $\tilde{d}(T)$ has an associated uncertainty based on the uncertainties of the slope $u(q)$ and the intercept $u(r)$. The formal application of the GUM to (16) would yield a result that contains no contribution of $u(r)$

$$u^2(\hat{d}) = u^2(d) + q^2 \cdot u^2(T) + (T_{\text{ref}} - T)^2 u^2(q). \quad (17)$$

Therefore, another model is proposed instead of (16) – namely,

$$\begin{aligned} \hat{d}(T_{\text{ref}}) &= d(T) + \Delta \tilde{d}(T_{\text{ref}} - T) \\ u^2(\hat{d}) &= u^2(d) + u^2(\Delta \tilde{d}), \end{aligned} \quad (18)$$

where

$$\begin{aligned} \Delta \tilde{d}(T_{\text{ref}} - T) &= \tilde{d}(T_{\text{ref}}) - \tilde{d}(T) \\ u^2(\Delta \tilde{d}) &= u^2(\tilde{d}(T_{\text{ref}})) + u^2(\tilde{d}(T)). \end{aligned} \quad (19)$$

This means that, instead of forcing the regression function to run through the uncorrected result, a

correction value $\Delta \tilde{d}$ is added to this result. This value can be calculated from the increase (or decrease for a negative slope) of the regression function related to the temperature deviation $T_{\text{ref}} - T$. This increase (or decrease) of the function is uncertain and has a value $u(\Delta \tilde{d})$. The uncertainties calculated according to (19) are larger than those yielded by (18), an effect caused mainly by the uncertainty of the intercept. For the example in Figure 3, the standard uncertainty of the corrected deflection ($\Delta T = 0.5$ K) in accordance with (18) is $3.03 \cdot 10^{-6}$ mV/V, whereas the same value calculated in accordance with (19) amounts to $4.08 \cdot 10^{-6}$ mV/V.

7. SUMMARY

The application of standard uncertainty calculation methods to the linear regression of measurement results using the least-squares approach was investigated. The method proposed can be used to calculate the uncertainty of a linear regression function based on the uncertainty of its slope and its intercept. The method can be applied to the correction of measurement results obtained at deviating temperatures but is not limited to force transducers: it can also be applied, for example, to torque or pressure transducers.

8. ACKNOWLEDGEMENTS



This work is part of the project 18SIB08, funded by the EMPIR programme.

I would like to thank my colleague Norbert Tetzlaff for his accurate measurements and calibrations in the 200 kN force standard machine at PTB.

9. REFERENCES

- [1] HBM: Data sheet of the Z30A force transducers. Online [Accessed 20200108]: <https://www.hbm.com/fileadmin/mediapool/hbmdoc/technical/B02075.pdf> (Registration necessary)
- [2] GTM: Data sheet of the KTN-D force transducers. Online [Accessed 20200108]: https://www.gtm-gmbh.com/fileadmin/media/dokumente/produkte/datenblaetter/de/Datenblatt_Serie_KTN-D_20170419.pdf
- [3] R. Kumme, H. Kahmann, F. Tegtmeier, N. Tetzlaff, D. Röske, PTB's New 200 kN Deadweight Force Standard Machine, Proc. of the IMEKO 23rd TC3, 13th TC5 and 4th TC22 International Conference, 30 May to 1 June 2017, Helsinki, Finland. Online [Accessed 20200108]:

<https://www.imeko.org/publications/tc3-2017/IMEKO-TC3-2017-032.pdf>

- [4] ISO/IEC Guide 98-3:2008, Uncertainty of measurement - Part 3: Guide to the expression of uncertainty in measurement (GUM:1995). Online [Accessed 20200108]: <https://www.iso.org/standard/50461.html>

PDF version:

https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf

- [5] Maxima, a Computer Algebra System. Online [Accessed 20200614]: <http://maxima.sourceforge.net>