ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS

Volume 64 59 Number 2, 2016

http://dx.doi.org/10.11118/actaun201664020503

CALIBRATION AND ITS USE IN MEASURING FUEL CONSUMPTION WITH THE CAN-BUS NETWORK

Adam Polcar[1], Jiří Čupera[1], Vojtěch Kumbár[1]

[1] Department of Technology and Automobile Transport, Faculty of AgriSciences, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

Abstract

POLCAR ADAM, ČUPERA JIŘÍ, KUMBÁR VOJTĚCH. 2016. Calibration and its Use in Measuring Fuel Consumption with the CAN-Bus Network. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis,* 64(2): 503–507.

The article deals with the description of calibration and its use in measuring fuel consumption via the CAN-Bus network. The CAN-Bus communication network holds great potential for the area of vehicle testing, since it allows the utilization of internal sensors of the vehicle to monitor its output parameters. In order to ensure proper accuracy of the measurements, however, it is best to perform calibration of the relevant sensors using certified measuring instruments. The main focus of the article is the calibration of the fuel consumption as measured by the CAN-Bus network based on fuel consumption measured by precision mass flow meters. Using the least squares method and a regression triplet, a calibration curve was created from the measured data to describe the relation between the two variables. The calibration curve was also used to calculate the reverse estimates along with confidence intervals for randomly chosen fuel consumptions obtained from the CAN-Bus network. To express the accuracy of the calibration methods, we also defined the limit values which express when the signal is still statistically significant enough to be distinguished from noise.

Keywords: least squares method, regression triplet, reverse estimate, estimate confidence interval, calibration limits

INTRODUCTION

CAN-Bus (Control Area Network) is the most common communication network used in vehicles. The parameters of the network and its data accuracy is set by standards ISO 11898, ISO 11783-2, SAE J2411 and SAE J1939. The network topology allows quick communication between the vehicle's control units. One of the big advantages of the communication network is that it can also be used to record the operating parameters of the vehicle for both diagnostic purposes and within a testing system (Bauer *et al.*, 2013; Vlk, 2006; Štěrba *et al.*, 2011). When testing a vehicle, it is thus not necessary to install external sensors, which could often be expensive and time consuming. A problem, however, lies in the accuracy of the data recorded this way. It is therefore advisable to calibrate the data from the CAN-Bus network via precision measuring instruments before using the internal sensors.

This article deals with the description of the calibration methods for compiling a calibration model which would describe the relationship between the mass consumption of fuel measured by laboratory sensors and the fuel consumption recorded by the CAN-Bus network in an agricultural tractor.

Generally speaking, calibration consists of two steps. First, the compilation of a calibration model and, second, the use of this model. The compilation of a calibration model is performed using regression methods. When using a calibration model, an inverse function is being solved, i.e. for a measured feedback y, a corresponding value x is being sought along with its static characteristics. There are two basic types of calibration: absolute and comparative. Absolute calibration searches for a relation between a measurable quantity, called signal, and a quantity which determines the state or the properties of the system. In comparative calibration, one device is calibrated in relation to another one (Meloun and Militký, 2012; Scheffé, 1973).

When building calibration or regression models, the most commonly used is the least squares method (Jarušková, 1996; Chatterjee and Hadi, 2006). However, the least squares method provides sufficient parameter estimates only when the assumptions regarding data and the regression model are met. If these assumptions are not met, the results calculated by the least squares method lose their properties. For this reason, the compilation of a regression model consists of a proposal and a subsequent regression diagnostic called regression triplet. The regression triplet consists of data critique, model critique and method critique. The process of constructing a regression model is described in detail for example in Meloun and Militký (2001).

MATERIALS AND METHODS

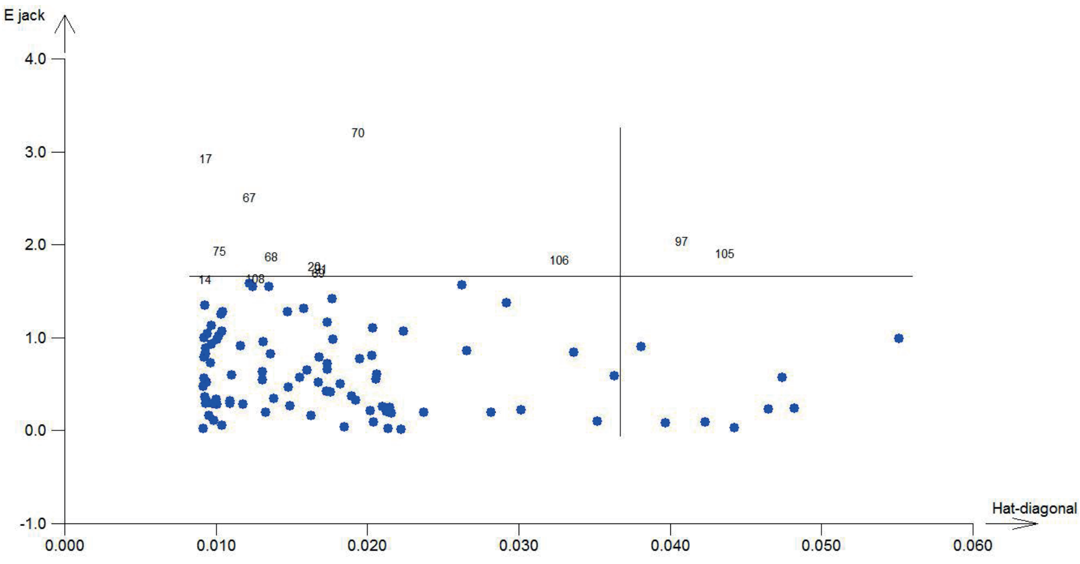
The data for calibration was obtained when measuring Claas Axion 850 Cebis tractor in a cylinder testing room in the laboratories of the Department of Technology and Automobile Transport of the Mendel University in Brno. The tests included the measurement of fuel consumption using mass flow meters Coriolis Sitrans FC MassFlo Mass 6000. In order to minimize the influence on the operation of the tractor's fuel system, two flow meters were used in differential connection into the low-pressure portion of the fuel system. The measurement accuracy of the flow meters is 0.1% of the measured value. The flow measurement range is 0–1,000 kg.h^{-1} with the range of the specific weight of measured agents being 0–2,900 kg.m^{-3} (Čupera *et al.*, 2010).

During the test, data was read from the CAN-Bus network by a computer connected to the tractor via DLC using proprietary software created at the Department of Technology and Automotive Transport. The communication speed was 250 kbps with the data reading frequency from the CAN-Bus network being 20 Hz.

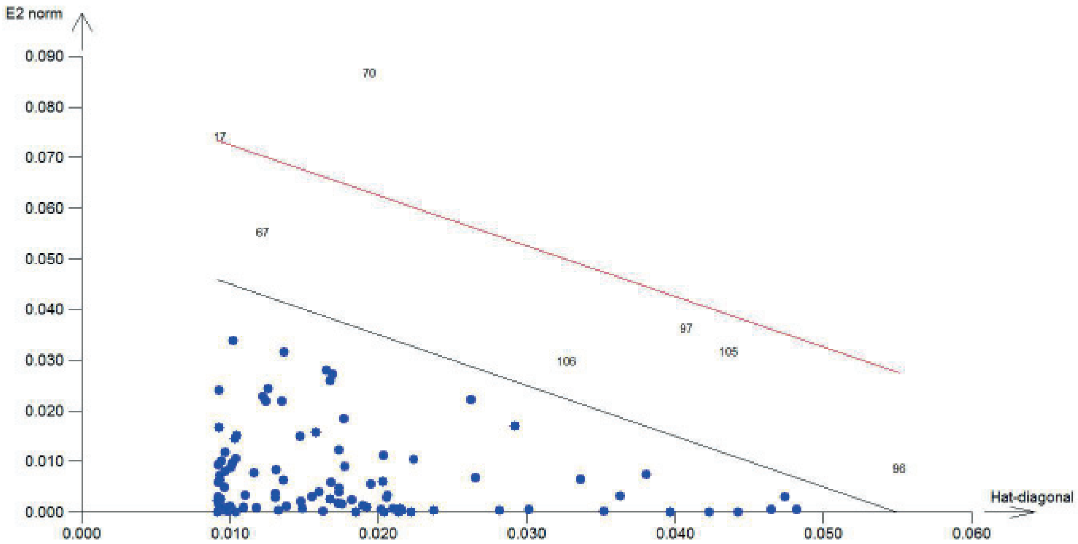
To be able to determine the fuel consumption in various modes of the engine, the tractor was connected (via the power take-off shaft) to a vortex dynamometer, type: V500. Using the vortex dynamometer, the engine of the tractor was stressed to the required torque and revolutions. Measurements were performed using a static method in line with the OECD standard.

Statistical analyses, creation of the regression model and the calculation of the calibration limits and reverse estimates were done using QC.Expert 3.3 and ADSTAT 1.25 software. Via regression analysis, the measured values were used to compile a regression model. Before the compilation itself, the data had to be critiqued and checked for outliers. As reported for example by Meloun *et al.* (2002) and Massart and Kaufman (1986), the presence of outliers in the data influences not only the correctness of the regression coefficient calculation, but also the accuracy of the subsequent calibration. Figs. 1 and 2 display the graphs of influential points, Williams graph and Pregibon graph, which are one of the many used to identify outliers. The graphs show the presence of several outliers in the measured fuel consumption data.

After removing the outliers from the data, the estimates were recalculated and a more accurate linear model was compiled. The significance level was selected before the calculation itself as $\alpha = 0.05$. The statistical significance of the individual regression coefficients and the correctness of the calculated model was subsequently verified using a number of statistical tests, such as: Student's t-test, Fisher-Snedecor test of regression model significance, Cook-Weisberg test of residues



1: *Williams graph*



2: *Pregibon graph*

heteroscedasticity (constancy of variance), Jarque-Berra normality test of residues etc. The calculated, more accurate model expressing the dependence of consumption in the CAN-Bus network in $l.h^{-1}$ (y) on the fuel consumption in the Coriolis flow meters in $kg.h^{-1}$ (x) is as follows:

$$y = -3.845(0.113) + 1.311(0.004)x. \qquad (1)$$

In parentheses, there are estimates of the standard deviations of given parameters. As the linear equation (1) shows, the absolute term has a negative value. That means, that a zero fuel consumption shown by the Coriolis flow meters corresponds to a CAN-Bus network measurement shown as –3.845 $l.h^{-1}$. However, a negative fuel consumption is not possible. It must thus be stated, that the calculated calibration curve is valid only for the range of measured values i.e. x ε <5.99; 37.85> in $kg.h^{-1}$ and y ε <4.38; 45.27> in $l.h^{-1}$. To be able to also describe the relation between x ε <0; 5.99) and y ε <0; 4.38), it is necessary to carry out additional measurements and complete the data matrix with these values. Since we can assume that the relation between the variables at low values will be parabolic (both consumptions will start at point 0; 0), the creation of the calibration model will require the use of a spline function.

RESULTS AND DISCUSSION

Using the completed equation (1) a reverse estimate can now be performed. The use of this equation or calibration model lies in the best possible estimation of the unknown value of x, i.e. in determining the mass consumption of fuel based on one or several repeated measurements of feedback y (fuel consumption from the CAN-Bus network). The reverse estimate is the main point of the calibration.

The reverse estimate and its confidence interval can be determined in several ways. Often a direct estimate and Naszodi's revised estimate is used (Meloun and Militký, 2012). A direct estimate is achieved based on the equation:

$$\hat{x}^* = \bar{x} + \frac{y^* - \bar{y}}{b_1}, \qquad (2)$$

where

y^* ... the measured value of the signal (or the average of $\bar{y}^*$ for M > 1 of repeated measurements),

b_1 ... the estimate of the calibration line direction,

$\bar{y}$ .... and $\bar{x}$ is the average of the variables.

The estimate using direct estimation is generally considered deviated. Correction of the deviation is performed using Naszodi's revised estimate (Meloun *et al.*, 2002):

$$\hat{x}^*_B = \bar{x} + \frac{(y^* - \bar{y})b_1}{b_1^2 + \dfrac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}. \qquad (3)$$

These point estimates, however, must be completed with a confidence interval. For the calculation of a 95% confidence interval for direct estimate and Naszodi's revised estimate, the following relations are used (at asymptotic normality):

$$L_L = \hat{x}^* - 1.96\sqrt{D(\hat{x}^*)}, \qquad (4)$$

$$L_U = \hat{x}^* + 1.96\sqrt{D(\hat{x}^*)}, \qquad (5)$$

where $D(\hat{x}^*)$ represents the dispersion of the x parameter estimate. If the estimate of the calibration line direction is sufficiently accurate, the approximate confidence interval can be used for the unknown parameter in the following form (Meloun *et al.*, 2002):

$$L_{L,U} = x^* \pm t_{1-\alpha/2}(n-2)\frac{\hat{\sigma}}{|b_1|}\sqrt{\frac{1}{n}+\frac{(y^*-\bar{y})^2}{b_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2}}. \quad (6)$$

For the calculation of reverse estimates, several fuel consumption values were set, which can be obtained from the CAN-Bus network. These values ranged from 5 to 45 $l.h^{-1}$ in increments of 5 $l.h^{-1}$. For these values, direct and Naszodi's estimates were calculated. Along with these estimates, 95% confidence intervals were calculated. Tab. I lists the results of the reverse estimates along with confidence intervals.

As the table shows, both estimates are the same. Therefore, the conclusions of the authors (Meloun and Militký, 2012) are confirmed → having sufficiently accurate data with small dispersion around the regression line, a simple classic direct estimate is sufficient.

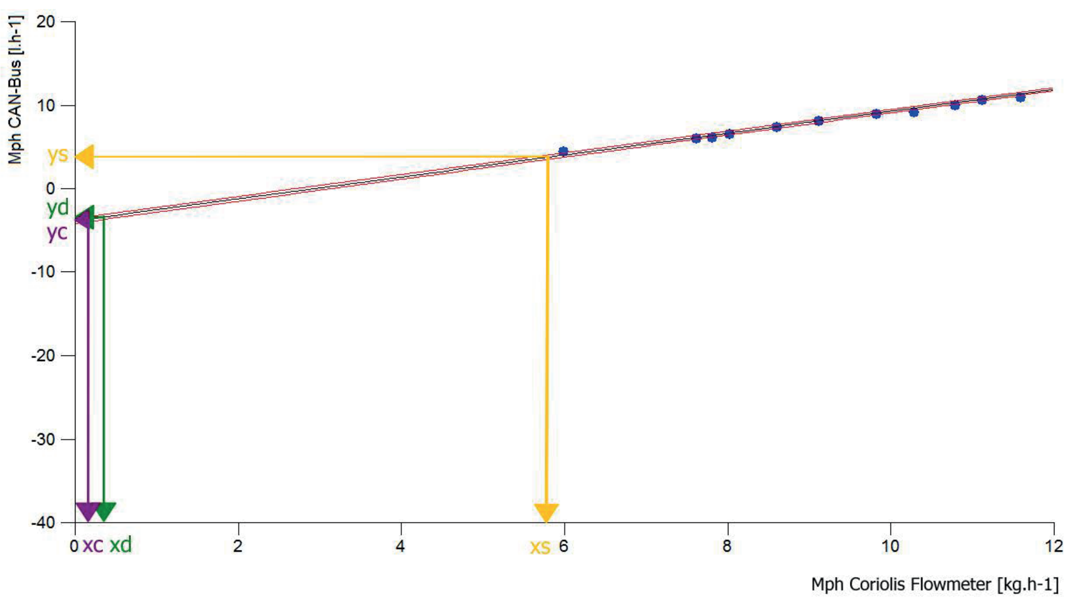
The next step in the calibration task is the calibration accuracy, expressed via the calibration limits. Calculating the calibration limits establishes the limit values for which the signal is still statistically differentiated from the noise (Zorn *et al.*, 1999; Meloun and Militký, 2012). The accuracy of the calibration is thus expressed by the critical level y_c, the detection limit y_D and the determination limit y_s. Fig. 3 shows a part of the calibration curve with the above calibration limits highlighted.

The critical level represents the upper limit of the confidence interval (in this case, 95% of the confidence interval) of the signal prediction from the calibration model, for fuel consumption in [$kg.h^{-1}$] equal to zero. Values lower than y_c are considered noise, or a blank (Kupka, 2000). In our case, the values are less than −3.621 $l.h^{-1}$, or less than 0.171 $kg.h^{-1}$(x_c).

The detection limit y_D corresponds to the fuel consumption provided by the CAN-Bus network for which the lower limit of 100(1 − α) % of the confidence interval of the signal prediction equals y_c. The detection limit for mass consumption of fuel x_D indicates the minimum fuel consumption measured by the Coriolis flow meter which is differentiable from the zero value with a probability of (1 − α).

I: *Calibration Table (C.I. – Confidence Interval)*

Measured Values (fuel consumption from CAN-Bus [$l.h^{-1}$]) y^*	Direct Estimate (fuel consumption [$kg.h^{-1}$]) x^*	"Naszodi's Estimate x^*B"	Lover Limit LL = x^* – C.I.	Upper Limit LU = x^* + C.I.
5	6.74	6.74	6.18	7.3
10	10.56	10.56	10	11.11
15	14.37	14.37	13.82	14.92
20	18.18	18.18	17.63	18.73
25	21.99	21.99	21.45	22.54
30	25.81	25.81	25.26	26.35
35	29.62	29.62	29.07	30.17
40	33.43	33.43	32.88	33.98
45	37.24	37.24	36.63	37.8



3: *Part of the calibration curve with highlighted calibration limits*

The detection limit was calculated as $y_d = -3,399$ l.h^{-1}, $x_d = 0.34$ kg.h^{-1}. As is evident from the critical level and the detection limit, the values of y_c and y_d are negative, which is, as already mentioned, unrealistic. These results also confirm the need for the use of a spline function when creating the calibration model.

The last calculated critical level was the determination limit y_s. The determination limit is the smallest value in the signal for which the relative standard deviation of prediction is still sufficiently small and equal to C. For C, a value of $C = 0.1$ is usually chosen (Meloun and Militký, 2012). The determination limit was calculated as $y_s = 3.768$ l.h^{-1}, $x_s = 5.805$ kg.h^{-1}.

Fig. 3 also shows the calculated Working-Hotelling confidence bands of the regression line (red curves).

CONCLUSION

The main objective of these measurements and the article itself was the calibration of the measured fuel consumption read by the CAN-Bus network according to the fuel consumption measured by the mass flow meters. Using the least squares method and regression triplet, the data measured were then compiled into a calibration line describing the relation between both variables. When creating the calibration model, it was discovered that the fuel consumption has a linear course for the range of values we measured. In case the input data would also include low fuel consumption values starting from the origin, a spline function would have to be used when creating the calibration model.

Using the calibration curve, we also calculated the reverse estimates along with confidence intervals for randomly selected fuel consumption values from the CAN-Bus network readings. To express the accuracy of the calibration methods, we also defined the limit values which express when the signal is still statistically significant as opposed to noise.

The CAN-Bus communication network holds great potential for the area of vehicle testing, since it allows the utilization of internal sensors of the vehicle to monitor its output parameters. Due to the need for accuracy of measurement, however, it is advisable to perform a calibration of the relevant sensors using certified measuring instruments in order to avoid inaccurate data readings.

Acknowledgement

The research has been supported by the project TP 4/2014 financed by IGA AF MENDELU.

REFERENCES

BAUER, F., SEDLÁK, P., ČUPERA, J., POLCAR, A., FAJMAN, M., ŠMERDA, T. and KATRENČÍK, J. 2013. *Tractors and their Use* [in Czech: *Traktory a jejich využití*]. Praha: Profi Press s. r. o.

ČUPERA, J., BAUER, F., SEDLÁK, P., HAVLÍČEK, M., FAJMAN, M., POLCAR, A., VYKYDAL, P. and TATÍČEK, M. 2010. Transport in Agriculture and its Impact on Environment [in Czech: Doprava v zemědělství a její dopad na životní prostředí]. *Zpráva o projektu*. Brno: Mendel University in Brno.

CHATTERJEE, S., HADI, A. S. 2006. *Regression Analysis by Example*. 4th edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

JARUŠKOVÁ, D. et al. 1996. Mathematical Statistics [in Czech: *Matematická statistika*]. Praha: České vysoké učení technické v Praze.

KUPKA, K. 2000. *QC-Expert ver. 3.3 – User Guide* [in Czech: *QC-Expert ver. 3.3 – Uživatelský manuál*]. Pardubice: TriloByte Statistical Software.

MASSART, D. L., KAUFMAN, L. 1986. Least median of squares: a robust method for outlier and model detection in regression and calibration. *Analytica Chimica Acta,* 187: 171–179.

MELOUN, M., MILITKÝ, J. 2012. *Interactive Statistical Data Analysis* [in Czech: *Interaktivní statistická analýza dat*]. Praha: Univerzita Karlova v Praze: Karolinum.

MELOUN, M., MILITKÝ J. 2001. Detection of single influential points in OLS regression model building. *Analytica Chimica Acta*, 439: 169–191.

MELOUN, M., MILITKÝ, J., KUPKA, K. and BRERETON, R. G. 2002. The effect of influential data, model and method on the precision of univariate calibration. *Talanta*, 57: 721–740.

SCHEFFÉ, H. 1973. A Statistical Theory of Calibration. *The Annals of Statistics*, 1: 1–37.

ŠTĚRBA, P., ČUPERA, J., POLCAR, A. 2011. *Cars: Diagnostic of Motor Vehicles II* [in Czech: *Automobily: Diagnostika motorových vozidel II*]. Brno: Avid, s. r. o.

VLK, F. 2006. *Diagnosis of Motor Vehicles* [in Czech: *Diagnostika motorových vozidel*]. Brno: prof. Ing. František Vlk, DrSc.

ZORN, M. E., GIBBONS, R. D. and SONZOGNI, W. C. 1999. Evaluation of Approximate Methods for Calculating the Limit of Detection and Limit of Quantification. *Environmental Science and Technology,* 33: 2291–2295.

Contact information

Adam Polcar: adam.polcar@mendelu.cz
Jiří Čupera: jiri.cupera@mendelu.cz
Vojtěch Kumbár: vojtech.kumbar@mendelu.cz