In-use Measurement of Ultrasonic Flowmeter based on Machine Learning

Mengna LI[1*], Chengze Lv[2], Wenli LI[2], Chunhui LI[1]

[1]National Institute of Metrology (NIM), Beijing, China
[2]China Jiliang University, Hangzhou, China
E-mail (corresponding author): limn@nim.ac.cn

Abstract

To guarantee the accuracy of ultrasonic flow meter, an in-use measurement system for ultrasonic flowmeter incorporating digital signal processors and machine learning approaches was proposed. Experimental analysis has been carried out to determine the variables affecting the accuracy of ultrasonic flowmeter. Based on random forest algorithm, we evaluated the contribution of different variables on the accuracy performance of ultrasonic flowmeter, and establish a model including variables extraction and prediction of flow deviation for in-use measurement of ultrasonic flowmeter. By obtaining data of the flowmeter signal index, flow rate characteristics, sound velocity and flow velocity etc., the flow deviation of ultrasonic flow meter is predicted using random forest algorithm, and the difference between predicated value and observed value is smaller than 0.76%. Furthermore, the degree of influence of different variables on the accuracy of ultrasonic flowmeter was analysed. The uncertainty of the prediction result was evaluated, with an extended uncertainty U = 0.92% ~ 0.22% (k=2).

1. Introduction

Ultrasonic flowmeters (USM), with advantages of high accuracy, no moving parts, and little or no pressure drop, are one of the fastest-growing technologies in natural gas flow measurement [1, 2]. Accuracy in flow rate measurements became a mandatory requirement for the trading. However, lack of efficient in-use measurement method is a major issue for ultrasonic flowmeters. The processing tasks for real-time performance are mandatory for guaranteeing the accuracy in flow rate measurements of ultrasonic flowmeter.

Machine learning, as a branch of artificial intelligence and computer science, focuses on the use of data and algorithms to imitate the way that humans learn and gradually improves its accuracy [3]. Few researchers have utilized the power of machine learning for flow measurements [4, 5]. We explored the potential of utilizing machine learning algorithms in metrology filed. We proposed an in-use measurement system for ultrasonic flowmeter incorporating digital signal processors and machine learning approaches. Experimental analysis has been carried out to determine the index affecting the accuracy of ultrasonic flowmeter. Based on random forest algorithm, we evaluate the contribution of different variables on the performance of ultrasonic flowmeter, and establish a model including feature extraction and prediction of flow deviation to guarantee the accuracy of ultrasonic flowmeter.

By obtaining data of the flowmeter signal index, flow rate characteristics, sound velocity and flow velocity etc., the flow deviation of ultrasonic flow meter is predicted using random forest algorithm. Based on the uncertainty of gas flow standard facilities and the performance of the algorithm, the article also elaborates the evaluation of uncertainty of the prediction result. The uncertainty of input features is analysed using the weight coefficients derived from the random forest model, and the performance of the algorithm are evaluated using time-series experimental data. Consequently, a comprehensive uncertainty assessment of in-use measurement of ultrasonic flowmeter is achieved.

2. Basic principles

2.1 Ultrasonic flowmeters

For ultrasonic flowmeter using the transit time method [6], the measurement is performed by transmitting a pulse from a transducer through the fluid to another transducer positioned downstream in the pipe, and back again. The flow velocity is obtained by measuring the difference in the time taken for the signal to travel up and downstream. Therefore, the transit time in the upstream direction t_u, and in the downstream direction t_d, can be expressed by Eqs. (1):

$$\begin{cases} t_{\mathrm{u}} = \frac{L}{c_{\mathrm{f}} - v\cos\theta} \\ t_{\mathrm{d}} = \frac{L}{c_{\mathrm{f}} + v\cos\theta} \end{cases} \tag{1}$$

where L is the path length, c_{f} is the velocity of ultrasound in the fluid, v is axial velocity measured along the path, and θ is the angle between the sound path and the axial velocity of flow.

From Eqs. (1), the velocity of ultrasound in the fluid and the axial velocity of flow can be calculated as follows

$$v = \frac{L}{2\cos\theta}\left(\frac{1}{t_{\mathrm{d}}} - \frac{1}{t_{\mathrm{u}}}\right) \tag{2}$$

$$c_{\mathrm{f}} = \frac{L}{2}\left(\frac{1}{t_{\mathrm{d}}} + \frac{1}{t_{\mathrm{u}}}\right) \tag{3}$$

An estimate of the averaged flowrate of the paths $\bar{v}$ can be obtained, multiplied by the area A, the volume flowrate q_{V} can be expressed as

$$q_{\mathrm{V}} = A\bar{v} \tag{4}$$

2.2 Random Forest algorithm (RF)

As for Machine Learning, bagging or bootstrap aggregation is a technique for reducing the variance of an estimated prediction function. Bagging seems to work especially well for high-variance, low-bias procedures. For regression, we simply fit the same regression tree many times to bootstrap sampled versions of the training data, and average the result [7, 8]. Random forest is a substantial modification of bagging that builds a large collection of de-correlated trees, and then averages them. Random forest is popular, and are implemented in a variety of packages. Some random forests reported in the literature have consistently lower generalization error than others [9].

An important feature of random forests is its use of out-of-bag (oob) samples. An oob error estimate is almost identical to that obtained by N-fold cross validation. Random forests can be fit in one sequence, with cross-validation being performed along the way. Furthermore, random forest “cannot overfit” the data, making it a feasible method to perform on different dataset of ultrasonic flowmeter.

Random forests can use the oob samples to construct a variable importance measure, apparently to measure the prediction strength of each variable. Based on the importance measure, we evaluate the contribution of different variables, which indicates the impact of variables on the accuracy performance of ultrasonic flowmeter.

3. In-use Measurement Model using RF

3.1 Variables collection and extraction

The information of USM is collected: the brand, nomination diameter, accuracy grade, flow range of the ultrasonic flowmeter, the instrument coefficient, the verification information, the on-site conditions, the installation conditions, and the basic information of the thermometer, manometer, component analyzer. The collected experimental data was processed by calculation and analysis, and three sets of variables are extracted as input variables of the flow prediction model for ultrasonic flowmeter, as shown in Table 1.

Table 1: Variables of flow deviation prediction for USM in-use measurement

Data type	Variables	References or Equations
Signal index	Signal-to-noise ratio (SNR)	USM
	Signal amplification (AGC)	USM
Flow characteristics	Velocity (VOG)	USM
	Averaged Velocity	USM
	Profile factor	$\frac{v_2 + v_3}{v_1 + v_4}$
	Symmetry	$\frac{v_1 + v_2}{v_3 + v_4}$
Metrology characteristics	Speed of sound (SOS)	USM
	Averaged SOS	USM
	Theoretical speed of sound	$C = \left[\frac{C_{\mathrm{p}}}{C_{\mathrm{v}}}\frac{R_{\mathrm{u}}T}{M}\left(Z + \rho_{\mathrm{m}}\left(\frac{\partial Z}{\partial \rho_{\mathrm{m}}}\right)_T\right)\right]^{0.5}$
	SOS deviation	$E = \frac{\bar{C}_f - C}{C} \times 100\%$
	Measurement Flow	USM

In Tab.1, USM refers reading values of flowmeter; Profile factor and Symmetry were calculated using 4-path USM as example; Theoretical speed of sound was obtained based on AGA report No.10.

3.2 Prediction model for flowr deviation

For 4 ultrasonic flow meters (one DN100, one DN150 and two DN200), input variables were extracted from experimental data. In addition, according to the latest real-flow calibration result of the ultrasonic flowmeter, the flow deviation was obtained, which served as the output value of the model.

The training sample set and validation sample set of 4 flow meters were obtained separately, ensuring the independence and randomness of the samples. The ratio of training sample and validation sample in the machine learning algorithm is generally set in the ratio of 2:1 to 4:1, so the number of data points of the training sample and verification sample of the 4 flow meters were shown in Table 2.

Table 2: Sample of the prediction model

No.	USM	Training sample number	Validation sample number
1	DN100	300	100
2	DN150	124	50
3	DN200-1	70	35
4	DN200-2	240	80

4. Result and evaluation

4.1 Predicted flow deviation

The training sample were used to train the model, and the total number of decision trees in the random forest algorithm is set to 500, and the number of features per node is set to 1/3 of the total number of input variables. Consequently, a prediction model of flow deviation and the importance measure of the variables were obtained. Using the established RF model, the predicted flow deviations using validation set were obtained. In addition, a comprehensive uncertainty assessment of in-use measurement of ultrasonic flowmeter is achieved.

For all 265 experimental points of the ultrasonic flowmeters, the absolute value of deviation between predicted values and observed values are smaller than 0.88%.

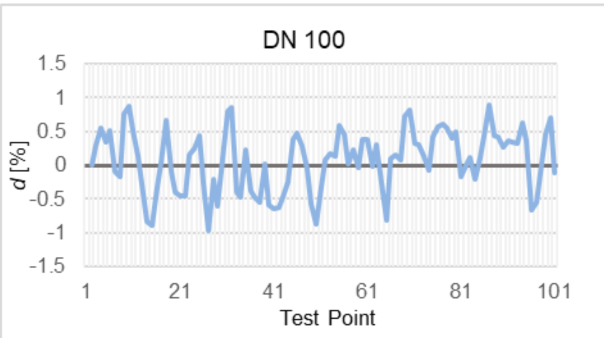

Figure 1: The deviation between predicted values and observed values for DN100

4.2 Importance measure of variables

A variable importance measure to indicate the prediction strength of each variable was constructed using oob samples. When the tree is grown, the oob samples are passed down the tree, and the prediction accuracy is recorded. Then the values for some certain a variable are randomly permuted in the oob samples, and the accuracy is again computed. The decrease in accuracy as a result of this permuting is averaged over all trees, and is used as a measure of the importance of that variable in the random forest.

Table 3: The importance measure of variables for DN100

Type	Variables	Importance Measure (Scores)
Metrology characteristics	Averaged VOG	6.72
	Measurement Flow	4.46
	SOS deviation	0.81
	Theoretical speed of sound	0.79
Flow characteristics	VOG of path 4	0.74
	VOG of path 3	0.69
	VOG of path 2	0.65
	VOG of path 1	0.64

According to the Importance Measure, the top 8 variables are obtained, as shown in Tab.3. (USM DN100). Among these variables, average VOG and measurement flow are the most important feature for the prediction of flow deviation in the performance of USM. Sound speed deviation and theoretical sound velocity also show high importance scores. In addition, the VOG of each path is also characteristic variables of importance.

4.3 Result evaluation

4.3.1 Prediction model performance

The performance of the RF algorithm is evaluated using variance explained (Var explained). The Var explained represents the relationship between the predictors and the response variables and indicates the goodness of fit of the model, which can be understood as the determinant of the fit (R^2) of the model [10,11]. In addition, predicted data and observed data residuals are the best quantitative indicators of the difference between the model and the real process, and they provide valuable information that can be used to assess the uncertainty of the model. The uncertainty of the output values of the prediction model is analyzed by using the residual distribution of the model and the relationship between the input variables and predictors [12~14]. Thus, the uncertainty $u_{r,fit}$ of the RF algorithm can be expressed as the residuals of the model:

$$u_{\mathrm{r,fit}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Q_{\mathrm{dev,fit}} - Q_{\mathrm{dev}})^2} \quad (5)$$

Where $Q_{\mathrm{dev,fit}}$ is the predicted value of the flow deviation of the ultrasonic flowmeter, and the Q_{dev} is the observation of the flow deviation. The goodness of fit of the four ultrasonic flow meters and the uncertainty of the model are shown in Table 4.

Table 4: Evaluation of the performance of RF model

USM	Var Explained [%]	Model Uncertainty ($u_{\mathrm{r,fit}}$) [%]
DN100	61.54	0.46
DN150	86.91	0.11
DN200-1	91.38	0.27
DN200-2	86.76	0.06

4.3.2 Uncertainty analysis of variables

- Theoretical speed of sound

Theoretical sound speed values were calculated using data of temperature, pressure, components, etc., according to the calculation formula provided by AGA Report No.10 [15]. The uncertainty of the theoretical sound including the uncertainty of the calculation method, pressure measurement, temperature measurement and Components measurement.

Table 5: Uncertainty of theoretical speed of sound

Variables	Source	Standard Uncertainty [%]
$u_{\mathrm{r}}(W)$	Calculation method	0.09
$u_{\mathrm{r}}(p)$	Pressure measurement	0.0577
$u_{\mathrm{r}}(T)$	Temperature measurement	0.005
$u_{\mathrm{r}}(M)$	Components measurement	0.013
Combined standard uncertainty $u_r(C)$= 0.108 %		
Combined expanded uncertainty $U_r(C)$= 0.22 %(*k*=2)		

- Deviation of sound speed

The deviation between each channel of the ultrasonic flowmeter does not exceed 0.035%, and it obeys uniform distribution within range. The sound speed deviation of the ultrasonic flowmeter is less than 0.06%. Considering the sound speed deviation uniformly distributed within the changing boundary. The uncertainty of deviation of sound speed is:

$$u_r(e) = \sqrt{(\frac{0.035\%}{\sqrt{3}})^2 + (\frac{0.06\%}{\sqrt{3}})^2} = 0.0401\% \qquad (6)$$

- Averaged velocity

According to equations (1) to (4), the common parameters for velocity and measured sound speed include: the length of the path and the transit time. The parameters of the involved in the volume flowrate calculation include the length of the path, angle between the sound path and the axial velocity of flow, and flow area, which are fixed parameters of the USM. The influencing factors of flowrate can be attributed to the sound wave transmission time. Therefore, the uncertainty of averaged velocity is mainly derived from the uncertainty of theoretical sound velocity and the uncertainty of sound speed deviation is:

$$\sqrt{u_{\mathrm{r}}(C)^2+u_{\mathrm{r}}(e)^2} = \sqrt{0.108^2 + 0.0401^2} = 0.115\% \quad (7)$$

- Flow measurement

The uncertainty of the flow measurement of the ultrasonic flowmeter needs to consider the uncertainty of the flow measurement of the meter to be calibrated, the repeatability of the measurement, and the temperature, pressure, compression factor etc. Based on the Law of conservation of mass, the volume flowrate is:

$$q_{\mathrm{v,MUT}} = q_{\mathrm{f}} \cdot \frac{p}{p_{\mathrm{MUT}}} \cdot \frac{T_{\mathrm{MUT}}}{T} \cdot \frac{Z_{\mathrm{MUT}}}{Z} \qquad (8)$$

Where $q_{\mathrm{v,MUT}}$ is the volume flowrate of meter being calibrated; q_{f} is the volume flowrate measured by the standard meter; p is the pressure at the standard meter; p_{MUT} is the pressure at the meter to be calibrated; T is the temperature at the standard meter; T_{MUT} is the temperature at the meter to be calibrated; Z is the compression factor of the gas flowing through the standard meter; Z_{MUT} is the compression factor of the gas flowing through the meter to be calibrated.

The uncertainty caused by temperature, pressure and compression factor is small and can be ignored in the calculation. Therefore, the uncertainty of the flow measurement can be calculated as follows

$$u_r(q_{\mathrm{f}}) = \sqrt{{u_r}^2(q_{\mathrm{v,MUT}})+{u_{\mathrm{R}}}^2(q_{\mathrm{f}})} = \sqrt{0.08\%^2 + 0.08\%^2} = 0.1132\% \qquad (9)$$

Based on the importance measure in Tab.3, the Weight coefficient is obtained using normalization of the values of importance measure. The standard uncertainty and weight coefficient of input variables are in Tab.6.

Table 6: The standard uncertainty and weight coefficient for 4 USM

Variables	Standard Uncertainty [%]	Weight coefficient
Averaged velocity	0.115	0.236 ~ 0.090
Flow measurement	0.114	0.199 ~ 0.059

Theoretical speed of sound	0.108	0.028 ~ 0.006
Deviation of sound speed	0.040	0.028 ~ 0.012

The combined standard uncertainty of input variables is $u_{\mathrm{r,input}} = 0.093\% \sim 0.048\%$.

4.3.3 Evaluation of the result

The standard uncertainty u of the prediction result includes the uncertainty caused by the measurement, that is, the uncertainty of the input variables $u_{\mathrm{r,input}}$, and the uncertainty of the model $u_{\mathrm{r,fit}}$, which can be expressed as:

$$u = \sqrt{{u_{\mathrm{r,input}}}^2 + {u_{\mathrm{r,fit}}}^2} \qquad (10)$$

Table 7: The standard uncertainty of prediction result

USM	**Standard uncertainty of input variables [%]**	**Standard uncertainty of model [%]**	**Standard uncertainty of result [%]**
DN100	0.048	0.46	0.46
DN150	0.080	0.11	0.14
DN200-1	0.071	0.27	0.28
DN200-2	0.093	0.06	0.11

The extended uncertainty of the in-use measurement of ultrasonic flowmeter based on random forest algorithm is U=0.92% ~ 0.22% (k=2).

5. Conclusion

To ensure the accuracy of ultrasonic flowmeter, an in-use measurement method for ultrasonic flowmeter was proposed by establishing an ultrasonic flowmeter flow deviation prediction and analysis model based on random forest algorithm. The main conclusions include:

(1) Establishing an ultrasonic flowmeter in-use measurement procedure to diagnose the performance of the ultrasonic flowmeter in use by obtaining parameters such as signal quality, flow rate index, and measurement characteristics of the ultrasonic flowmeter.

(2) Proposing a flow deviation prediction model of ultrasonic flowmeter based on random forest algorithm. The variables affecting the accuracy of ultrasonic flowmeter in use are analysed. Uncertainty evaluation of the prediction results and a comprehensive assessment of the in-use measurement method for the ultrasonic flowmeter were completed.

The research results can provide a basis for the promotion and application of the in-use measurement method for ultrasonic flowmeter. In addition, the prediction model based on the random forest algorithm can not only ensure the accuracy of the ultrasonic flowmeter at work, but also provide support for the research of the online calibration method of the ultrasonic flowmeter. In future research, the in-use measurement procedure of ultrasonic flowmeter will be further improved. A machine learning model and related evaluation system with stronger adaptability will be established to ensure the accuracy of the of ultrasonic flowmeter in use.

References

[1] Lynnworth L C, Yi L. Ultrasonic flowmeters: half-century progress report, 1955-2005. [J]. Ultrasonics, 2006, 44(4): e1371-e1378.

[2] Drenthen J. G., Boer G. d. The manufacturing of ultrasonic gas flowmeters[J]. Flow Measurement and Instrumentation, 2002, 12 (2): 89-99.

[3] Mitchell T M. Machine Learning[M]. McGraw-Hill, 2003.

[4] Zhao H , Peng L , Takahashi T , et al. ANN Based Data Integration for Multi-Path Ultrasonic Flowmeter[J]. IEEE Sensors Journal, 2013, 14(2):362-370.

[5] D Zheng, Zhang P , Xu T . Study of acoustic transducer protrusion and recess effects on ultrasonic flowmeter measurement by numerical simulation[J]. Flow Measurement & Instrumentation, 2011, 22(5):488-493.

[6] D. W. Splitzer, ed., Flow Measurement, ISA 2001, p. 515–571

[7] Hastie, Trevor & Tibshirani, Robert. (2001). The Elements of Statistical Learning Theory. 10.1007/978-0-387-21606-5_1.

[8] Dietterich T G. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization[J]. Machine Learning, 2000, 40(2):139-157.

[9] Breiman L. Random Forest [J]. Machine Learning, 2001, 45:5-32.

[10] Ling J, Templeton J. Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty[J]. Physics of Fluids, 2015, 27(8):042032-94.

[11] Solomatine D P, Shrestha D . A novel method to estimate model uncertainty using machine learning techniques[J]. Water Resources Research, 2009, 45(12).

[12] Xiao-Yong Pan H B S. Robust prediction of B-factor profile from sequence using two-stage SVR based on random forest feature selection. [J]. Protein & Peptide Letters, 2009, 16(12):1447-1454.

[13] Al-Mukhtar M. Random forest, support vector machine, and neural networks to modelling suspended sediment in Tigris River-Baghdad[J]. Environmental Monitoring and Assessment, 2019, 191(11):673.1-673.12.
[14] Lie M , Glaser B , Huwe B . Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models[J]. Geoderma, 2012, 170(3):70-79.
[15] AGA Transmission Measurement Committee Report No.10, Speed of Sound in Natural Gas and Others Related Hydrocarbon Gases[S].2003