Metrol. Meas. Syst., Vol. XXII (2015), No. 3, pp. 383–392.

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

SOFT SENSING METHOD OF LS-SVM USING TEMPERATURE TIME SERIES FOR GAS FLOW MEASUREMENTS

Weiqing Xu, Zichuan Fan, Maolin Cai, Yan Shi, Xiaomeng Tong, Junpeng Sun

Beihang University (BUAA), School of Automation Science and Electrical Engineering, Xueyuan Road 37, Haidian 100191, Beijing, China (xuweiqing2010@gmail.com, ✉ *cain8492@126.com,* +86 010 8233 9160, *caimaolin@buaa.edu.cn)*

Abstract

This paper proposes a soft sensing method of least squares support vector machine (LS-SVM) using temperature time series for gas flow measurements. A heater unit has been installed on the external wall of a pipeline to generate heat pulses. Dynamic temperature signals have been collected upstream of the heater unit. The temperature time series are the main secondary variables of soft sensing technique for estimating the flow rate. A LS-SVM model is proposed to construct a non-linear relation between the flow rate and temperature time series. To select its inputs, parameters of the measurement system are divided into three categories: blind, invalid and secondary variables. Then the kernel function parameters are optimized to improve estimation accuracy. The experiments have been conducted both in the single-pulse and multiple-pulse heating modes. The results show that estimations are acceptable.

Keywords: gas flow, soft sensor, support vector machine, temperature time series.

1. Introduction

The gas flow rate is an important parameter concerned with performance of many industrial systems. To monitor these systems, flow measurements are required [1–4]. In order to detect and quantify gas flow rates, non-invasive measurement techniques whose sensors are installed outside the gas pipelines and indirectly contact with gases are recommended, as these techniques have an evident advantage that it is not necessary to break pipelines to install sensors when measuring. Hence, the non-invasive measurement techniques have a potential to be widely used in many industrial applications [5–8].

The current non-invasive techniques based on static temperature fields, mechanical waves, radioactive rays and electric waves [5–14], focus on establishing simple and clear relations (*e.g.* the linear function) between the gas flow rates and parameters of physical fields. However, the relations are deteriorated because the physical fields are severely impeded when propagate through the pipeline wall while being slightly affected by gases. Moreover, to ensure the physical fields have sufficient energy to penetrate the pipeline wall so that the parameters could be captured effectively, elaborate and expensive devices are required. Thus, the current measurements are restrained in gas flow applications.

To deal with problems resulting from the non-invasive measurements, a soft sensing technique is introduced. The core of soft sensor is constructing a nonlinear function between indirectly measured primary variables (*e.g.* the gas flow rate) and measurable secondary variables [15, 16], and achieving optimal estimation of primary variables through transforming and computing secondary variables. By using the soft sensing technique, a measurement system which can use convenient and inexpensive sensors – though containing a nonlinear model – could be successfully applied to obtain gas flow rates.

Article history: received on Jan. 23, 2015; accepted on Apr. 13, 2015; available online on Sep. 21, 2015; DOI: 10.1515/mms-2015-0028.

To establish the nonlinear model, the supported vector machine (SVM), a learning algorithm with the statistic theory proposed by Vapnik in 1990s, has been adopted [17–19]. It is based on the theory of structural risk minimization and kernel function, and can reach the best compromise between the complexity of model and the ability of study. Thus, the best generalization ability in limited sample information can be obtained. To make the SVM more pragmatic, the least squares support vector machine (LS-SVMs) has been formulated and successfully applied to such problems, like classification and function estimation [20, 21].

This paper is organized as follows: in Section 2, a convenient measurement system based on the dynamic thermal characteristics of a pipeline is introduced, and followed by the LS-SVM regression model. In Section 3, a main procedure of the soft sensing method of LS-SVM using the temperature time series is presented. The experimental results are discussed in Section 4. Finally, Section 5 gives the conclusion of this work.

2. Basic principles

2.1. Non-invasive measurement system

A non-invasive measurement method is proposed in reference [22]. A heater unit is placed outside the pipeline, and it takes a long time for the temperature field (mostly affected by the pipeline wall) to become stationary. The methods based on the stationary temperature field are rejected [22] due to their long time delay (350 s). Such a delay is very bad for use, especially for the case when the flow rate is changing. Besides, such a delay does not exist in a traditional invasive calorimetric flowmeter (delay < 1 s). In this study, a novel measurement system based on the dynamic thermal characteristics of a pipeline is introduced for measuring the gas flow rate. The measurement is achieved by using dynamic temperature signals, *i.e.*, the temperature time series. Fig. 1 shows a schematic diagram of the novel measurement system.

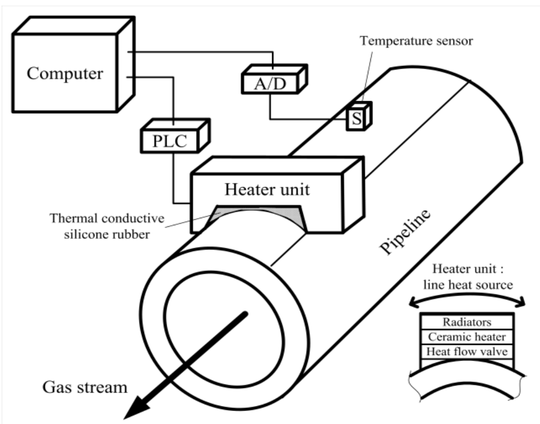

Fig. 1. A schematic diagram of the measurement system.

A heater unit which generates heating pulses is installed on the external wall of the pipeline. The gas stream flows in the direction of the arrow. A temperature sensor is attached upstream of the heater unit to detect the temperature of pipeline cooled by the gas.

It is known that the flow rate affects the intensity of convection heat transfer, and the temperature of pipeline can be a function of the flow rate. As a result, the flow rate can be obtained by computing the data captured by the temperature sensor.

Metrol. Meas. Syst., Vol. XXII (2015), No. 3, pp. 383–392.

2.2. LS-SVM regression model

Given the samples $\{x_i, y_i\}$, where $x_i \in R^m$, $i = 1,2, \ldots, N$ are the input data of soft sensor model, $y_i \in R$ are the output data, and N is the number of samples, the nonlinear regression function for LS-SVM can be expressed as [20]:

`

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x_i, x_j) + b, \tag{1}$$

where x_i is the input of LS-SVM regression model (*e.g.* the temperature signals of pipeline and heater unit), and $f(x)$ is estimation of its output, *i.e.*, the estimation of flow rate. b is the offset [18–19]. $\alpha_i \in R^{N\times 1}$ are Lagrange multipliers. The kernel function $K(x_i, x_j)$ is a symmetric function satisfying the Mercer condition. There are a lot of kernel functions, such as linear, polynomial, radial basic (RBF) and multi-layer perception ones. It has been proven by many experiments that RBF has good performance for a small range of parameters and a low space complexity. Therefore, the RBF $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2/2\sigma^2)$ is used in this paper.

3. Procedure of soft sensing method

The basic structure of soft sensing method of SVM is shown in Fig. 2. The inputs (the secondary variables shown in Table 1) of the SVM model include: the measurable variable X (*e.g.* ambient temperature), the measurable output variable θ (*e.g.* temperature of the pipeline) and the control variable u (*e.g.* temperature of the heater unit). The value of θ is affected by u. The output of SVM model is the optimal estimation $\hat{Y}$ (*e.g.* flow rate). Besides, Y^* is a priori information (a computed off-line value or an estimated variable sampled at a large sampling interval [15]), and can be used in the parameter identification or on-line self-correction of SVM model. In this study, Y^* means the label of the train data, *i.e.*, the known flow rate. X' denotes the blind variables shown in Table 1.

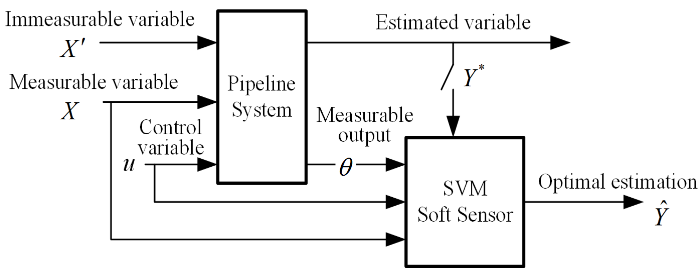

Fig. 2. The basic structure of SVM soft-sensor.

The steps of soft sensing method of SVM are summarized as:

1) Determine the sensor installation, the input and output variables, including the variable type and the number of variables.
2) Collect the sample data and divide them into two parts: the training set (validation set included) and the test set.
3) Establish the SVM soft sensor with the training set: a combination of regularization and kernel parameters, training SVM.
4) Perform cross validation and return to 3), until the combination is completed.
5) Select the optimal combination parameters to create the final SVM model and estimate the flow rates at the test points.

3.1. Sensor installation

The proposed soft sensing method is principally utilizing the temperature time series to estimate the gas flow rate. And the location of the sensor installation is important. The downstream area is not appropriate for the sensor to install in, because the sensor temperature varies non-monotonically with the flow rate. If the sensor is installed downstream of the heater unit, the sensor temperature is affected by two factors: the heat brought in by the gas from the heater unit and the heat taken away by the gas flowing downwards. An increase of flow rate could result in an increase of the heat brought in and taken away. Consequently, the variation of temperature is complicated, and not monotonous. However, if the sensor is installed upstream, the sensor temperature is affected only by the heat taken away and the variation is monotonous. The details about sensor installation are discussed in our previous work [22]. Hence, the sensor should be installed upstream of the heater unit to guarantee the monotonic change of upstream temperature with the flow rate, as the principle of our method illustrated in Fig. 3.

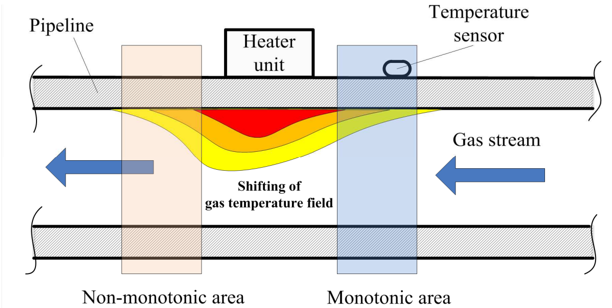

Fig. 3. A schematic diagram of sensor installation.

3.2. Selection of secondary variables

The soft sensing model of LS-SVM is a black-box model and its input, output and transfer characteristics are all accomplished by LS-SVM without considering internal workings. In the process of modelling first the input and output data were confirmed, and then the nonlinear relation of input and output was obtained in step (5).

There are many variables in the non-invasive measurement system. The variables which ought to be considered are classified into three categories, including secondary, blind and invalid variables, as shown in Table 1.

Table 1. Classification of all parameters.

Classification	**Parameters**
Secondary variables	temperature of the pipeline (θ), ambient temperature (X), temperature of the heater unit (u)
Blind variables (X')	heat capacity of the gas, application pressure of the gas, initial temperature of the unheated gas, density of the gas, dynamic viscosity of the gas, thermal conductivity of the gas
Invalid variables	wall thickness of the pipeline, nominal diameter of the pipeline, density of the pipeline material, thermal conductivity of the pipeline material, specific heat of the pipeline material

Metrol. Meas. Syst., Vol. XXII (2015), No. 3, pp. 383–392.

The secondary variables, as the input of SVM soft sensing model, must be selected according to such fundamentals as: sensitivity, specificity, applicability, accuracy, measurability and so on [23].

The blind variables cannot be measured or obtained from non-invasive measurements. Further, some variables could be regarded as constants during a measurement, with no influence on the soft sensing processing. They are called "invalid variables", *e.g.* the dimensions of a pipeline, the specific heat of pipeline material, the density of pipeline material, *etc*. As a result, variables which can be used as the secondary variables are limited in number. Additionally, to generate a temperature time series, two heating modes of measurement are discussed, as follows: a low frequency pulse produced by the heater unit in the single pulse mode; and a group of high frequency pulses produced in the multiple pulse mode. The single-pulse mode is applied for infrequent measurements, while the multiple-pulse mode is used for frequent measurements, and it requires more energy consumption. To define the temperature time series, simulation results of the temperature field model are taken from [22]. The curves in Figs. 4 and 5 show variations of pipeline temperature after the heat pulse is generated. A series of temperature sample points (time, temperature) can be obtained from each curve. It is defined as the corresponding temperature time series. The coloured blocks show how the pulses are generated by the heater unit. The red block indicates the heater unit operation in the heating state, while the blue one indicates it in the cooling state.

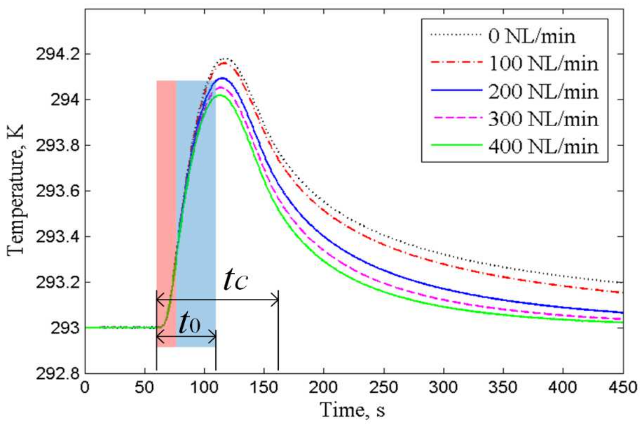

Fig. 4. The temperature time series at different flow rates for the single pulse mode.

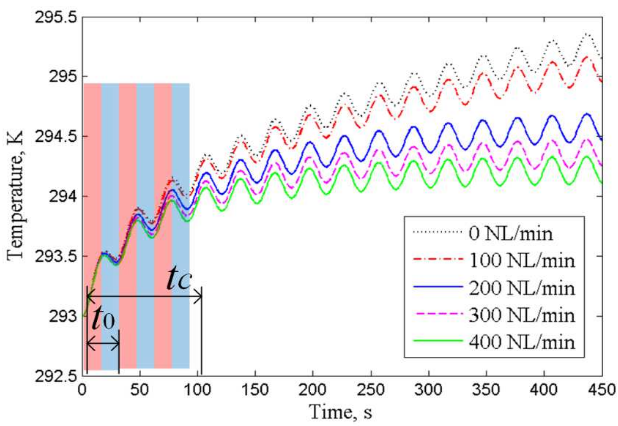

Fig. 5. The temperature time series at different flow rates for the multiple pulse mode.

In Fig. 4, the temperature time curves first rapidly rise and then slowly decline. These curves at different flow rates are harder to distinguish during the heating period than in the cooling period. In Fig. 5, the temperature time curves are harder to be distinguished during the early period compared with the later period for different flow rates.

Additionally, there are big differences between the temperature time curves at 100 NL/min and 200 NL/min, due to a huge variation of the convective heat transfer coefficient with changing flow states from the laminar to the turbulent one. These differences obviously influence estimation of the flow rate, which is illustrated in Figs. 6 and Fig. 7.

Thus, to distinguish the temperature signals with different flow rates, the period of signal collecting t_c should be a little longer than that of the heating signal t_0. In this paper, t_c is set to $2t_0$, and should not be much longer, so as not to significantly extend the total measurement time.

3.3. Parameter optimization

The performance indicators (*e.g.* processing time, root mean squared error and correlation coefficient of estimation) of SVM depend on selection of the kernel function, the kernel's parameters, and the regularization parameter. In this paper, a *radial basic function* (RBF) has been chosen as the kernel function, and the regularization parameter and kernel parameter are optimized to keep the error between estimation and expectation within an acceptable range. Considering the estimation accuracy, optimization processing of the kernel function parameters is performed.

Firstly confirm the aggregates of regularization parameter and kernel parameter. Then choose parameters from the aggregate to fit together, carry them through training by LS-SVM, check up with the forecast aggregate, and obtain optimization of the regularization and kernel parameters. Finally, these optimized parameters can be used to establish the LS-SVM model for estimation. Further, there are some popular algorithms for parameter optimization processing, such as the *grid search algorithm* (GSA), the *genetic algorithm* (GA), and the *particle swarm optimization* (PSO). In this paper, we do not focus on selection of these algorithms. Thus, GSA is used just for simplicity.

4. Simulations and experiments

4.1. Simulation results

Simulation is performed to validate the soft sensing method. i.e., the pipeline block (Fig. 2) is accomplished with a temperature field model [22] instead of a real pipeline. As the secondary variable, θ is built up with several temperature time series. They are divided into two sets: the training set (including the validation set) and the test set. The training set is used to establish the LS-SVM soft sensing model, and the test set is used to test accuracy of the model. In this study, θ is defined as 12 temperature time series (Increase of the number of temperature time series could improve accuracy of the model). 6 of them are chosen as the training set, other 6 ones build up the test set. X is the ambient temperature (equals to 293 K), $u(t)$ is the temperature of heater unit. The wave pattern of $u(t)$ is sinusoidal; its amplitude is 5 K, and period is 30 s. For parameter optimization processing, the grid search algorithm (GSA) is selected.

Simulation results are illustrated in Figs. 6 and 7 according to the two heating modes, respectively. To show the performance of estimation, a reference line representing actual values is given to easily confirm the results. Also, the relative root mean squared error (rRMSE) and the correlation coefficient (R) are used to quantify the performance. The relative root mean squared error is defined as:

Metrol. Meas. Syst., Vol. XXII (2015), No. 3, pp. 383–392.

$$E = \sqrt{\frac{\sum_{i=1}^{n}\left(f(x_i) - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}}, \; \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i . \tag{2}$$

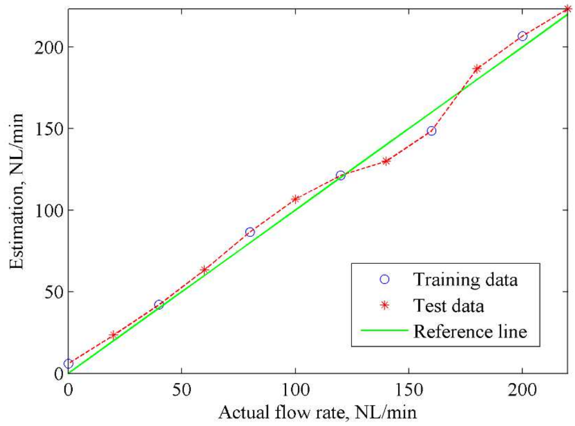

Fig. 6. The gas flow estimation using simulation signals for the single pulse mode (rRMSE = 8.909%, R = 99.662%).

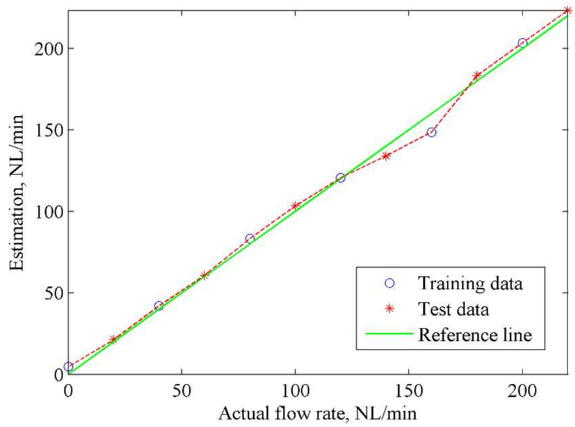

Fig. 7. The gas flow estimation using simulation signals for the multiple pulse mode (rRMSE = 5.055%, R = 99.891%).

It can be seen from Figs. 6 and 7 that: both methods for the single pulse and the multiple-pulse modes can reflect the effectiveness of soft sensing of LS-SVM, and the estimations are close to the reference line with acceptable deviations; there is a buckle at the flow rate of about 150 NL/min. This is because the temperature drops a lot when the flow state changes from the laminar to turbulent one. The training points are generated with the training set. Every element of the training set (temperature time series) is generated at an actual flow rate (horizontal coordinate). The vertical coordinate of the training points are estimated by the (1). The actual flow rates are not equal to the estimated ones and the training points are not at the reference line. The reason is described as follows:

The aim of the LS-SVM method is to minimize the generalization error which measures how well it generalizes onto unseen data. The neural network only cares about empirical risk minimization, which puts the training data (seen data) at the reference line and may work bad for the unseen data.

In addition, it should be emphasized that the samples of training data should be representative and widely distributed within the measurement range. Extrapolation of the LS-SVM model is not recommended for use. As shown in Fig. 8, although the samples for training are great in number, the estimations by extrapolation have significant deviations from the reference line at large flow rates.

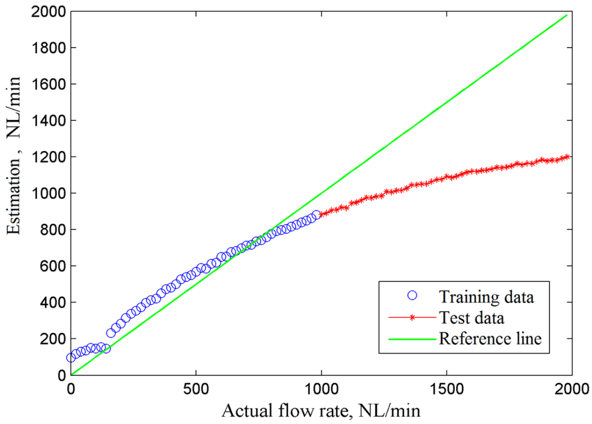

Fig. 8. A failure of estimation using simulation signals (rRMSE = 162.163%, R = 98.775%).

4.2. Experiment results

A convenient device is used to implement the measurement [22]. The heater unit applied in the experiments consisted of a heat flow valve, a ceramic heater and radiators, as shown in Fig. 1. The heat flow valve can be seen as a thermoelectric cooler. It utilizes the Peltier effect to generate a heat flux between the junctions of two different types of materials. The ceramic heater is a positive thermal coefficient ceramic. The radiators are constructed of cooling fins and Peltier units, the Peltier units are used to improve the efficiency of the radiators.

In addition, the temperature sensor used is a Pt 100 sensor (0~200°C, 0.2% FS, $\tau_{0.5} \leq 0.5$ s), and is placed at 10 mm away from the heater unit. The measured pipeline (DN 32 × 3) is a non-galvanized steel pipe (l = 1 m). More importantly, the signals obtained by the non-invasive measurement system are drowned by the intense environmental noise. Thus, the temperature signals are denoised using the wavelet algorithm before being input into the soft sensing model [24]. Consider an experimental example: the ambient temperature is 292 K, the temperature of heater unit is sinusoidal and its period is 10 s. A series of specific flow rates varies from 0 to 200 NL/min at the gauge pressure of 0.2 MPa, and the sampling time for temperature signals is 0.2 S. Then 6 samples of the temperature time series are chosen as the training set, and 6 samples are used for testing the estimation. The measurement results according to two heating modes are shown in Figs. 8 and 9.

It can be seen from Figs. 9 and 10 that: the estimations for both methods are approximations of the actual values of flow rate with obvious deviations, and relative root mean squared errors are more than 10%. Compared with the simulations, the measurements using real signals feature lower estimation performance. The variations of some blind variables, such as the initial temperature and gas pressure, have a significant influence on the temperature signals, and these parameters could be regarded as the system errors which are difficult to eliminate.

Metrol. Meas. Syst., Vol. XXII (2015), No. 3, pp. 383–392.

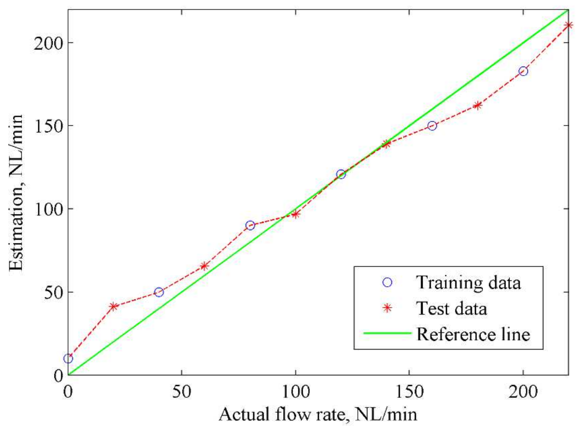

Fig. 9. The gas flow estimation using real signals for the single pulse mode (rRMSE = 17.899%, R = 99.523%).

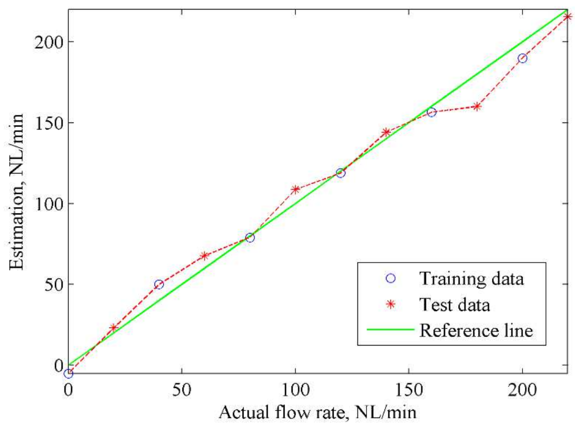

Fig. 10. The gas flow estimation using real signals for the multiple pulse mode (rRMSE = 14.357%, R = 99.257%).

As a result, estimations using real signals are acceptable. The accuracy and reliability of the measurement results can be improved by keeping some blind variables steady during the measurement. Thus, the measurement system is suggested to be installed away from the entrance of pipeline, to operate in the fully developed flow.

5. Conclusion

In this paper, a soft sensing method using temperature time series based on LS-SVM is proposed for gas flow measurement. The method uses the non-invasive technique whose sensor is installed outside the pipeline and not in direct contact with the gas. Experiments are performed according to two heating modes for infrequent and frequent measurements, respectively. The results show that both methods can reflect effectiveness of the soft sensing of LS-SVM. Furthermore, shortening the time for signal collecting (not less than the period of the heating signal) can diminish the measurement time without an excessive impact on results. In addition, the flow rate estimations still have some deviations, and the method proposed still requires improvement in terms of accuracy, if necessary. To keep some blind variables (*e.g.* initial temperature and pressure of gas) steady during the measurement is an effective way. As for the measurement time, the measurement can be accelerated significantly by upgrading the heater unit and shortening the period of the heating signal. Last but not least, although the measurements of the soft sensing method of LS-SVM are currently lower in accuracy compared those obtained by the invasive flowmeters, the new method could be used for range estimations and installation guidance for permanent flowmeters.

Acknowledgement

The authors wish to express their gratitude for the financial support from a Grant (51375028) of the National Natural Science Foundation of China.

References

[1] Cai, M.L., Kawashima, K., Kagawa, T. (2006). Power Assessment of Flowing Compressed Air. *Trans. ASME J. Fluids Eng.*, 128(2006), 402–405.

[2] Olczyk, A. (2008). Specific mass flow rate measurement in a pulsating flow of gas. *Metrol. Meas. Syst.*, 15(2), 165–176.

[3] Benson, J.M., *et al.* (1970). Thermal mass flowmeter. *Instrum. Control Syst.*, 43, 85–87.

[4] Laub, J.H. (1957). Measuring mass flow with the boundary-layer flowmeter. *Control Engng.*, 4, 112–117.

[5] Kolahi, K., Gast, T., Röck, H. (1994). Coriolis mass flow measurement of gas under normal conditions. *Flow Meas. Instrum.*, 5, 275–383.

[6] Keita, N.M. (1994). Behaviour of straight pipe Coriolis mass flowmeters in the metering of gas: theoretical predictions with experimental verification. *Flow Meas. Instrum.*, 5, 289–294.

[7] Webster, J.G. (1999). The Measurement, Instrumentation and Sensors Handbook. *Boca Raton*, 28.1–28.11.

[8] Viswanathan, M., Kandaswamy, A., Sreekala, S.K., Sajna, K.V. (2002). Development, modeling and certain investigations on thermal mass flow meters. *Flow. Meas. Instr.* 12, 353–360.

[9] Kei, T., Isao, S., Koichi, I. (1996). Simple temperature compensation of thermal air-flow sensor. *Sensors and Actuators A: Physical*, 57, 197–201.

[10] Kim, D.K., Han, I.Y., Kim, S.J. (2007). Study on the steady-state characteristics of the sensor tube of a thermal mass flow meter. *International Journal of Heat and Mass Transfer.*, 50, 1206–1211.

[11] Sazhin, O. (2013). Novel mass air flow meter for automobile industry based on thermal flow microsensor. I. Analytical model and microsensor. *Flow Measurement and Instrumentation*, 30, 60–65.

[12] Abdul, R., Rahiman, M.H., Chan, K.S., Nawawi, S.W. (2007). Non-invasive imaging of liquid/gas flow using ultrasonic transmission-mode tomography. *Sensors and Actuators A: Physical.*, 135, 337–345.

[13] Janka, K. (1984). Ion deflection air flow meter with constant deflection. *Rev. Sci. Instrum.*, 55, 976–982.

[14] Dyakowski, T. (1996). Process tomography applied to multi-phase flow measurement. *Meas. Sci. Technol.*, 7, 343–353.

[15] Joseph, B. (1999). Tutorial on inferential control and its applications. *Proc. IEEE American Control Conference*, San Diego, US, 3106–3118.

[16] Liu, L.C., Kuo, S.M., Zhou, M.C. (2009). Virtual Sensing Techniques and Their Applications. *Proc. IEEE Int. Conf. on Networking, Sensing and Control*, Okayama, Japan.

[17] Cherkassky, W., Mulier, F. (1998). *Learning from Data.* US: John Wiley & Sons.

[18] Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer Verlag.

[19] Saunders, C., Gammerman, A., Vovk, V. (1998). Ridge regression learning algorithm in dual variables. *Proc. Int. Conf. on Machine Learning*, Madison, Wisconsin.

[20] Suykens, J.A.K., Vandewalle, J. (1999). Least Squares Support Vector Machine classifiers. *Neural Processing Letters*, 9, 293–300.

[21] Muller, K., *et al.* (1997). Predicting Time Series with Support Vector Machines. *Proc. Int. Conf. on Artificial Neural Networks*, Lausanne, Switzerland, 999–1004.

[22] Fan, Z.C., Cai, M.L., Xu, W.Q. (2012). Non-invasive and non-intrusive gas flow measurement based on the dynamic thermal characteristics of a pipeline. *Meas. Sci. Technol.*, 23, 105303.

[23] Joseph, B. (1978). Brosilow Inferential control of processes. Part I. *Steady state analysis and design*, 24, 485–492.

[24] Fan, Z.C., Cai, M.L., Wang, H.H. (2012). An improved denoising algorithm based on wavelet transform modulus maxima for non-intrusive measurement signals. *Meas. Sci. Technol.*, 23, 045007.