Yuxue XU, Yun WANG, Tianhong YAN, Yuchen HE, Jun WANG, De GU, Haiping DU, Weihua LI, 2021. Quality-related locally weighted soft sensing for non-stationary processes by a supervised Bayesian network with latent variables. *Frontiers of Information Technology & Electronic Engineering,* 22(9):1234-1246. https://doi.org/10.1631/FITEE.2000426

Quality-related locally weighted soft sensing for non-stationary processes by a supervised Bayesian network with latent variables

Key words: Soft sensor; Supervised Bayesian network; Latent variables; Locally weighted modeling; Quality prediction

Corresponding authors: Tianhong YAN, Yuchen HE
E-mail: thyan@cjlu.edu.cn; yche@cjlu.edu.cn
ORCID: https://orcid.org/0000-0003-3916-3926
https://orcid.org/0000-0002-0528-2778

Motivation

1. Because of the complexity of industrial processes, model-driven soft sensors are usually impractical, cost-prohibitive, and time-consuming.
2. Nonlinear relationships among different process variables are very common in the process industry which usually contains complex structures. Locally weighted learning (LWL) is often used to solve nonlinear relations. To build an accurate model with LWL, the similarity needs to be properly defined.
3. To cope with noise interference, process variables are reasonably treated as random variables to construct a model in a probability frame.

Main idea

1. To extract latent information in complex process variables, a supervised Bayesian network (SBN) is proposed.

2. To improve the performance of soft sensing in non-stationary systems, a double-layer similarity measurement strategy based on latent variables is proposed.

Method

1. The node *T* in the SBN helps take quality-related information into the hidden layer construction

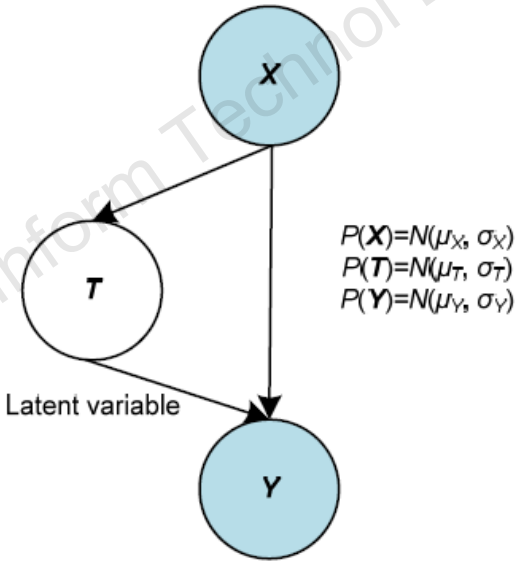


Fig. 3 Structure of a supervised Bayesian network with latent variables

Method (Cont'd)

From the structure of the SBN proposed above, latent variables can be extracted from process variables and quality variables of the training samples.

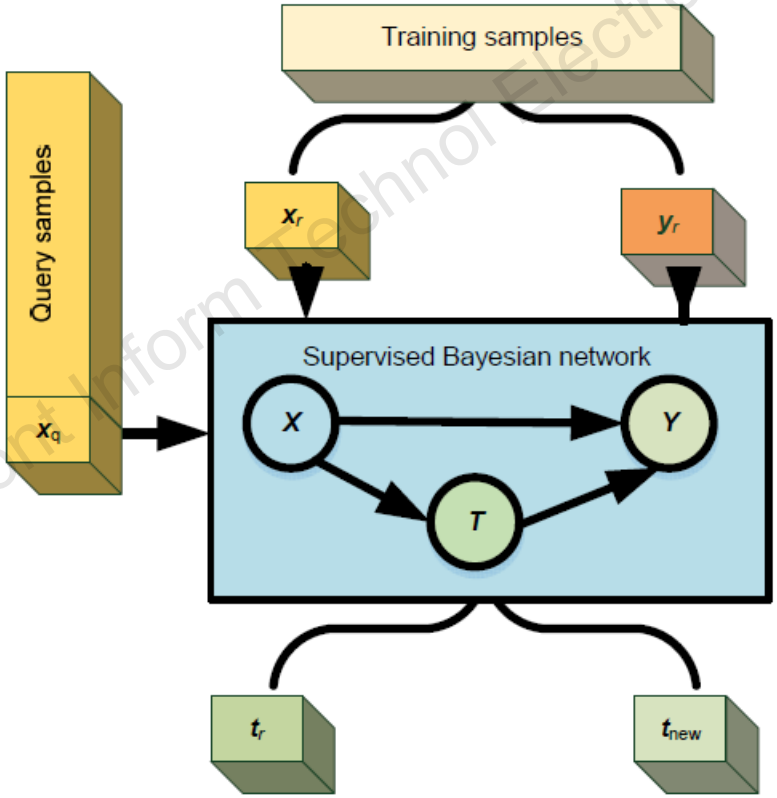


Fig. 4 Schematic of latent variable calculation

Method (Cont'd)

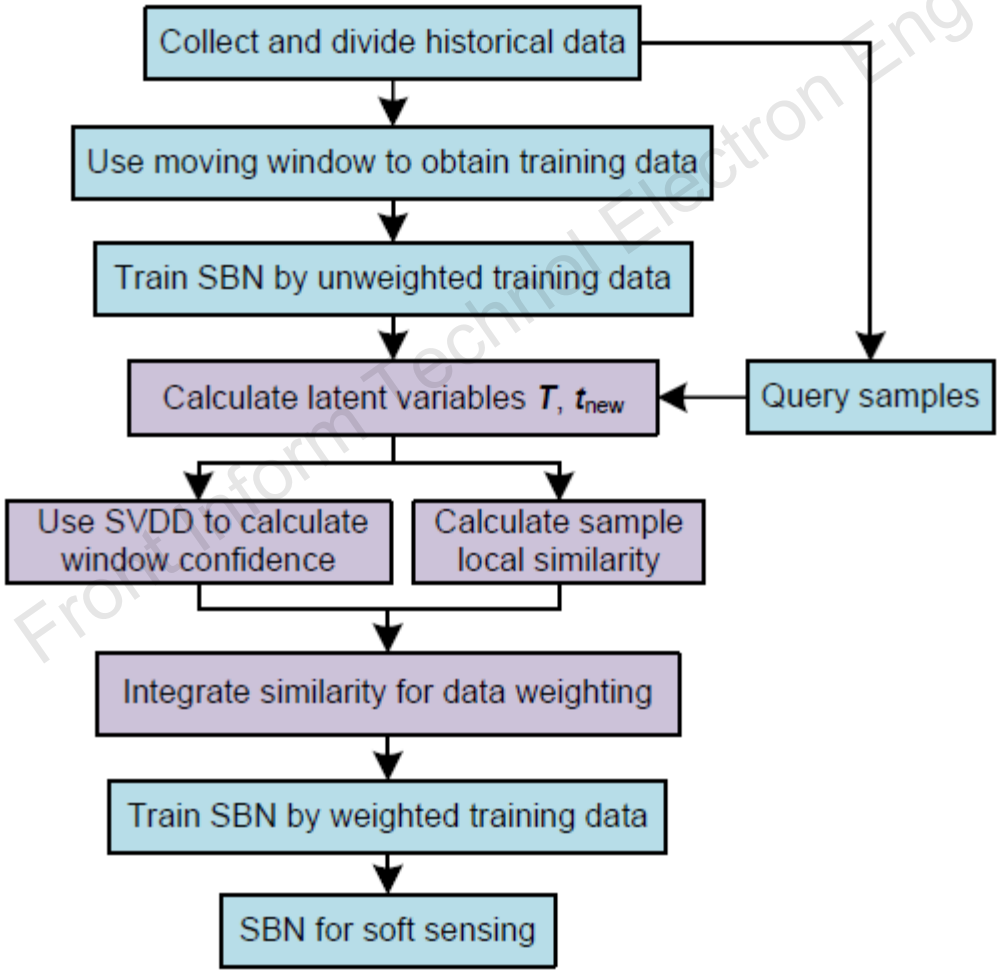
2. This structure is composed of three major parts, including sample local similarity, window confidence, and integration similarity.

Table 1 Procedure of similarity measurement

Name	Function
Sample local similarity	The Euclidean distance between the query sample and each training sample
Window confidence	The similarity of the query sample to the whole of the window by SVDD
Integration similarity	Used to assign weights to training data

Method (Cont'd)

3. The main steps of LWSBN for quality prediction



Major results

The proposed LWSBN is tested using a numerical example and a real industrial process. For comparison, the LWBN and LWLR mentioned above and a locally weighted supervised latent factor analysis model (LWSLFA) are also tested.

Major results (Cont'd)

1. Numerical example

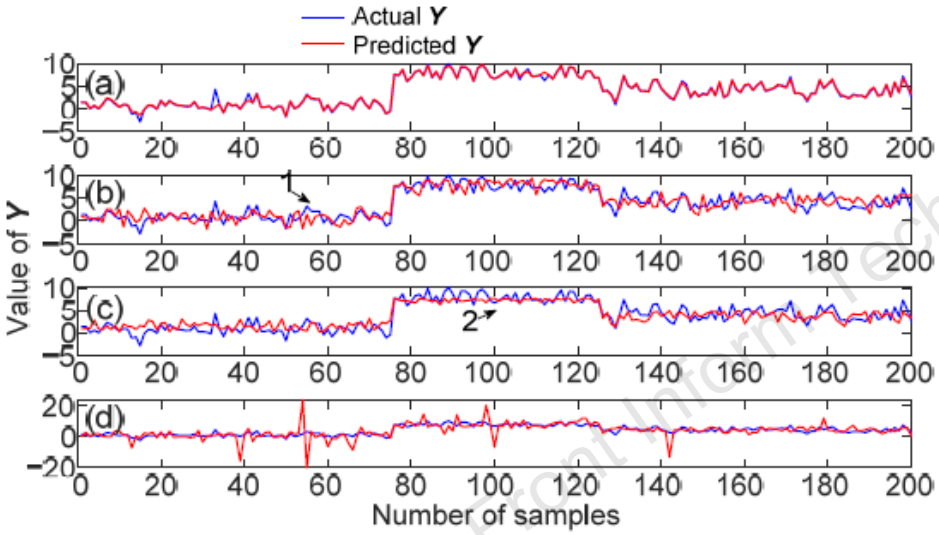


Fig. 8 Predicted *Y* of the four methods in a numerical example: (a) LWSBN; (b) LWBN; (c) LWSLFA; (d) LWLR

Table 2 Prediction of precision of the four methods in the numerical example

Index	Precision			
	LWSBN	LWBN	LWSLFA	LWLR
RMSE	0.3862	0.5225	0.9361	**0.1057**
R^2	**0.9837**	0.9712	0.9021	0.4640

RMSE: root mean squared error. The best results are in bold

Major results (Cont'd)

2. Industrial application: debutanizer column

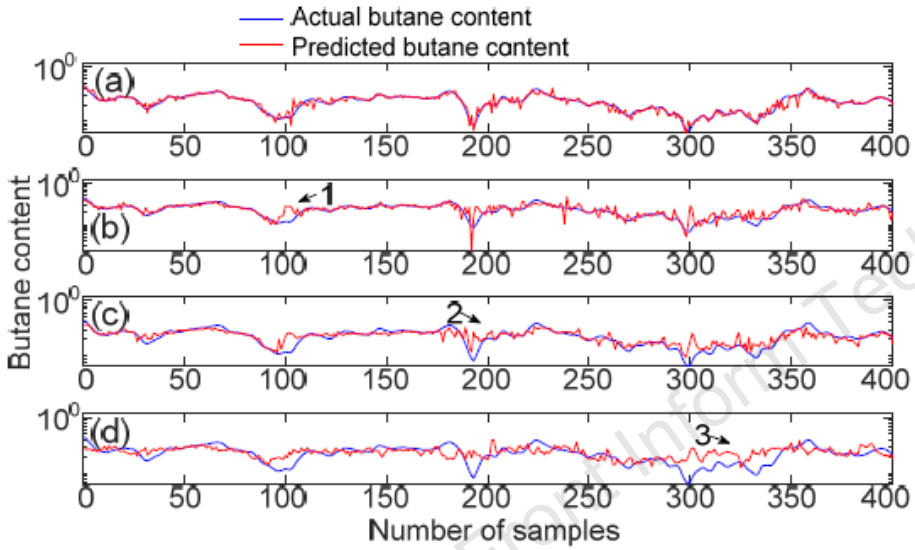


Fig. 11 Prediction results in the debutanizer column: (a) LWSBN; (b) LWBN; (c) LWSLFA; (d) LWLR

Table 4 Prediction of precision of the four methods in the debutanizer column

Index	Precision			
	LWSBN	LWBN	LWSLFA	LWLR
RMSE	**0.0321**	0.0525	0.0438	0.0936
R^2	**0.9506**	0.8679	0.9077	0.5790

RMSE: root mean squared error. The best results are in bold

Conclusions

1. Different from previous BNs, latent variables were extracted based on the SBN and a corresponding double-layer similarity calculation strategy has been proposed where the integration of local sample similarity and window confidence helps capture local and global similarity information.

2. To test the validity of this method, a numerical example and a real case were used where the prediction accuracy and reliability of the proposed method have been emphasized against previous LWBN, LWLR, and LWSLFA. The similarity measurement method involved in this paper can further be extended to other BNs.