Adaptive Soft-sensor for Sudden Changes in Process Characteristics Based on Transfer Learning

Kaito Katayama,a Kazuki Yamamoto,b Koichi Fujiwaraa\*

aDepartment of Materials Process Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan

bAGC Inc., 1-1, Suehiro-cho, Tsurumi-ku, Yokohama, 230-0045, Japan

fujiwara.koichi@hps.material.nagoya-u.ac.jp

Abstract

Soft-sensors have been widely utilized in various manufacturing processes for predicting important process variables, such as product quality or variables relating to process safety, from other easily measured variables when important process variables cannot be measured online. Highly adequate soft-sensors are necessary for efficient and safe process operation; however, when process characteristics are altered due to a variety of factors, the performance of soft-sensors may deteriorate. Because it is burdensome to update or re-build soft sensors frequently, they should be adapted to such changes automatically. Thus, automatic soft-sensor updating methods for adapting to current process characteristics, such as Just-In-Time (JIT) modelling, have been proposed. Transfer learning (TL) has also been used for updating soft-sensors. The soft-sensors updating methods based on TL have been utilized for adapting quickly to changes in process characteristics, in which operation data collected prior to changes in process characteristics are used as the source domain data, and the current process data are regarded as the target domain data. However, previous methods using TL cannot be applied to sudden and unpredictable changes in process characteristics because TL needs information about the exact timings of characteristic changes.

This study proposes a new adaptive soft-sensor updating technique called the latest sample targeting FEDA (LST-FEDA). In the proposed method, a fixed number of samples close to the query in time are defined as the target domain of FEDA, which is considered a TL method, because samples measured at close sampling points may have similar characteristics. Using a JIT modelling manner with LST-FEDA, soft-sensors are constructed and target variables are predicted for every sample measurement by focusing on some latest samples that may be similar to the query as the target domain in TL. Locally-weighted partial least squares (LWPLS) can be applied for JIT modelling. The proposed method enables soft-sensors to quickly adapt to sudden changes as well as slow changes in process characteristics. The proposed method was applied to the operation data of a vinyl acetate monomer (VAM) process, which showed that the proposed method could maintain its prediction performance even after malfunction occurrence. The prediction performance of the proposed method after the malfunction improved the correlation coefficient by 6.1 % and the root means squared errors (RMSE) by 32.9% on average compared with the previous method. In conclusion, the proposed LST-FEDA can potentially contribute to efficient process operation and reduce the maintenance or updating of soft-sensors in the practical stage.

**Keywords**: Soft-sensor, Just-In-Time modelling, Transfer learning, VAM process

* 1. Introduction

In manufacturing processes, various types of hardware sensors have been used to measure process variables; however, product quality or other important variables often have to be measured with periodic offline analysis because some variables, such as product composition, cannot be measured online. It takes some time to perform offline analysis, resulting in delays in process control.

Soft-sensors have been widely used in various manufacturing processes to address this issue (Kadlec *et al.*, 2011). A general soft-sensor uses a regression model predicting a key variable that is difficult to analyze online with hard-sensors from process variables easy to measure online. By inputting a newly measured sample into the constructed soft-sensor, the objective variable is predicted in real-time, which contributes to realizing stable and efficient process operation. However, as process characteristics change due to the aging of manufacturing facilities or periodical process maintenance, the prediction performance of the soft-sensor may deteriorate, leading to declining operation efficacy and safety (Kano and Fujiwara, 2013). It is important for the soft-sensor to rapidly adapt to the current process condition for effective and safe process operation.

To address this issue, an adaptive modelling method, Just-In-Time (JIT) modelling was developed (Atkeson *et al.*, 1997). JIT modelling can handle sudden changes in process characteristics as well as gradual changes by utilizing past process operation information; however, it does not always adapt to abrupt changes when the new process condition is completely unknown, for example, facility malfunction.

Recently, the application of transfer learning (TL) to soft-sensor adaption has been proposed. TL is a machine learning method that transfers information from an experienced similar problem (source domain) to a new problem (target domain) to make the problem easier to solve (Pan and Yang, 2010). Yamada *et al.* (2022) applied a TL technique called frustratingly easy domain adaptation (FEDA) (Daumé III, 2007) to construct soft-sensors for process manufacturing products with multiple grades. They reported achieving accurate predictions of the quality of the target grade product. However, it is difficult to apply conventional TL-based methods to abrupt changes in process characteristics because the onset of such changes should be identified in order to transfer information appropriately.

In order to solve these problems of JIT-based and TL-based methods, this study proposes a new TL method, referred to as latest sample targeting FEDA (LST-FEDA), for JIT modelling. Because it is expected that samples measured at close measurement time have similar characteristics to each other in manufacturing processes, LST-FEDA designates a fixed number of recent samples as the target domain of TL and other past samples as the source domain, which is expected to select suitable samples for soft-sensor update in JIT modelling. Whenever new samples are acquired, the data in the target domain are refreshed by LST-FEDA, and the soft-sensor is reconstructed following the JIT modelling manner. By applying the proposed LST-FEDA to JIT modelling, the soft-sensor can cope handle gradual changes in process characteristics over time, as well as abrupt changes even when there is no historical data similar to the current process condition.

In this study, the usefulness of the proposed method is demonstrated through its application to simulation data of a vinyl acetate monomer (VAM) production process (Machida *et al.*, 2016) generated by a process simulator provided by Omega Simulation Co., Ltd (https://www.omegasim.co.jp/, accessed 2023/08/05).

* 1. Methods
		1. Frustratingly easy domain adaptation (FEDA)

FEDA is a domain adaptation learning method, which can be described by a simple expansion of an input variable space (Daumé III, 2007). In soft-sensor design, the source and target domains are operation data before and after the change in process characteristics. FEDA extends the source domain data $X\_{S}\in R^{N\_{S}×M}$ and target domain data $X\_{T}\in R^{N\_{T}×M}$ as follows:

|  |  |
| --- | --- |
| $$Φ\left(X\_{S}\right)=\left[\begin{matrix}X\_{S}&X\_{S}&0\end{matrix}\right]$$ | (1) |
| $$Φ\left(X\_{T}\right)=\left[\begin{matrix}X\_{T}&0&X\_{T}\end{matrix}\right]$$ | (2) |

where $N\_{S}$ and $N\_{T}$ are the numbers of source and target domain samples, respectively. $0$ denotes a zero matrix whose size is the same as $X\_{S}$ or $X\_{T}$. The common input variable space can be trained from $X\_{S}$ and $X\_{T}$ while spaces specific to the source and target domains be from $0$ and $X\_{S}$ or $0$ and $X\_{T}$, respectively.

While process characteristics gradually change due to aging, JIT modelling is expected to cope with such slow changes. However, process characteristics sometimes suddenly alter by process maintenance or malfunction, with which JIT modelling cannot always cope. Since the former is scheduled, specifying the source and target domains in TL is easy, and FEDA can be used to adapt soft-sensors to the process condition after maintenance. On the other hand, the latter situation is usually unpredictable, and identifying its precise onset is not easy. It is difficult to apply FEDA for such cases because appropriate information transferring requires a proper definition of the source and target domains.

* + 1. Latest Sample Targeting FEDA (LST-FEDA)

It can be assumed that samples close to each other in time have similar characteristics to each other except for abrupt changes in process characteristics. The latest adjacent samples for model adaptation can be always selected in JIT modelling while process characteristics gradually change. When abrupt changes in process characteristics occur, historical samples similar to the current condition should be used for model update.

According to the above consideration, this study modifies FEDA for JIT modelling and extends the input matrix $X\in R^{N×M}$ as follows:

|  |  |
| --- | --- |
| $$Φ\left(X\right)=\left[\begin{matrix}X\_{past}&X\_{past}&0\\X\_{new}&0&X\_{new}\end{matrix}\right]$$ | (3) |

where $N$ and $M$ are the numbers of stored samples and measured variables, respectively. $X\_{new} \in R^{k×M}$ is the latest *k* samples in $X$ as the target domain, and $X\_{past} \in R^{(N-k)×M}$ is the set difference of $X$ and $X\_{new}$. That is, this modification constraints that the target domain must be the latest $k$ samples only. Following $Φ\left(X\right)$, the query $ x\_{q}\in R^{M}$ needs to be also modified as follows:

|  |  |
| --- | --- |
| $$Φ\left(x\_{q}\right)=\left[\begin{matrix}x\_{q}^{⊤}&0^{⊤}&x\_{q}^{⊤}\end{matrix}\right]^{⊤}$$ | (4) |

Using $Φ\left(X\right)$ and $Φ\left(x\_{q}\right)$, a soft-sensor is updated with JIT modelling manner. This formula is referred to as LST-FEDA.

* 1. Case study

To evaluate the performance of the proposed LST-FEDA with JIT modelling, we used simulation process data generated from a vinyl acetate monomer (VAM) process model. Forty hours of VAM process operation data were generated, which contained a situation in which a sudden malfunction occurred in the process. At the start of the measurement, the process was operating in a steady state, and 33 hours after the start of the measurement, process malfunctions occurred. The first 30 hours of data were used for training and the remaining 10 hours for test. To verify the prediction performance of the soft-sensors for various types of malfunctions, test data containing five different types of malfunctions were prepared. MAL1 and MAL2 are malfunctions that reduce the feed composition of the feedstock ethylene and acetic acid, respectively; MAL3 and MAL4 are the malfunctions that reduce the feed pressure of the feedstock Ethylene and oxygen, respectively; MAL5 is a malfunction that decreases reaction activity of the reactor due to degradation and sintering of the catalyst bed. The 65 process variables of the VAM process measured every 10 seconds were adopted as input variables of a soft-sensor, and the output to be predicted by the soft-sensor was the mass percent concentration of VAM in the product, which is measured every 30 minutes offline.

We adopted locally-weighted partial least squares (LWPLS) (Kim *et al.*, 2013) as a JIT modeling method, and constructed soft-sensors with the proposed LST-FEDA. In addition, LWPLS-based soft-sensors with and without the conventional FEDA were constructed for comparison However, it is important to note that it is difficult to use the conventional FEDA in real processes when an unpredicted malfunction occurs. Thus, it was assumed that the occurrence of the unpredicted malfunctions was observed for comparison, and the conventional FEDA was applied only when the malfunctions occurred. On the other hand, new information was always transferred in the proposed LST-FEDA. The number of the target domain samples in the proposed LST-FEDA, the number of latent variables $Z$, and the localization parameter $φ$ in LWPLS have to be determined appropriately as hyper-parameters. In this case study, they were tuned through three-fold time-series cross-validation (S. Arlot and A. Celisse, 2010). In this study, $Z=3$ and $φ=2$ were selected for LWPLS and LWPLS with FEDA, and $k=15$, $Z=35,$ and $φ=0.25$ for LWPLS with LST-FEDA.

The prediction results for the test datasets of MAL1 and MAL2 are illustrated in Figures 1 and 2. The vertical dashed line at 180 min denotes the onset of the malfunction. The prediction by any soft-sensor fluctuated significantly shortly after the malfunction occurrence, which lasted for approximately one hour. However, the errors between the measurement and the prediction by LST-FEDA were the smallest of the three soft-sensors.

Table 1 summarizes the prediction performance of the constructed soft-sensors for five test datasets including malfunctions. In order to account for the differences in behaviors after the malfunction, the prediction performances were evaluated in three periods: 0-600



Figure 1: Prediction results of the soft-sensors in VAM process with MAL1



Figure 2: Prediction results of the soft-sensors in VAM process with MAL2

Table 1: Prediction performances in VAM process

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Malfunction | Model | 0-600min(the whole period) | 180-240min(shortly after malfunction) | 240-600min(after malfunction) |
| $$R$$ | RMSE($×10^{-4}$) | $$R$$ | RMSE($×10^{-4}$) | $$R$$ | RMSE($×10^{-4}$) |
| MAL1 | - | 0.49 | 23.6 | 0.67 | 74.6 | 1.00 | 0.74 |
| FEDA | 0.75 | 8.90 | 0.56 | 28.0 | 0.99 | 0.96 |
| LST-FEDA | **0.99** | **1.22** | **0.91** | **3.49** | **1.00** | **0.44** |
| MAL2 | - | 0.59 | 3.27 | -0.51 | 9.87 | 0.90 | 1.19 |
| FEDA | **0.76** | **2.06** | -0.45 | **5.86** | 0.95 | 1.07 |
| LST-FEDA | 0.73 | 2.28 | **-0.45** | 6.86 | **0.96** | **0.72** |
| MAL3 | - | 0.14 | 8.18 | -0.55 | 25.8 | 0.80 | 0.81 |
| FEDA | 0.33 | 4.11 | -0.53 | 12.7 | 0.74 | 0.96 |
| LST-FEDA | **0.60** | **2.13** | **-0.35** | **6.48** | **0.92** | **0.54** |
| MAL4 | - | 0.95 | 0.84 | 0.92 | 1.67 | 0.96 | 0.68 |
| FEDA | 0.94 | 0.79 | **0.97** | **0.36** | 0.93 | 0.88 |
| LST-FEDA | **0.96** | **0.66** | 0.93 | 0.48 | **0.96** | **0.66** |
| MAL5 | - | 0.92 | 17.5 | -0.27 | 11.7 | 0.92 | 22.1 |
| FEDA | 0.93 | 17.3 | -0.57 | 9.83 | 0.92 | 22.0 |
| LST-FEDA | **0.93** | **15.8** | **0.19** | **9.08** | **0.93** | **20.1** |

min (whole period), 180-240 min (shortly after the malfunction), and 240-600 min (after the malfunction). In this study, the correlation coefficient ($R$) and the root mean squared error (RMSE) were adopted as evaluation metrics. The best values in each test data are highlighted in bold in Table 1.

The proposed LST-FEDA achieved the best performance in almost all malfunction scenarios except for the whole period in MAL2. According to Figure 1, LST-FEDA successfully suppressed significant fluctuation of the prediction shortly after the malfunction in comparison with other two methods. Moreover, LST-FEDA could appropriately follow the measurement one hour after the malfunction occurrence.

In MAL2, RMSE of LST-FEDA for the whole period was slightly worse than the conventional FEDA. According to Figure 2, LST-FEDA was not able to avoid incorrect prediction deterioration shortly after MAL2 occurred; however, the conventional FEDA also did not make accurate predictions during this period, which suggests that it is impossible for any method to appropriately cope with MAL2 shortly after its occurrence. In regard to the prediction performance after the malfunction, the soft-sensors with the proposed LST-FEDA achieved the highest performance in all malfunction cases.

In summary, the proposed LST-FEDA improved $R$ by 22.4 % and RMSE by 29.8 % on average for the whole period, and $R$ by 6.1 % and RMSE by 32.9 % on average for the after malfunction period in comparison with the conventional FEDA. These results confirm that the proposed LST-FEDA can adapt soft-sensors more quickly and accurately to changes in process characteristics due to various process malfunctions suddenly occurring.

* 1. Conclusion

In this study, we propose a new TL method for JIT modeling, referred to as LST-FEDA. The proposed LST-FEDA is an extension of FEDA defining a fixed number of recent samples close to a query in time as the target domain data and other data as the source domain data. By updating the target domain data for every new sample collection, it is possible to construct soft-sensors emphasizing characteristics of recent data more than past data. Thus, soft-sensors constructed with LST-FEDA quickly adapted to sudden changes in process characteristics that cannot be addressed by conventional TL methods.

To verify the performance of LST-FEDA, the proposed LST-FEDA was applied to the operation data of the VAM process. This result clearly showed that LST-FEDA achieved the best prediction performance in almost all malfunction cases in comparison with the conventional methods. The proposed method will contribute to efficient and safe process operation in the future.

In future works, we plan to apply LST-FEDA to real process data and assess its performance comprehensively. Additionally, we aim to refine LST-FEDA further by allowing for more flexible determination of the number of target domain samples.

References

P. Kadlec, R. Grbić, and B. Gabrys, 2011, Review of adaptation mechanisms for data-driven soft sensors, Computers & Chemical Engineering, 35, 1, pp.1-24

M. Kano and K. Fujiwara, 2013, Virtual sensing technology in process industries: trends and challenges revealed by recent industrial applications, Journal of chemical engineering of Japan, 46, 1, pp.1-17

C. G. Atkeson, A. W. Moore, and S. Schaal, 1997, Locally Weighted Learning, Springer Netherlands, Lazy Learning, pp.11-73

S. J. Pan and Q. Yang, 2010, A survey on transfer learning, IEEE Transactions on knowledge and data engineering, 22, 10, pp.1345-1359

N. Yamada and H. Kaneko, 2022, Adaptive soft sensor based on transfer learning and ensemble learning for multiple process states, Analytical Science Advances, 3, 5-6, pp.205-211

H. Daumé III, 2007, Frustratingly Easy Domain Adaptation, ACL Anthology, pp.256-263

Y. Machida, S. Ootakara, H. Seki, Y. Hashimoto, M. Kano, Y. Miyake, N. Anzai, M. Sawai, T. Katsuno, and T. Omata, 2016, Vinyl Acetate Monomer (VAM) Plant Model: A New Benchmark Problem for Control and Operation Study, IFAC-PapersOnLine, 49, 7, pp.533-538

Omega Simulation Co. Lid., https://www.omegasim.co.jp, (accessed 2023/08/05)

S. Kim and R. Okajima and M. Kano and S. Hasebe, 2013, Development of soft-sensor using locally weighted PLS with adaptive similarity measure, Chemometrics and Intelligent Laboratory Systems, 124, pp.43-49

S. Arlot and A. Celisse, 2010, A survey of cross-validation procedures for model selection, Statistics Surveys, 4, pp.40-79