# Dioxin Soft Measuring Method in Municipal Solid Waste Incineration Based on Virtual Sample Generation

Jian Tang, Junfei Qiao, Ke Gu, Aijun Yan
Faculty of Information Technology
Beijing University of Technology
Beijing, China
freeflytang@bjut.edu.cn

*Abstract*—**Municipal solid waste incineration (MSWI) becomes the most popular technique to enhance environment protection. This process produces one of the most toxic chemicals in the world, i.e., polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs). The dioxin (DXN) production should be restricted rigidly by using operation optimization and control of MSWI process based on present industrial devices. However, it is difficult to realize the on-line real-time continuous measuring of DXN duo to the complexity formation mechanism and high-cost long-time off-line detection approach. In this paper, a soft measuring method based on virtual sample generation (VSG) is used to address this problem at the first time. A few numbers of true training samples are used to produce virtual training samples based on feasibility-based programming (FBP) model using selective ensemble kernel partial least squares (SENKPLS) and prior knowledge. Simulation result based on dataset in reference [31] for a HL MSWI process shows effectiveness of the proposed method.**

***Keywords—Municipal solid waste incineration (MSWI); dioxins (DXN); virtual sample generation (VSG); selective ensemble kernel partial least squares (SENKPLS); soft measuring***

## I. INTRODUCTION

Normally, polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) are termed as Dioxins (DXN) [1]. Municipal solid waste incineration (MSWI) becomes one of the most production sources of DXN in most of large cities. Generally, two process phases in MSWI can form DXN, which are homogeneous in the temperature range of 500-800 °C and heterogeneous in the temperature range of 200-400 °C. As the complicated formation mechanisms and difficulty-to-measure character of PCDD/Fs, a universal result can be not be obtained until now [2,3]. More researchers are focus on heterogeneous PCDD/Fs because of the discovery of heterogeneous formation mechanisms [4,5]. More advanced industrial dust removal equipments are used to reduce DXN emission [6,7]. Recently, studies of the formation mechanisms of homogeneous PCDD/Fs are also increased [8]. However, formation mechanisms and factors influence of DXN are still remain debated [1]. Studies about DXN indicators are also developed in terms of replace measuring method of DXN [9,10]. However, the online monitoring indicator in terms of high accurately, low cost, simplicity and on-line real-time are still far more realization [11,12]. Moreover, all these studies of measuring DXN are not in prospective of operational optimization and control of MSMI process. The main bottleneck to realize of the above objective is that the real-time continuous detection of DXN cannot be reached. Thus, it is necessary to find an alternative approach.

Some difficulty-to-measure process parameters can be measured based on data-driven soft measuring techniques [13,14]. These soft measuring models have been used in broad fields in terms of their inferential estimation capability [15]. The most commonly used methods include artificial neural networks (ANN) and support vector machines (SVM) [16,17]. Normally, which type modeling algorithm to be used depends on characteristics of the being measured process parameters in different application backgrounds. Projection to latent structure or partial least squares (PLS) and its kernel version, i.e., KPLS, can construct prediction model with a small number latent variables [18]. Thus, they are suitable to address co-linear data in industrial process [19]. Normally, selective ensemble (SEN) modeling method can use multiple sub-models to obtain better generalization performance than that of the single-one. Aim at selective information fusion problem of multi-source features, brand and bound (BB)-based SEN(BBSEN) approach is proposed, which is used to model frequency spectra data [14, 20]. SEN model based on "manipulating training samples" strategy is proposed to selective fusion different training samples [21]. Fatherly, a uniform SEN-model by selective fusion multi-condition samples and multi-source features are proposed in [22] to overcome shortcomings of [20] and [21]. However, these modeling methods cannot solve the rare samples-based modeling problem.

Based on prior knowledge, the concept of creating virtual samples, i.e., virtual sample generation (VSG) method is proposed to improve image recognition accuracy [23]. Then, VSG-based multi-layer perception networks and distributed network are proposed to construct regression model [24,25]. Recently, some intelligent optimization algorithms, such as GA and PSO, are used to in VSG [26,27,28,29]. Most of these approaches are based on traditional single-model to produce virtual samples. In our former studies, SEN-based

VSG method by using multi-scale mechanical vibration and acoustic frequency spectra is proposed [21]. However, the above SEN methods are all used to high dimensional spectral data with same measurement unit. Therefore, it is necessary to build DXN soft measuring model based on VSG. Moreover, it is necessary to find metric to select the number of virtual samples.

Motivated by the above problem, a soft measuring method based on VSG is used to model dioxin in MSWI at the first time. At first, small number true training samples are used to build feasibility-based programming (FBP) model using selective ensemble kernel partial least squares (SENKPLS) algorithm. Then, linear interpolation and prior knowledge are combined to produce virtual sample inputs, and these inputs are fed into FBP to produce virtual sample outputs. A simple statistical similarity criterion is proposed to select the a group of virtual samples. Finally, these true and virtual samples are mixed to construct soft measuring model of DXN. Simulation result based on dataset in References [30,31] for a PH MSWI plant shows effectiveness of the proposed method.

## II. DXN Soft Measuring Stratedy of MSWI

The proposed DXN soft measuring strategy consists of two phases, which are the off-line training phase and on-line measuring phase. It is shown in Fig. 1.

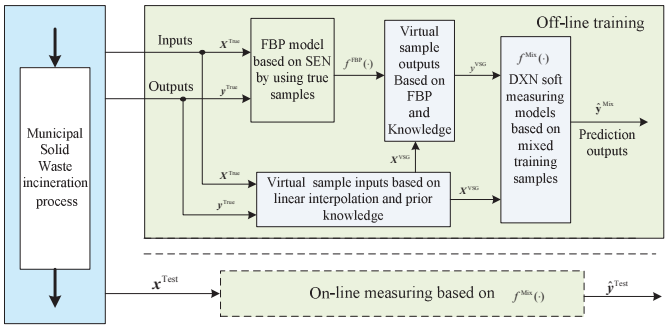


Fig. 1 DXN soft measuring strategy of MSWI

Fig. 1 shows that the off-line training phase consists of at least four modules. Their functions are shown as follows:

(1)FBP model based on SEN by using true samples: In the MSWI process, limit true samples can obtain based on the off-line analyzer with high cost and long time. Use the limited true input and output training sample, FBP model is build based on SENKPLS by using "manipulating training samples" ensemble construction strategy. It is used to produce virtual sample outputs based on virtual sample inputs.

(2)Virtual sample inputs based on linear interpolation and prior knowledge: To the practical MSWI process, some information, such as physical meaning and distribution of the true training samples, is known. That is to say, the intervals among different input variables of the actual training samples are prior knowledge to us. Thus, the simple average linear interpolation method is used to obtain virtual sample inputs in this study.

(3)Virtual sample outputs based on FBP and knowledge: In nature, this is a prediction output process based on FBP model. With different linear interpolation parameter, a series of virtual sample inputs are generated. As a results, different groups' virtual sample outputs are generated. A simple statistical similarity criterion is used to select the appropriate virtual samples group. How to select the optimized virtual samples number in mixed sample-based soft measuring modeling is a open issue.

(4)DXN soft measuring models based on mixed training samples: It is a similar process same as that of FBP model construction. Moreover, learning algorithms that difference from FBP can also be used to construct DXN model.

## III. Relization of the Proposed Stratedy

### A. Off-line Training Phase

The original true training samples are denoted as $\boldsymbol{X}_{\text{ori}}^{\text{True}} \in \Re^{k\times P}$ and $\boldsymbol{y}_{\text{ori}}^{\text{True}} \in \Re^{k\times 1}$, which are pre-processed with:

$$\boldsymbol{X}^{\text{True}} = (\boldsymbol{X}_{\text{ori}}^{\text{True}} - \boldsymbol{1}^k \boldsymbol{u}^{\text{T}}) \cdot \sum^{-1} \tag{1}$$

where, $\boldsymbol{u}$ and $\sum$ represent mean and variance of the true training data. The same process also is done on $\boldsymbol{y}_{\text{ori}}^{\text{True}}$ to obtain $\boldsymbol{y}^{\text{True}}$. The construction process of FBP can be represented as:

$$\left.\begin{matrix}\boldsymbol{X}^{\text{True}}, \boldsymbol{y}^{\text{True}} \\ M_{\text{para}}\end{matrix}\right\} \xrightarrow{\text{SENKPLS}} f^{\text{FBP}}(\cdot) \tag{2}$$

where, $M_{\text{para}}$ represents the modeling parameters of SENKPLS. Details of SENKPLS are shown in [21], which are omitted in here.

For convenience, the original true training samples can be re-denoted as:

$$\begin{aligned}\mathbf{X}_{\text{ori}}^{\text{True}} &= \{\mathbf{x}_l\}_{l=1}^k \\ &= \{[x_1,\cdots,x_p,\cdots,x_P]_l\}_{l=1}^k = \{[x_l^p]_{p=1}^P\}_{l=1}^k\end{aligned} \tag{3}$$

where, $x_l^p$ represents the $p$th feature variable of the $l$th sample. We partition the interval between two adjacent feature variables $x_l^p$ and $x_{l+1}^p$ into $N_{\text{VSG}}$ sub-parts with simple linear interpolation. Then, virtual sample inputs with number $(N_{\text{VSG}}-1)$ are obtained. These virtual sample inputs can be calculated with:

$$(x_{l'}^{\text{p}})^{\text{VS}} = x_l^p + \frac{\left(x_{l+1}^p - x_l^p\right)l'}{N_{\text{VSG}}} \qquad 1 \le l' < N_{\text{VSG}} \tag{4}$$

where $N_{\text{VSG}} \ge 2$, it is the virtual sample production parameter, and $x_{l'}^{\text{VS}}$ is the $l'$th virtual sample input of the $p$th feature variable. With the same method, the $l'$th virtual sample input of all feature variables be denoted as:

$$\mathbf{x}_{l'}^{\text{VS}} = [(x_{l'}^p)^{\text{VS}}]_{p=1}^P \tag{5}$$

There are total $k_{\mathrm{VSG}}$ intervals in the true training samples. The number of virtual sample input is represented as:

$$k' = (N_{\mathrm{VSG}} - 1) \times k_{\mathrm{VSG}}, \qquad 2 \le N_{\mathrm{VSG}} . \tag{6}$$

Thus, there are total $(N_{\mathrm{VSG}} - 1)$ groups virtual sample inputs.

Based on the constructed FBP model, virtual sample outputs based on $\mathbf{X}^{\mathrm{VSG}} = \{\mathbf{x}_{l'}^{\mathrm{VS}}\}_{l'}^{k'}$ is calculated as:

$$\{y_{l'}^{\mathrm{VSG}}\}_{l'}^{k'} = f^{\mathrm{FBP}}\left(\{\mathbf{x}_{l'}^{\mathrm{VS}}\}_{l'}^{k'}\right) \tag{7}$$

By making the re-scaled process with:

$$\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}} = \mathrm{Recale}(\{y_{l'}^{\mathrm{VSG}}\}_{l'=1}^{k'}) \tag{8}$$

the virtual sample outputs is obtained. Finally, there are total $(N_{\mathrm{VSG}} - 1)$ groups virtual sample outputs with number $k_{\mathrm{VSG}}$, $2k_{\mathrm{VSG}}$, …, $(N_{\mathrm{VSG}} - 1) \times k_{\mathrm{VSG}}$. In order to select an appropriate group, the normal statistical indices, such as maximum, average and minimum values, are used to judge the similarity between different group virtual sample outputs and the true training samples. It can be represented as:

$$\mathrm{Sim}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}}, \mathbf{y}_{\mathrm{ori}}^{\mathrm{True}}) = \frac{\mathrm{Max}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}})}{\mathrm{Max}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{True}})} + \frac{\mathrm{Mean}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}})}{\mathrm{Mean}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{True}})} + \frac{\mathrm{Min}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}})}{\mathrm{Min}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{True}})} \tag{9}$$

Based on (8), the group virtual sample outputs with the $\mathrm{Sim}(\mathbf{y}_{\mathrm{ori}}^{\mathrm{VSG}}, \mathbf{y}_{\mathrm{ori}}^{\mathrm{True}})$ value closed to three are selected, which is denoted as $\mathbf{y}^{\mathrm{VSG}}$ for convenience. Then, all the virtual samples can be represented as:

$$\begin{aligned} S^{\mathrm{VSG}} &= \{\mathbf{X}^{\mathrm{VSG}}, \mathbf{y}^{\mathrm{VSG}}\} = \{\{\mathbf{x}_{l'}^{\mathrm{VS}}\}_{l'=1}^{k'}, \{y_{l'}^{\mathrm{VSG}}\}_{l'=1}^{k'}\} \\ &= \{S_{l'}^{\mathrm{VSG}}\}_{l'=1}^{k'} \end{aligned} \tag{10}$$

By combining the original true training samples and the virtual samples, we can obtain the mixed training samples $S^{\mathrm{Mix}}$. It can be re-written as:

$$\begin{aligned} S^{\mathrm{Mix}} &= \{S^{\mathrm{True}}; S^{\mathrm{VSG}}\} \\ &= \{\mathbf{X}^{\mathrm{True}}, \mathbf{y}^{\mathrm{True}}; \mathbf{X}^{\mathrm{VSG}}, \mathbf{y}^{\mathrm{VSG}}\} \\ &= \{\{\mathbf{x}_l^{\mathrm{True}}\}_{l=1}^{k}, \{y_l^{\mathrm{True}}\}_{l=1}^{k}; \{\mathbf{x}_{l'}^{\mathrm{VSG}}\}_{l'=1}^{k'}, \{y_{l'}^{\mathrm{VSG}}\}_{l'=1}^{k'}\} \\ &= \{\{\mathbf{X}_{l^{mix}}^{\mathrm{mix}}\}_{l^{mix}=1}^{k+k'}, \{y_{l^{mix}}^{\mathrm{mix}}\}_{l^{mix}=1}^{k+k'}\} \\ &= \{\mathbf{X}^{\mathrm{mix}}, \mathbf{y}^{\mathrm{mix}}\} \end{aligned} \tag{11}$$

By using SENKPLS [21], the DXN soft sensor model is represented as:

$$\hat{\mathbf{y}}^{\mathrm{Mix}} = f^{\mathrm{Mix}}(\mathbf{X}^{\mathrm{mix}}) \tag{12}$$

### *B. On-line Measuring Phase*

When the new sample input $\mathbf{x}_{\mathrm{ori}}^{\mathrm{Test}}$ is obtained, the scaling operation based on old mean and variance must be done. It can be realized with:

$$\mathbf{x}^{\mathrm{Test}} = (\mathbf{x}_{\mathrm{ori}}^{\mathrm{Test}} - \mathbf{1}^k \mathbf{u}^{\mathrm{T}}) \cdot \sum^{-1} \tag{13}$$

Then, the prediction output based on the testing sample is calculated with:

$$\hat{\mathbf{y}}^{\mathrm{Test}} = f^{\mathrm{Mix}}(\mathbf{x}^{\mathrm{Test}})$$

## IV. SIMULATION RESULTS

### *A. Experimental Data*

The MSWI process is same as that of Ref [30], which is omitted in here. The experimental data under 11 different operation conditions are acquired, which are shown in Table 1. In Table I, the odd ordinal number cases are used as the true training samples and all the datasets are used as testing samples. Therefore, there are only 6 true training samples are used to produce the virtual samples.

TABLE I. EXPERIMENTAL DATA UNDER 11 OPERATION CONDITIONS

| Case | Operation parameters | | | | | | Normal pollution factors(mg/Nm$^3$) | | | | | | | | DXN (ng/Nm$^3$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Furnace Temp(℃)* | *Boiler Outlet Temp(℃)* | *Bag Outlet Temp(℃)* | *Flu (Nm$^3$/h)* | *$O_2$ (%)* | *$H_2O$ (%)* | *$CO_2$ (%)* | *$SO_2$* | *NOx* | *HCL* | *CO* | *Dust* | *HF* | *$NH_3$* | |
| 1 | 1015 | 218 | 161 | 42194 | 9.3 | 27.5 | 9.1 | 7.2 | 251 | 56.9 | 0.5 | 7.7 | 0.6 | 0.3 | 4.63 |
| 2 | 1048 | 200 | 145 | 39785 | 9.4 | 15.6 | 6.1 | 16.9 | 153.7 | 24.6 | 0.5 | 8.3 | 0.6 | 6.0 | 1.04 |
| 3 | 1030 | 201 | 145 | 43722 | 9.4 | 16.5 | 5.9 | 18.6 | 147.5 | 22.0 | 0.5 | 8.9 | 0.6 | 6.0 | 1.24 |
| 4 | 1050 | 180 | 145 | 28033 | 11.9 | 8. 5 | 5.7 | 20.8 | 43.2 | 6.8 | 0.6 | 10.6 | 0.4 | 25.2 | 0.655 |
| 5 | 1050 | 210 | 158 | 38892 | 6.0 | 36.1 | 10.6 | 6.8 | 175 | 37.8 | 6.0 | 6.8 | 0.0 | 3.2 | 2.64 |
| 6 | 955 | 225 | 161 | 37458 | 9.2 | 33.6 | 12. 4 | 8.4 | 60.1 | 25.3 | 5.4 | 7.3 | 0.0 | 5.4 | 3.52 |
| 7 | 983 | 225 | 159 | 41468 | 8.9 | 35.5 | 8.8 | 9.7 | 122.9 | 22.1 | 6.0 | 8.2 | 0.0 | 5.9 | 2.25 |
| 8 | 1006 | 214 | 159 | 37469 | 9.0 | 34.2 | 8.9 | 10.1 | 81.4 | 21.0 | 7.0 | 7.4 | 0.0 | 5.3 | 2.17 |
| 9 | 910 | 242 | 170 | 41000 | 10.0 | 38.7 | 11.0 | 1.5 | 162.0 | 51 | 5.5 | 8.5 | 12 | 11 | 6.00 |
| 10 | 925 | 240 | 166 | 46010 | 9.0 | 39.8 | 12.0 | 2.0 | 117.0 | 39.0 | 0.8 | 12.0 | 11.0 | 14.0 | 3.83 |
| 11 | 1070 | 190 | 155 | 32435 | 10.2 | 21.2 | 7.1 | 15.6 | 94.6 | 30.2 | 0.3 | 0.0 | 0.6 | 0.1 | 2.97 |

### B. FBP Model Consruction Results

In order to show characteristics of the DXN modeling data, PLS are used to analyze the true training samples. It shows that the former 3 latent variables (LVs) capture 82.36% 99.91% variance of the 14 input feature variables and the output predicted DXN. Thus, it shows that contribution of input features can be improve further with nonlinear PLS. By using KPLS with RBF kernel radius 700, contribution of the input features is improved to 93.22%. Thus, it is reasonable to build FBP model with nonlinear modeling method.

FBP-model based on SENKPLS are constructed with the following modeling parameters: RBF kernel radius 70, kernel LVs (KLVs) number 6, candidate sub-model number 20. The validation and testing datasets are same by using all the samples in Table I. The modeling process is repeated 100 times. It shows that the prediction outputs aren't stable. There are almost 20 times cannot obtain the prediction results. The minimum and maximum RMSEs of the prediction output for testing samples are 0.3163 and 2.1909, respectively.

### C. Virtual Sample Generation Results

By taking the strategy "generate virtual samples between two adjacent true samples", the number of virtual samples are 5, 10, … , 45 at $N_{VSG}$ = 2, 3, … , 10. The prediction output ranges with different values of $N_{VSG}$ are shown in Table II.

TABLE II. PREDICTION OUTPUT RANGES WITH DIFFERENT $N_{VSG}$

| | True | $N_{VSG}$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* |
| Min | 1.240 | 2.216 | 0.286 | 2.321 | 1.524 | 1.065 | 1.481 | 0.489 | 0.340 | 1.405 |
| Mean | 3.288 | 3.403 | 3.233 | 3.123 | 3.148 | 3.412 | 3.284 | 3.270 | 2.998 | 3.197 |
| Max | 6.000 | 4.978 | 7.312 | 4.707 | 5.247 | 6.463 | 5.290 | 5.393 | 5.265 | 5.787 |
| Sim(•) | 3 | | | | | | | | | |

Table 2 show that ranges of prediction output with different $N_{VSG}$ are difference. The ranges with $N_{VSG}$ =6 is close to true values and its Sim(•) value is close to 3. Thus, the virtual samples with $N_{VSG}$ =6 is selected as the virtual sample output. The outputs of the true samples and the virtual ones with $N_{VSG}$ = 6 are shown in Fig. 2.

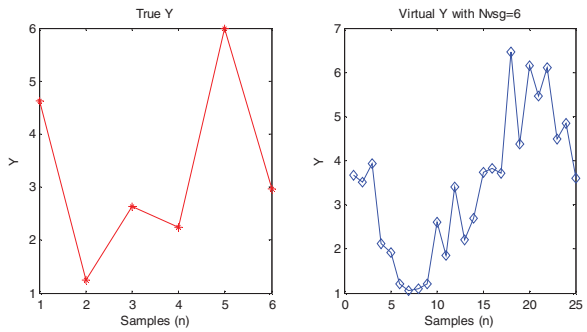


Fig. 2 True and virtual outputs at $N_{VSG}$ = 6

Fig. 2 shows that the proposed VSG is effective in terms of the shape of the true and virtual sample outputs' trend. The range of virtual samples is a little larger than that of that of the true ones. It is necessary to control it in a reasonable range. More researches would be done in the future.

### D. Soft Measuring Results Based on Mixed Samples

In the modeling phases, the mix training samples are used to construct DXN prediction model. The number of training samples is 31, in which 25 are virtual ones. Total 11 experimental data are used as testing dataset. That to say, the 6 true training samples are also used in the testing data set. Three different approaches, i.e., PLS, KPLS and SENKPLS, are used to construct DXN soft measuring model with the mixed training samples.

To PLS method, by using the leave-one-out-cross-validation (LOOCV) method, 5 latent variables (LVs) are used to construct final DXN model. Contributions of different LVs are shown in Table III.

TABLE III. CONTRIBUTIONS OF DIFFERENT LVS WITH PLS METHOD

| LV # | X-Block | | Y-Block | |
|---|---|---|---|---|
| | *This LV* | *Total* | *This LV* | *Total* |
| 1 | 48.25 | 48.25 | 71.97 | 71.97 |
| 2 | 19.03 | 67.28 | 12.51 | 84.48 |
| 3 | 14.72 | 82.00 | 0.58 | 85.06 |
| 4 | 6.45 | 88.45 | 0.72 | 85.79 |
| 5 | 11.55 | 100.00 | 0.13 | 85.92 |

Table III shows that the former 3 LVs capture 82.00% and 85.06% variables of the input and output data. It is lower than that without virtual samples. The RMSEs of the training and testing samples are 0.5942 and 0.5463, respectively. The prediction curve based on PLS is shown in Fig. 3.

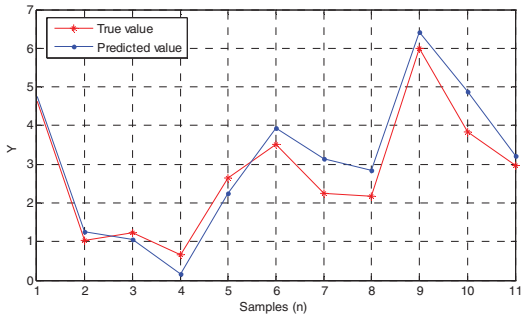


Fig. 3 Prediction curves of testing data based on PLS at $N_{VSG}$ =6

To KPLS method, the same RBF kernel radius as the FBP model is used. The number of Kernel LVs (KLVs) also is 5, which is selected by LOOCV approach. Contributions of different KLVs are shown in Table 4.

TABLE IV. CONTRIBUTIONS OF DIFFERENT LVS WITH KPLS METHOD

| LV # | X-Block | | Y-Block | |
|---|---|---|---|---|
| | *This LV* | *Total* | *This LV* | *Total* |
| 1 | 73.97 | 73.97 | 72.02 | 72.02 |
| 2 | 11.25 | 85.21 | 12.47 | 84.49 |
| 3 | 7.94 | 93.15 | 0.60 | 85.08 |
| 4 | 2.76 | 95.92 | 0.74 | 85.82 |
| 5 | 4.08 | 100.00 | 0.14 | 85.96 |

Table IV shows that the former 3KLVs capture 93.15% and 85.08% variables of the input and output data. The captured variable of DXN data is lower than that of the only true training samples. The RMSEs of the training and testing samples are 0.5933 and 0.5425, respectively. Thus, the predication performance is almost same as that of PLS method. However, with optimized selection of kernel parameters, the better results can be obtained. The prediction curve of the testing samples based on KPLS is shown in Fig. 4.

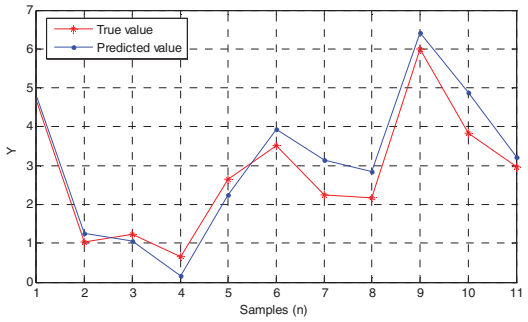


Fig. 4 Prediction curves of testing data based on KPLS at $N_{VSG}$ =6

To SENKPLS-based DXN model, the modeling process is repeated 100 times with kernel radius 70 and KLV number 6 and candidate sub-model number 20. The statistically results of different modeling method are shown in Table V.

TABLE V. STATISTICALLY RESULTS OF DXN MODEL

| Methods | Samples | Max RMSE | Mean RMSE | Min RMSE |
|---|---|---|---|---|
| PLS | Mixed training | 0.5942 | 0.5942 | 0.5942 |
| | Testing | 0.5463 | 0.5463 | 0.5463 |
| KPLS | Mixed training | 0.5933 | 0.5933 | 0.5933 |
| | Testing | 0.5425 | 0.5425 | 0.5425 |
| SENKPLS | Mixed training | 0.7036 | 0.5989 | 0.5564 |
| | Testing | 0.6492 | 0.4031 | 0.2545 |

Table V shows that the proposed method has improved the prediction ability of DXN model. Same as that of PLS and KPLS methods, the testing dataset has lower RMSE than that of training samples. Thus, PLS based method can avoid over-fitting problem effectively. The prediction curves of the training and testing datasets based on SENKPLS are shown in Fig. 4 and Fig. 5.

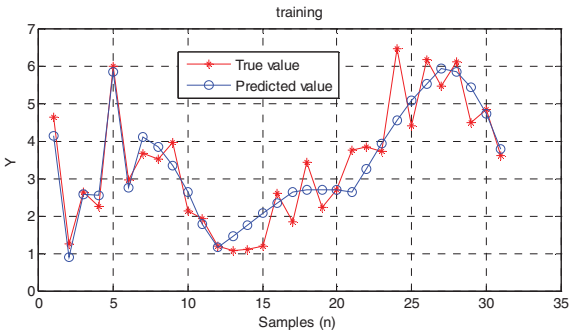


Fig. 5 Prediction curve of training data based on SENKPLS

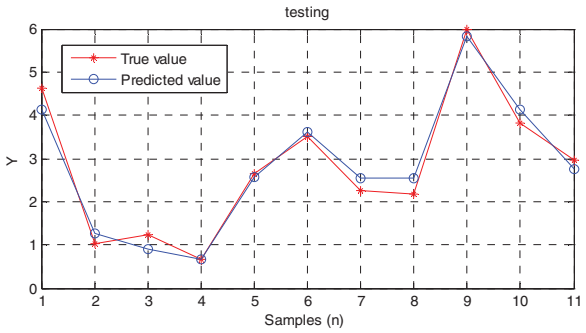


Fig. 6 Prediction curves of testing data based on SENKPLS

Fig. 5 shows that there are big prediction errors on the most of virtual samples. Fig. 6 shows that SENKPLS-based DXN model can obtain better prediction result than that of PLS/KPLS-based method.

### *E. Disscussions*

With the proposed method, the following problems should be focused: (1)How to give a reasonable range of different input feature variables of the DXN model? Except the prior knowledge, some constrain should be added to the output of FBP model; (2) How to select a more reasonable $N_{VSG}$ value? And how to modify the simple linear interpolation method to a nonlinear one based on characteristics of modeling data; (3) In the modeling phase based on mixed training samples, how to select virtual sample by considering the prediction performance.

## V. CONCLUSIONS

On-line real-time measurement of polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/Fs) can help to realize the operation optimization and control of MSWI industrial process. It has been a bottleneck due to the complex formation mechanism and high-cost long-period detection of PCDD/Fs. In this paper, a soft measuring method based on virtual sample generation (VSG) is used to address this problem at the first time. The little true training samples are used to produced virtual training samples based on feasibility-based programming (FBP) model using selective ensemble kernel partial least squares (SENKPLS)

algorithm. A new similarity-based criterion is used to determine the VSG parameter. Simulation result based on experimental shows effectiveness of the proposed method. However, there are still many open issues for VSG. More researches should be done fatherly.

## ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China (61573364, 617030879, 61640308, 61503066, 61573249), the State Key Laboratory of Synthetically Automation for Process Industries (PAL-N201504, PAL-N201605).

## REFERENCES

[1] H. Zhou, A. Meng, Y. Q. Long, Q. H. Li, and Y. G. Zhang, "A review of dioxin-related substances during municipal solid waste incineration," Waste Management,Vol.36, pp. 106-118, February 2015.

[2] B. R. Stanmore, "Modeling the formation of PCDD/F in solid waste incinerators," Chemosphere, vol. 47, pp. 565-773, May 2002.

[3] B. R. Stanmore, "Adsorption and partitioning of PCDD/F on flyash fromincinerators," Environmental Engineering Science, Vol. 19, pp. 69-78, July 2004.

[4] S. L. Alderman, G. R. Farquar, E. D. Poliakoff, and B. Dellinger, "An infrared and Xray spectroscopic study of the reactions of 2-chlorophenol, 1,2-dichlorobenzene, and chlorobenzene with model CuO/silica fly ash surfaces," Environmental Science & Technology, vol. 39, pp. 7396-7401, July, 2005.

[5] J. Vehlow, "Reduction of dioxin emissions from thermal waste treatment plants: a brief survey," Reviews in Environmental Science & Bio/technology, vol. 11, pp. 393-405, September 2012.

[6] B. Kim, S. Lee, S. Maken, H. Song, J. Park, and B. Min, "Removal characteristics of PCDDs/Fs from municipal solid waste incinerator by dual bag filter (DBF)system," Fuel, Vol. 86, pp. 813-819, March-April 2007.

[7] W. Lin, L. Wang, Y. Wang, H. Li, and G. Changchien, "Removal characteristics ofPCDD/Fs by the dual bag filter system of a fly ash treatment plant," Journal of Hazardous Materials, Vol. 153, pp. 1015-1022, May 2008.

[8] X. Qu, H. Wang, Q. Zhang, X. Shi, F. Xu, and W. Wang, "Mechanistic and kinetic studies on the homogeneous gas-phase formation of PCDD/Fs from 2,4,5-trichlorophenol," Environmental Science & Technology, Vol. 43, pp. 4068-4075, May 2009.

[9] M. Pandelova, D. Lenoir, and K.W. Schramm, "Correlation between PCDD/F, PCB and PCBz in coal/waste combustion Influence of various inhibitors," Chemosphere, Vol. 62, pp. 1196-1205, February 2006.

[10] B. K. Gullett, L. Oudejans, D. Tabor, A. Touati, and S. Ryan, "Near-real-time combustion monitoring for PCDD/PCDF indicators by GC – REMPI – TOFMS," Environmental Engineering Science, Vol. 46, pp. 923-928, December 2012.

[11] Z. Zhou, B. Zhao, H. Kojima, S. Takeuchi, Y. Takagi, and N.Tateishi, "Simple and rapid determination of pcdd/fs in flue gases from various waste incinerators in china using dr-ecoscreen cells," Chemosphere, Vol. 102, pp. 24-30, May 2014.

[12] H. Kojima, S. Takeuchi, M. Iida, S. F. Nakayama, and T. Shiozaki, "A sensitive, rapid, and simple dr-ecoscreen bioassay for the determination of pcdd/fs and dioxin-like pcbs in environmental and food samples," Environmental Science & Pollution Research, pp. 1-12, June 2015.

[13] W.Wang, "Modeling component concentrations of sodium aluminate solution via hammerstein recurrent neural networks," IEEE Transactions on Control Systems Technology, Vol. 20, pp. 971-982, July 2012 .

[14] J. Tang, T. Y. Chai, W. Yu, and L. J. Zhao, "Modeling load parameters of ball mill in grinding process based on selective ensemble multisensor information," IEEE Transactions on Automation Science & Engineering, Vol. 10, pp. 726-740, July 2013.

[15] M. Kano, and K Fujiwara, "Virtual sensing technology in process industries: trends & challenges revealed by recent industrial applications," Journal of Chemical Engineering of Japan, Vol. 46, pp. 1-17, March 2013.

[16] C. Shang, F.Yang, D. Huang, and W. Lyu, "Data-driven soft sensor development based on deep learning technique," Journal of Process Control, Vol. 24, pp. 223-233, March 2014.

[17] B. Gu, and V. S. Sheng, "A robust regularization path algorithm for ν-support vector classification," IEEE Transactions on Neural Networks & Learning Systems, Vol. 28, pp. 1241-1248, May 2017.

[18] R. Rosipal, and L. J. Trejo, "Kernel partial least squares regression in reproducing kernel Hilbert space," Journal of Machine Learning Research, Vol. 2, pp. 97-123, March 2002.

[19] J. Tang, J. Zhang, Z. W Wu, Z. Liu, T. Y. Chai, and W. Yu, "Modeling collinear data using double-layer GA-based selective ensemble kernel partial least squares algorithm," Neurocomputing, Vol. 219, pp. 248-262, January 2017.

[20] J. Tang, W. Yu, T.Y. Chai, Z. Liu, and X. J. Zhou, "Selective ensemble modeling load parameters of ball mill based on multi-scale frequency spectral features and sphere criterion," Mechanical Systems & Signal Processing, Vol. 66-67, pp. 485-504, January 2016.

[21] J. Tang, T. Y. Chai, W. Yu, Z. Liu, and X. J. Zhou, "A Comparative study that measures ball mill load parameters through different single-scale and multi-scale frequency spectra-based approaches," IEEE Transactions on Industrial Informatics, Vol. 12, pp. 2008-2019, December 2016.

[22] J. Tang, J. F. Qiao, Z. W. Wu, T. Y. Chai, J. Zhang, and W. Yu, "Vibration and acoustic frequency spectra for industrial process modeling using selective fusion multi-condition samples and multi-source features," Mechanical Systems & Signal Processing, Vol. 99, pp. 142-168, January 2018,

[23] T. Poggio, T. Vetter, "Recognition and structure from one 2d model view: observations on prototypes, object classes, and symmetries," A. I. Memo 1347, Artificial Intell. Lab., MIT, Cambridge, MA, 1992.

[24] S. Cho, M. Jang, and S. Chang, "Virtual sample generation using a population of networks," Neural Processing Letters, Vol. 5, pp. 21-27, April 1997.

[25] C. F. Huang, and C. Moraga, "A diffusion-neural-network for learning from small samples," International Journal of Approximate Reasoning, Vol. 35, pp. 137-161, February 2004.

[26] D. C. Li, and I. H. Wen, "A genetic algorithm-based virtual sample generation technique to improve small data set learning," Neurocomputing, Vol. 143, pp. 222-230, November 2014,

[27] Z. S. Chen, B. Zhu, Y. L. He, and L. A.Yu, "PSO based virtual sample generation method for small sample sets: applications to regression datasets," Engineering Applications of Artificial Intelligence, Vol. 59, pp. 236-243, March 2017.

[28] H. F. Gong, Z. S. Chen, Q. X. Zhu, and Y. L. He, "A monte carlo and PSO based virtual sample generation method for enhancing the energy prediction and energy optimization on small data problem: an empirical study of petrochemical industries," Applied Energy, Vol. 197, pp. 405-415, July 2017.

[29] G. Coqueret, "Approximate norta simulations for virtual sample generation," Expert Systems with Applications, Vol. 73, pp. 69-81, May 2017.

[30] G. Zhang, J. Hai, and J. heng, "Characterization and mass balance of dioxin from a large-scale municipal solid waste incinerator in china," Waste Management, Vol. 32, pp. 1156-1162, June 2012.

[31] G. Zhang, J. Hai, M. Ren, S. Zhang, J. Cheng, and Z. Yang, "Emission, mass balance, and distribution characteristics of pcdd/fs and heavy metals during cocombustion of sewage sludge and coal in power plants," Environmental Science & Technology, Vol. 47, pp. 2123-2130, February 2013.