Soft Sensor Modeling Method for Sulfur Recovery Process Based on Long Short-Term Memory Artificial Neural Network (LSTM)

Qing Tan and Bolun Li*
Qingdao Jeri Automation Co., LTD., Shenzhen Road, Qingdao City, China.

Abstract. In the process of sulfur recovery, H₂S and SO₂ concentrations reflect the effectiveness and reliability of the recovery process. However, the concentration of H₂S and SO₂ in the process of sulfur recovery is difficult to be measured by online analysis instrument, so the soft sensing modeling method is often used to analyze the system. Because SRU system has strong nonlinear characteristics and dynamic process characteristics, traditional soft sensing modeling methods are often limited in use. Long Short-Term Memory (LSTM) Artificial neural networks show strong ability in processing nonlinear data and dynamic data. Therefore, LSTM soft sensing method is used in this paper to systematically analyze the sulfur recovery process.

1 Introduction

With the development of industry, the scale of oil use is increasing. In the process of oil production, the processing of crude oil is the first and key step. The production process and products of crude oil are directly exposed to the public. The content of sulfur compounds has an important impact on the difficulty of crude oil production, crude oil quality and environmental pollution. There are still a lot of environmental problems in petroleum processing, especially the recovery of hydrogen sulfide (H₂S) and sulfur dioxide (SO₂) is becoming more and more important. These acidic exhaust gases. Once discharged into the atmosphere, it will cause serious pollution to the environment.

Sulfur Recovery Unit (SRU) is used to treat sulfur-containing acidic waste gas produced in petroleum processing. In SRUs, H₂S is converted into sulfur elemental and SO₂ is produced at the same time. Therefore, the exhaust gas from SRUs contains residual H₂S and SO₂, and the content of these two components should be strictly monitored. However, both H₂S and SO₂ are acid gases with strong corrosion, which will cause great damage to be measuring instruments, resulting in frequent maintenance and replacement of analysis instruments, and greatly increasing production costs of enterprises [1]. Therefore, it is very important to solve the measurement problem in SRU system.

In SRU system, variables not only have nonlinear characteristics, but also strong dynamic process characteristics. The output of the system at the current moment is not only related to the input at the current moment, but also to the input and output at the previous moment. The activation module of Long Short-Term Memory (LSTM) [2] artificial neural network mainly has four "gates", which are respectively input gate, output gate, memory gate and forgetting gate. These modules with memory function can maintain their state for a long time and regulate the information of incoming and outgoing units. LSTM networks are therefore more efficient at handling long-term time dependencies.

In view of the characteristics of SRU system and the application range of LSTM artificial neural network, this paper will combine theory and practice, and use LSTM to conduct soft-sensor modeling of H₂S concentration and SO₂ concentration in SRU system [3]. The validity of the LSTM soft sensor model is verified by comparing the parameter optimization of the LSTM soft sensor model through theoretical analysis and experimental methods.

2 Experimental auxiliary variable selection

At present, the mainstream SRUs adopt the Klaus process, which mainly takes the acidic gas of the upstream low-temperature methanol cleaning unit, the tail gas of the conversion unit and the flash vapor of the gasification unit as the main source of the sulfur-containing gas and converts the sulfur element in the sulfur-containing gas into single sulfur [4]. SRUs are mainly used to deal with two kinds of gas: one is H₂S produced by air purification equipment, also known as MEA gas. The other is H₂S and NH₃ produced by acid water vapor extraction device, also known as SWS gas. Fig. 1 shows the simplified SRU solution:
Fig. 1. Process flow chart of sulfur recovery plant

Table 1. Main variables of the SRU

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Variable name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MEA Gas flow rate</td>
</tr>
<tr>
<td>2</td>
<td>First-line air flow</td>
</tr>
<tr>
<td>3</td>
<td>Second-line air flow</td>
</tr>
<tr>
<td>4</td>
<td>SWS gas flow rate</td>
</tr>
<tr>
<td>5</td>
<td>SWS region air flow</td>
</tr>
</tbody>
</table>

The SRU input MEA gas and SWS gas are burned in the reactor and the gas is converted to pure sulfur. However, residual H₂S and SO₂ still exist in the exhaust gas produced by the treatment. Since the sensor is prone to corrosion in acid gas, soft sensing technology is needed to estimate the concentration of H₂S and SO₂ in the output stream [5-6]. Table 1 lists the auxiliary variables used in the SRU soft sensor model:

The data used in this experiment were taken from the SRU system data set of the coastal (Mediterranean) petroleum refineries of ERG Petrochemical Group in Italy. The dataset contains 10081 samples, 5 procedure variables and 2 target variables. Considering the dynamic operating conditions, data were collected with a continuous sampling period of 1 minute. Data preprocessing does not require data sorting, but only normalization processing. Among them, changes of MEA gas flow, first-line air flow and second-line air flow are shown in Fig. 2(a); changes of SWS gas flow and air flow in SWS area are shown in Fig. 2(b). For the SRU system, there is a large coupling relationship between auxiliary variables [7]. Therefore, to simplify the calculation and improve the accuracy of the model, it is necessary to carry out certain screening and other processing of auxiliary variables. This design uses the first 7000 groups of data as experimental data.

Mass variables in SRU system are H₂S concentration and SO₂ concentration. In this test, the changes of H₂S concentration and SO₂ concentration over time are shown in Fig. 3:

As can be seen from Fig. 3, H₂S and SO₂ concentrations present significant changes within the
working range.

Modeling requires training samples to establish a soft sensing model, and the pre-processed data are the measured values of auxiliary variables in the industrial production process. Since the accuracy of the measuring device and the process environment will affect the measurement data, resulting in errors and affecting the detection model, data processing is required. Firstly, data preprocessing needs to eliminate the influence of some obviously wrong values. Secondly, it is necessary to deal with the random error \[8\]. To eliminate the influence of different dimensions of data, normalized or standardized data preprocessing is usually carried out. The main data preprocessing methods include Min-Max method and Z-score method. The formula of Min-Max method is as follows:

\[
x_t = \frac{x_{\text{max}} - x_{\text{min}}}{\bar{x}_{\text{max}} - \bar{x}_{\text{min}}} \cdot (x_t - \bar{x}_{\text{min}}) + x_{\text{min}}
\]

(1)

Type (2-1) \(x_t\) said after processing data, processing \(\bar{x}_t\) said former data, \(\bar{x}_{\text{min}}\) said before processing in data collection, the minimum value in \(x_{\text{max}}\) said the highest value in the data set before processing. \(x_{\text{max}}\) indicates the maximum value and \(x_{\text{min}}\) indicates the minimum value of the data after processing.

The Z-score formula is as follows \[9\]:

\[
x_t = \frac{x_t - \bar{x}}{\sigma_x}
\]

(2)

Type (2-2) \(x_t\) said after processing data, processing \(\bar{x}_t\) said former data, the average of the vector \(x\) \(\bar{x}\) said, vector sigma \(\sigma_x\) said the standard deviation of \(x\).

The correlation coefficient is calculated by the following formula:

\[
R_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}
\]

(3)

In Formula (2-3), \(R_{x,y}\) represents the correlation coefficient between two vectors \(x\) and \(y\), \(\text{cov}(x,y)\) represents the covariance between two vectors \(x\) and \(y\), \(\sigma_x\) represents the standard deviation of vector \(x\), and \(\sigma_y\) represents the standard deviation of vector \(y\).

The final LSTM neural network parameters are shown in Table 2:

<table>
<thead>
<tr>
<th>Parameter name</th>
<th>The optimal parameter value of training set</th>
<th>Check the optimal parameter value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input dimension</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Output dimension</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Hidden layer number</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hidden layer dimension</td>
<td>250</td>
<td>150</td>
</tr>
<tr>
<td>Initial learning rate</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Learning rate decline cycle</td>
<td>175</td>
<td>150</td>
</tr>
<tr>
<td>Number of iterations</td>
<td>1500</td>
<td>700</td>
</tr>
</tbody>
</table>

Since the test set data cannot be known in the actual working condition, and the parameters of the training set are often different from the parameters to be tested, the optimal parameters of the check set are selected as the prediction model of the neural network.

3 Model prediction result

The LSTM soft sensor model uses the first 3500 sets of data in the SRU system as the training set, and the 3501-5600 sets of data as the verification set. Since the LSTM model requires a certain transition process, the data in the 5601-5650 group is selected as the transition process for prediction, and the data in the 5651-7000 group is finally predicted. The prediction results are shown in Fig. 4 and Fig. 5.

![LSTM model H2S predicted output and error](image1)

![LSTM model SO2 predicted output and error](image2)

(a) LSTM model H2S predicted output and error  (b) LSTM model SO2 predicted output and error

Fig. 4. LSTM model forecast output and error
4 Conclusion

In this paper, the effectiveness of soft sensor modeling method for sulfur recovery process based on LSTM was verified by experiments.

The prediction results of the LSTM model show that the RMSE value is smaller and the $R^2$ value is larger, which indicates that the LSTM model has higher accuracy and stronger generalization ability.

Compared with traditional soft sensor modeling methods, LSTM-based soft sensor modeling method has the following advantages:

(1) Strong generalization ability. LSTM can predict based on a small amount of data, and has low requirements on data time step, high prediction accuracy and strong generalization ability.

(2) Strong adaptability. LSTM is not only suitable for traditional linear systems, but also suitable for nonlinear systems and time-varying systems and can be widely used in various complex chemical industry fields [10].

However, LSTM model also has the following disadvantages:

(1) More sensitive to parameters. In the LSTM model, different parameters may lead to great differences in measurement results. The most important parameters are the number of hidden layer neurons and the selection of learning rate. When the high precision LSTM model is required, the parameters need to be constantly optimized to obtain the best effect.

(2) High performance requirements for arithmetic carrier. Because the training computation load of neural network model is large, the performance of computer is required to be higher.

References


