Online Flooding Prognosis in Packed Columns by Monitoring Parameter Change in EGARCH Model

Yi Liu, Bo-Fan Hseuh, Zengliang Gao, and Yuan Yao*

Abstract—In the chemical industry, packed columns are commonly used operating units for separation. However, the flooding phenomenon often reduces the efficiency of packed columns and interferes with the performance of the system. Due to this reason, research on the real-time prognosis of flooding becomes a necessity in practice. Pressure drop is a key factor that indicates flooding phenomenon in packed columns. In this paper, the trajectory of pressure drop in each time window is modeled with an exponential generalized autoregressive conditional heteroskedastic (EGARCH) process. The onset of flooding is then implied by the parameter change of the model. To capture the change in an efficient manner, a nonparametric charting technique is adopted for statistical process control (SPC). The feasibility and efficiency of the proposed method are illustrated by the experimental results.

I. INTRODUCTION

For separation applications such as absorption and distillation, packed columns have been widely adopted in chemical industrial processes, in which gas and liquid flow countercurrent. As known, gas and liquid flow rates in packed towers are limited by the tendency of the columns to flood. With the increase of either liquid or gas flow rate, the liquid holdup in the column increases, causing a decrease in the free area available for gas flow. Consequently, the pressure drop in the column increases. The so-called flooding point is finally reached, when gas bubbles rise violently through the liquid. In such a situation, the pressure drop in the column increases sharply even with the slightest increase in gas flow rate. At the same time, some liquid may be transported outside by the gas leaving through the top of the column [1]. The flooding phenomenon often leads to poor column efficiency, and may even shut down the entire production line. Therefore, packing columns cannot be operated or controlled at flooding conditions. On the other hand, high gas/vapor velocity often means high column capacity. The closer the column operation is to its maximum possible capacity, the less the energy consumption will be. Therefore, the packed columns function at their highest efficiency when they are operated close to the flooding point.

In order to prevent the operation of packed columns from flooding, various empirical models have been used to predict the operating limits [1-7]. However, the prediction accuracy always depends on empirical parameters related to the packed column under consideration, which are difficult to obtain [8]. Additionally, when the feed composition changes due to its availability, the predictive models may be unreliable. The inability of accurately predicting and preventing flooding may result in loss of operating hours, decrease of product purity, equipment damages, safety hazards, etc. In industrial operations, a conservative setting of gas velocity is often selected for safety reasons. This being about 70-85% of the flooding point velocity, it leads to low production rates and high energy consumption. Therefore, in order to ensure an efficient and a safe operation of packed columns, research on online flooding prognosis and detection becomes necessary.

When a packed column is operated approaching the flooding conditions, the slope of the trajectory of the pressure drop inside the column changes dramatically. Therefore, it is generally recognized that the pressure drop is the most informative variable for flooding prognosis. However, to the best of our knowledge, no statistical process control (SPC) strategy has been developed till date for online flooding prognosis. The main reason is that during the process of approaching the flooding points, the mean and the variance of the pressure drop change with fluctuations in either gas or fluid flow rates, violating the basic assumption of conventional SPC. Consequently, the direct application of SPC to flooding prognosis is not suitable.

In order to achieve real-time prognosis of the flooding phenomenon, the concept of profile monitoring is adopted in this paper. In each time window, the trajectory of the pressure drop is modeled with an exponential generalized autoregressive conditional heteroskedastic (EGARCH) process. Then, the indicative information of flooding is extracted by monitoring the parameter change of the EGARCH model with a nonparametric charting technique.

II. METHODOLOGY

A. EGARCH Process for Trajectory Modeling

The basic motivation behind the proposed strategy is that significant difference exists in the patterns of the pressure drop trajectory before and after the occurrence of flooding. Therefore, the first step of flooding prognosis is to model the variable trajectory with an appropriate time series representation.
As shown in Fig. 1, the trajectory of the pressure drop in a packed column where the liquid flow rate was approximately fixed at 0.89 m³/h and the gas flow rate was increased continuously until flooding was observed by process engineers. Here, the flooding phenomenon was identified after the 2550th sampling interval. This experiment exhibits the scenario in which the heat load on the reboiler at the bottom of the column is increased, so as to improve the output concentration at the column bottom. Obviously, this time series is nonstationary.

The most widely employed model for modeling nonstationary time series is autoregressive integrated moving average (ARIMA) [9]. However, this type of model is not suited to describe the pressure drop trajectory in packed columns as illustrated in the following. The detrended variable trajectory of the pressure drop shown in Fig. 1 is illustrated in Fig. 2. Additionally, the corresponding data variance of this detrended variable trajectory is shown in Fig. 3. It can be found that the variance of the data keeps changing even though the system may be operated below the flooding point, yet ARIMA only works for time series with constant variance.

For modeling data series exhibiting a time-varying variance, autoregressive conditional heteroscedasticity (ARCH) [10] is a common option, which can be regarded as an autoregressive (AR) process for the variance. Its extension is named generalized autoregressive conditional heteroscedasticity (GARCH) [11], which represents the data volatility in a more compact way by adopting an autoregressive moving average (ARMA) form. Although useful for various applications, the conventional GARCH process requires the effects of positive and negative shocks to be symmetric. Such a requirement makes it unsuitable for modeling the pressure drop data collected in packed columns. In the studied situation, the exponential generalized autoregressive conditional heteroskedastic (EGARCH) model [12] is more appropriate.

The formula of an EGARCH process is as below:

\[
\ln(\sigma_t^2) = \alpha_0 + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) + \sum_{j=1}^{q} \gamma_j \left( \frac{\epsilon_{t-j}}{\sigma_{t-j}} - \frac{E_t\left(\epsilon_{t-j}\right)}{\sigma_{t-j}} \right)
+ \sum_{j=1}^{q} \xi_j \left( \frac{\epsilon_{t-j}}{\sigma_{t-j}} \right)
\]  

For each time point \( t \), \( \sigma_t^2 \) is the conditional variance, \( \alpha_0 \), \( \beta \), \( \gamma_j \) and \( \xi_j \) are the coefficients, \( \epsilon_{t-j} \) are the residuals, and \( p \) and \( q \) are the lags in the model. Among the coefficients, \( \alpha_0 \) is a constant term, \( \beta \) are named GARCH coefficients, and \( \gamma_j \) are named ARCH coefficients [12].

The following offline processing steps should be taken to check whether it is proper to model the trajectory of the pressure drop with an EGARCH process.

1. A detrending step is conducted prior to the following analysis on the variable trajectory, because the variance information is more useful for flooding prognosis than the changes in the mean values.

2. Calculate the autocorrelation function (ACF) and the partial-autocorrelation function (PACF) [9] for the data series, to check whether the detrended data series is nonstationary.
3. Perform Ljung-Box-Pierce Q-test [13] and Engle’s ARCH test [10]. These tests can reveal whether an ARCH/GARCH class of model should be used.

4. Use sign bias test [14] to detect the existence of asymmetric effects.

If all the above mentioned tests show significant results, then the EGARCH model should be used to describe the process. Due to the page limitation, the details of these tests are not formulated and discussed in this paper. For more comprehensive information, please refer to the related references.

B. Procedure of Online Information Extraction with EGARCH Model

To extract the status information of the packed column, the variable trajectory of the pressure drop in each time window is modeled with the EGARCH model. In doing this, flooding prognosis can be performed by detecting the changes in model coefficients. The detailed procedure of online information extraction is introduced below.

1. Select the size \( D \) and the step length \( S \) of the moving window.

2. In the current window, the trajectory of the pressure drop is detrended. Different approaches can be utilized for detrending. Here, a best-fit line is simply subtracted from the original data series.

3. Fit an EGARCH model to the detrended data series. The lags in the model can be determined based on the Akaike information criterion (AIC) [15] or the Bayesian information criterion (BIC) [16].

4. Record the model coefficients, move the window forward, and return to Step 2.

C. Nonparametric Control Chart for Flooding Prognosis

After information extraction, SPC should be conducted to detect the changes in the important model parameter(s). A difficulty in doing this is that there is no prior knowledge of the distribution of each parameter. Here, nonparametric or distribution-free charting techniques [17-19] are adopted, which are useful when there is limited or lack of knowledge about the underlying data distribution. In comparison with traditional control charts, the distribution-free charts do not require estimating the distribution parameters, so that the influence of the estimation errors can be avoided. In this paper, the nonparametric control chart based on the Mann-Whitney (MW) test [19] is adopted.

The main idea of the MW control chart is as follows. Suppose that a training sample set of size \( m_1 \), \( X = \{X_1, \cdots, X_{m_1}\} \), is collected from an in-control process and that the test sample of size \( m_2 \) is denoted as \( Y^h = \{Y_1^h, \cdots, Y_{m_2}^h\} \). Here, \( h \) denotes the \( h \)-th test sample. In the case of flooding prognosis, \( X \) contains \( m_1 \) number of model parameter values calculated during an operation period without flooding, while \( Y^h \) consists of \( m_2 \) number of parameter values calculated online. The MW statistic \( M_{X,Y}^h \) summarizes the total number of \((X, Y)\) pairs where the value of \( Y_j^h \) is larger than that of \( X_i \), i.e.:

\[
M_{X,Y}^h = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} I(X_i < Y_j^h),
\]

where \( I(X_i < Y_j^h) \) is the indicator function for the event \( \{X_i < Y_j^h\} \). It is clear that the value of \( M_{X,Y}^h \) is between 0 and \( m_1 m_2 \). Large values of \( M_{X,Y}^h \) indicate the occurrence of a positive shift, whereas small values correspond to a negative shift. The control limits can be found in a table provided in the reference [19]. If the upper control limit is considered, the chart triggers an alarm if the following relationship is satisfied:

\[
M_{X,Y}^h > U_{m_1 m_2}.
\]

For developing an MW control chart, there are two parameters, \( m_1 \) and \( m_2 \), that should be specified in advance. The guidelines to be followed are: the size of the training set \( (m_1) \) should be large enough to adequately reflect the systematic variations during the normal operation without flooding. At the same time, a trade-off should be made in the determination of the size of each test sample \( (m_2) \), because too large a value may delay the prognosis, while too small a value may lead to inefficient detections.

III. APPLICATION RESULTS

In order to illustrate the proposed method, experiments were conducted on a lab-scale packed column shown in Fig. 4. The apparatus consisted of an acrylic column with a diameter of 0.22 m and a height of 2.20 m. The heights of both the upper and the lower packing layers are both 0.46 m. In the experiments, air is introduced from the bottom of the column, while water is fed from the top. The structured packing is CY1700 and the spray density ranges from 7-24 m³/(m²-h).

Because the body of the column is transparent, the process engineers/operators can observe the process status directly, providing a cross reference to the prognosis results.

A. Statistical Test

Taking the variable trajectory of the pressure drop plotted in Fig. 1 as an example, the statistical test results are as below. Fig. 5 and Fig. 6 show the ACF and the PACF of the detrended data series. It is clear that the ACF/PACF remains significant after a large number of lags, indicating the data series is nonstationary. The results of the Ljung-Box-Pierce Q-test and the Engle’s ARCH test in different time windows are listed in Table I and Table II, respectively, from which significant ARCH/GARCH data properties are observed. Table III shows the results of the sign bias test in the 1st time window, which confirms the necessity of using EGARCH models. The symbol “****” in the last column in Table III means that the test results are significant. In summary, it is reasonable to use EGARCH to model the trajectory of the pressure drop in the packed column.
B. EGARCH Model Fitting

Following the steps described in Section II.B, an EGARCH model was built for the variable trajectory of the pressure drop in each time window, where $D = 500$ and $S = 25$. The model orders were selected by AIC/BIC. Taking the data series in the 1st window as an example, Fig. 7 and Fig. 8 illustrate the order pair selection results. In each window, an EGARCH (1,1) model was selected.

The model coefficients calculated in different windows were plotted in Fig. 9 and Fig. 10, respectively. It is clear that the ARCH coefficient $\alpha_1$ shown in Fig. 9 behaves like a constant and does not contain any information about the flooding phenomenon. In contrast to the ARCH coefficient $\alpha_1$, the GARCH coefficient shown in Fig. 10 is more informative about the process dynamics. Therefore, in the next step, the MW control chart was utilized to monitor $\beta_1$ for flooding prognosis.

C. Flooding Prognosis Results

Due to the lack of distribution information of $\beta_1$, the nonparametric charting technique, MW control chart, was utilized for flooding prognosis. The size of the training sample, $m_1$, was selected as 50, while the size of the test sample, $m_2$, was 5. The control limit corresponding to $\text{ARL}_0 = 500$ was determined by looking up the table in [19], which is equal to 217. The step size of the moving window for online monitoring was chosen to be 1 so that to ensure the monitoring efficiency. The monitoring result is shown in Fig. 11. As observed in this figure, the control chart alarmed before the

<table>
<thead>
<tr>
<th>Window Index</th>
<th>p Value</th>
<th>Test Statistic</th>
<th>Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>765.749</td>
<td>31.4104</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>687.118</td>
<td>31.4104</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>768.469</td>
<td>31.4104</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>828.707</td>
<td>31.4104</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>1328.2677</td>
<td>31.4104</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>1007.466</td>
<td>31.4104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window Index</th>
<th>p Value</th>
<th>Test Statistic</th>
<th>Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>290.759</td>
<td>3.84145</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>255.854</td>
<td>3.84145</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>303.432</td>
<td>3.84145</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>295.114</td>
<td>3.84145</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>295.697</td>
<td>3.84145</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>202.13</td>
<td>3.84145</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sign Bias Test</th>
<th>t Value</th>
<th>Prob.</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign Bias</td>
<td>15.350</td>
<td>1.716e-39</td>
<td>***</td>
</tr>
<tr>
<td>Negative Sign Bias</td>
<td>9.755</td>
<td>1.145e-19</td>
<td>***</td>
</tr>
<tr>
<td>Positive Sign Bias</td>
<td>12.164</td>
<td>7.273e-28</td>
<td>***</td>
</tr>
<tr>
<td>Joint Effect</td>
<td>263.473</td>
<td>7.973e-57</td>
<td>***</td>
</tr>
</tbody>
</table>
monitoring result of GARCH coefficient for flooding prognosis process engineers recognized the flooding phenomenon. In other words, the proposed method detected the change in the process dynamics before the flooding did harm to the system and achieved early prognosis of flooding.

IV. CONCLUSION

This study addresses the issue of developing an efficient monitoring method for online flooding prognosis in packed columns. A statistical process monitoring approach has been proposed, based on the finding that the dynamic behavior of the pressure drop in the column changes significantly when the system is approaching the flooding conditions. In this method, the variable trajectory of the pressure drop in each time window is represented by a time series model with the EGARCH type. This model captures the volatility information that is indicative of the process status. By monitoring the changes in the GARCH coefficient with a nonparametric charting technique, early prognosis of the flooding phenomenon is achieved. The proposed method can be implemented in a straightforward manner. Its feasibility was illustrated with the experimental results.
ACKNOWLEDGMENT

This work was supported in part by Ministry of Science and Technology, ROC under Grant No. MOST 105-2622-8-007-009 and Jiangsu Key Laboratory of Process Enhancement & New Energy Equipment Technology (Nanjing University of Technology).

REFERENCES