Computers and Chemical Engineering 106 (2017) 96–105

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Multi-mode operation of principal component analysis with k-nearest neighbor algorithm to monitor compressors for liquefied natural gas mixed refrigerant processes

Daegeun Ha[a], Usama Ahmed[a], Hahyung Pyun[a], Chul-Jin Lee[b], Kye Hyun Baek[c,*], Chonghun Han[a,*]

[a] School of Chemical and Biological Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul, 151-742, Republic of Korea
[b] School of Chemical Engineering and Materials Science, Chung-Ang University, Seoul 06974, Republic of Korea
[c] Semiconductor R&D Center, Samsung Electronics Co., Ltd, 1, Samsungjeonja-ro, Hwaseong-si, Gyeonggi-do, 18448, Republic of Korea

ARTICLE INFO

Article history:
Received 4 March 2017
Received in revised form 24 May 2017
Accepted 31 May 2017
Available online 2 June 2017

Keywords:
Principal component analysis
k-nearest neighbor
Liquefied natural gas
Mixed refrigeration process
Multi-mode operation

ABSTRACT

LNG mixed refrigeration (MR) process is usually used for liquefying natural gas. The compressors for refrigerant compression are associated with the high-speed rotating parts to create a high-pressure. However, any malfunction in the compressors can lead to significant process downtime, catastrophic damage to equipment and potential safety consequences. The existing methodology assumes that the process has a single mode of operation, which makes it difficult to distinguish between a malfunction of the process and a change in mode of operation. Therefore, *k*-nearest neighbor algorithm (*k*-NN) is employed to classify the operation modes, which is integrated into multi-mode principal component analysis (MPCA) for process monitoring and fault detection. When the fault detection performance is evaluated with real LNG MR process data, the proposed methodology shows more accurate and early detection capability than conventional PCA. Consequently, proposed k-NN integrated multi-mode PCA methodology will play an important role in monitoring the LNG process.

1. Introduction

With an increase in the complexities in the industrial systems and high level process integrations, the need for more accurate process monitoring techniques to detect the instrumental malfunctions are required to improve product yield, reduce operational risk, ensure safe operation and achieve sustainable profit. In the industrial process, thousands of sensors, analyzers and control loops have strong interactions among each other in terms of process variables. In order to efficiently manage such complex processes, data-driven process monitoring methods have been utilized to handle large dimensional data from thousands of sensors. In this regard, MacGregor and Kourti (1995), Kano et al. (2001), Kourti (2002), Joe Qin (2003), and Venkatasubramanian et al. (2003) have developed multivariate statistical tools for process control.

Process operating condition changes frequently due to set point changes, fluctuations in raw materials, composition of feed material, equipment aging and seasoning effects. In these situations, the application of traditional process monitoring methods based on the assumption that the process has only one stable operation region may cause false alarms, when the process is operated under another steady-state nominal operating mode. This is because different modes of process usually have different statistical properties such as mean value, variance, and correlation between variables (Chiang et al. (2000), Lee et al. (2004), and Yin et al. (2014)).

To cope with these situations, multi-mode process monitoring methods were developed, which can be categorized into (a) global modeling, (b) multiple modeling, and (c) adaptive modeling approaches. Global modeling is an approach to developing the uniform model which can be applied to all operation modes. Hwang and Han (1999) proposed hierarchical clustering and super-principal component analysis (PCA) model. Lane et al. (2001) built a group model capable of generating the cluster of processes in terms of grades. Multiple modeling is a method to build local models which can match each operation mode with satisfying specific process operation condition. Zhao et al. (2004) and Yoo et al. (2007) proposed a multiple PCA model based process monitoring methodology. Adaptive modeling is a sort of a model update scheme for mode changes. Jin et al. (2006) and Choi et al. (2006) proposed a

* Corresponding authors.
E-mail addresses: kyehyun.baek@samsung.com (K.H. Baek), chhan@snu.ac.kr (C. Han).

http://dx.doi.org/10.1016/j.compchemeng.2017.05.029
0098-1354/

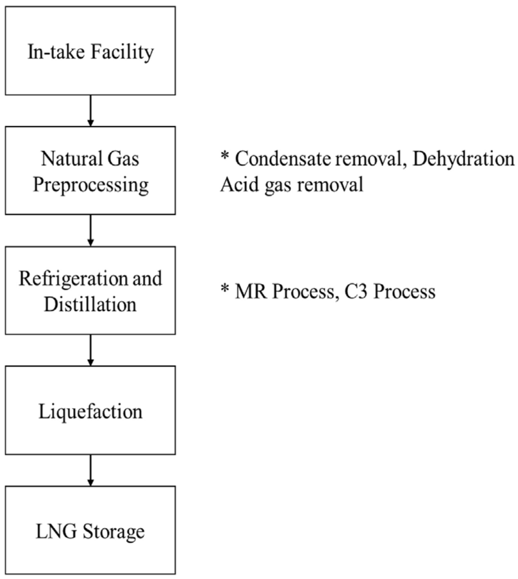

Fig. 1. Process flow of a typical LNG process.

Table 1
Variable description in the LNG MR Process.

Module	Variable	Description
Compressor #1	PI-001	Compressor #1 Inlet Pressure
	TI-001	Compressor #1 Inlet Temperature
	FI-001	Compressor #1 Inlet Flowrate
Compressor #2	PI-002	Compressor #2 Inlet Pressure
	TI-002	Compressor #2 Inlet Temperature
Compressor #3	PI-003	Compressor #3 Inlet Pressure
	TI-003	Compressor #3 Inlet Temperature
Compressor #4	PI-004	Compressor #4 Inlet Pressure
	TI-004	Compressor #4 Inlet Temperature
	PI-005	Compressor #4 Outlet Pressure
	TI-005	Compressor #4 Outlet Temperature
Compressor #5	PI-006	Compressor #5 Inlet Pressure
	TI-006	Compressor #5 Inlet Temperature
	FI-006	Compressor #5 Inlet Flowrate
Compressor #6	PI-007	Compressor #6 Inlet Pressure
	TI-007	Compressor #6 Inlet Temperature
	TI-008	Compressor #6 Outlet Temperature

robust recursive PCA monitoring methodology for a process that includes frequent operation mode changes.

There are certain limitations of those approaches. For instance, the global model which is based on multi-mode monitoring frequently shows low detection resolutions for a particular mode because it employs statistical mean of the whole data (Tan et al. (2011) and Ge et al. (2013)). In addition, a priori process knowledge is required in the preliminary step to manually segment the historical operating data according to different operating modes. Moreover, a similarity threshold should be predefined by user to incorporate the similar data groups. Those conditions are not desirable for automatic process monitoring in industrial practice. Multiple modeling has a detection capability with high resolution but the false alarm rate also increases in monitoring between-mode transition situations because it ignores cross-mode correlation (Zhang and Li (2013) and Haghani et al. (2014)). Adaptive modeling shows low reliability of monitoring results when process disturbance is included in model update (Joe Qin (1998) and Li et al. (2000)). Thus, it is still required to develop monitoring techniques of multi-mode operation cases.

This paper utilizes the k-nearest neighbor (k-NN) classification and multiple PCA models to enhance process monitoring performance under frequent operation mode change environments. The proposed methodology is evaluated with collected data from a real liquefied natural gas (LNG) mixed refrigeration (MR) process. The rest of paper is organized as follows. In Section 2, a LNG MR process is briefly described and the data collection procedure is explained. After describing the target process and data collection, theoretical background of PCA and k-NN is described in Section 3. In Section 4, mode identification techniques and fault detection models proposed in this study are explained as well as how they are integrated. In Section 5, performances of the proposed techniques are compared with those of general PCA by using real LNG MR process data.

2. Target process and data description

Fig. 1 shows a flow of a general LNG process, which consists of pre-processing, refrigeration, liquefaction, and storage sections. In the pre-processing stage, impurities including CO_2 and H_2S in natural gas are removed in acid gas removal unit, whereas the moisture is removed in the dehydration unit. In the refrigeration and distillation stage, mixed refrigerant (MR) which comprises of nitrogen, methane, ethane, propane, and sometimes butane is also used to cool down natural gas in a single main cryogenic heat exchanger (MCHE). In the liquefaction stage, liquefied natural gas reduces its volume by a factor of more than 600, making it more economical to transport over long distances by transport ships or pipelines.

This study focuses the MR compression process for process monitoring. The simplified MR compression process used in a LNG liquefaction plant is shown in Fig. 2. The storage suction drums are placed before the 1st and 5th stage of the compressor. They separate the liquid refrigerant from gas and liquid mixture so as to prevent the liquid phase of refrigerant from entering the compressors. In the MR compression, multiple stages of gas compression are required to achieve the desired pressure (27–34 bar) to make compressed mixed refrigerant. The MR compressor operates at 3.8 bar and continually compresses until the pressure reaches to 20 bar at fourth stage. The discharge flow from the 4th stage compressor is then cooled to 32 °C by the 1st cooler and passes to the 5th stage compressor via the 5th stage suction drum. The discharge flow is then compressed to the designed pressure level in the 5th and 6th stage compressors. The discharge from this compressor is cooled again by the 2nd cooler and is partially condensed in propane and mixed refrigerant evaporators. The mixed refrigerant liquid is subcooled while the mixed refrigerant vapor is condensed in MCHE. The partially condensed mixed refrigerant liquid is routed to a MR separator through the warm Joule-Thompson valve while the condensed mixed refrigerant vapor is sent to MCHE.

Major process variables described in Fig. 2 are summarized in Table 1. Total 17 variables are collected including compressor's pressure, temperature and flowrate variables in this study. Each pressure, temperature, and flow gauge is located at inlet or outlet of compressors. Among variables from each gauge, pressure variables can represent compressors' efficiency and flowrate variables can monitor refrigerant flow. If the operating conditions of compressors fluctuate or some compressors have a malfunction, pressure and flowrate are directly affected.

3. Theoretical background

3.1. Principal component analysis based fault detection

3.1.1. Principal component analysis

Principal component analysis (PCA) is a linear data reduction method by capturing the optimal variability in multivariate dataset.