Applied Thermal Engineering 130 (2018) 997–1003

Contents lists available at ScienceDirect

Applied Thermal Engineering

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

Research Paper

Modelling of a post-combustion CO_2 capture process using deep belief network

Fei Li [a], Jie Zhang [a,*], Chao Shang [b], Dexian Huang [b], Eni Oko [c], Meihong Wang [c]

[a] School of Chemical Engineering and Advanced Materials, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
[b] Department of Automation, Tsinghua University, Beijing 100084, China
[c] Department of Chemical and Biological Engineering, University of Sheffield, Sheffield S1 3JD, UK

ARTICLE INFO

Article history:
Received 27 December 2016
Revised 13 October 2017
Accepted 16 November 2017
Available online 20 November 2017

Keywords:
CO_2capture
Chemical absorption
Deep belief network
Restricted Boltzmann machine

ABSTRACT

This paper presents a study on using deep learning for the modelling of a post-combustion CO_2 capture process. Deep learning has emerged as a very powerful tool in machine learning. Deep learning technique includes two phases: an unsupervised pre-training phase and a supervised back-propagation phase. In the unsupervised pre-training phase, a deep belief network (DBN) is pre-trained to obtain initial weights of the subsequent supervised phase. In the supervised back-propagation phase, the network weights are fine-tuned in a supervised manner. DBN with many layers of Restricted Boltzmann Machine (RBM) can extract a deep hierarchical representation of training data. In terms of the CO_2 capture process, the DBN model predicts CO_2 production rate and CO_2 capture level using the following variables as model inputs: inlet flue gas flow rate, CO_2 concentration in inlet flue gas, pressure of flue gas, temperature of flue gas, lean solvent flow rate, MEA concentration and temperature of lean solvent. A greedy layer-wise unsupervised learning algorithm is introduced to optimize DBN, which can bring better generalization than a single hidden layer neural network. The developed deep architecture network models can then be used in the optimisation of the CO_2 capture process.

1. Introduction

Global climate change, as a result of the accelerated build-up of greenhouse gas (GHG) emission in atmosphere, has become a key concern of our society. The main component of GHG gas is carbon dioxide (CO_2). In the past few decades, numerous climate change policies were launched, but nonetheless, annual GHG emission still increased by 1.0 $GtCO_2$-eq (2.2%) per year from 2000 to 2010, compared to 0.4 $GtCO_2$-eq(1.3%) per year, from 1970 to 2000 [1]. A rapidly growing population plus industrialization, with corresponding increase in energy demand, is likely to result in increasing amount of GHG emission. Consequently, the Intergovernmental Panel on Climate Change has proposed that, compared to the emission levels in 1990, a 50% reduction of CO_2 emission is needed by 2050 [2].

The main source of worldwide CO_2 emission is the combustion of fossil fuel, such as petroleum, crude oil, natural gas and coal [3]. Amongst them, coal-fired power plants offer some advantages to operators, not only because of high availability of coal compared to other nature fuels, but also due to its flexible operation to changes in supply and demand [4]. However, the amount of CO_2 emission per unit of electricity released by coal-fired power plants is twice as much as their natural gas counterparts [5]. As a result, many researches have been explored to reduce the CO_2 gas emission from coal-fired power plants. Carbon capture and sequestration (CCS) is identified as an appropriate technique for the sustainability of coal-fired power plant, because of its efficiency and effectiveness in reducing CO_2 emission [6]. Amongst the various technologies of CCS, the post-combustion carbon capture technology with chemical absorption has been considered as the most suitable way to reduce CO_2 emission. This is because it can retrofit the existing coal-fired power plant easily and treat flue gas stream with low CO_2 partial pressure [4]. However, it still has some disadvantages, one of which is the large energy requirement for absorbent regeneration. The thermal energy for regeneration usually comes from extracted steam from the low pressure steam turbines, which will reduce the efficiency of the coal-fired power plant. Therefore, it is particularly important to find out the trade-off

* Corresponding author.
E-mail address: jie.zhang@newcastle.ac.uk (J. Zhang).

https://doi.org/10.1016/j.applthermaleng.2017.11.078
1359-4311/

Nomenclature

Abbreviations

BANNs	bootstrap aggregated neural networks
BAELM	bootstrap aggregated extreme learning machine
CO_2	carbon dioxide
CCS	carbon capture and sequestration
DBN	deep belief networks
GHG	greenhouse gas
MSE	mean squared error
MEA	monoethamine
RBM	restricted Boltzmann machines
RMSE	root mean squared errors
SLNNs	single-hidden layer feedforward neural networks

Letters

a	mean
$\mathbf{b}$	bias vector of visible layer
$\mathbf{c}$	bias vector of hidden layer
E	energy function
$\mathbf{h}$	vector of hidden unit outputs
m	mass fraction
P	probability distribution function
t	discrete time
u	input variables
$\mathbf{v}$	vector of visible unit outputs
$\mathbf{w}$	weight vector between visible units and hidden units
y	output variables

Greek symbols

α	learning rate
ε	flow rate
η	capture level
θ	parameter set of a neural network
μ	CO_2 production rate
σ	standard deviation

between CO_2 capture level and energy consumption by using process optimisation. In order to carry out process optimisation, it is necessary to develop an accurate model for the post-combustion carbon capture process.

The models for CO_2 capture processes can be categorized into three groups: mechanistic, statistic and artificial intelligence based models. Lawal et al. [4] have developed a mechanistic model to present how disturbances affect the carbon capture process performance. In their study, the mechanistic model is able to predict the column temperature profile and CO_2 loading with high accuracy. On the other side, it consumes a lot of time consuming and also requires extensive knowledge of the process underlying physics. Further, Zhou et al. [7] have proposed a statistical model to instead of mechanistic model, for predicting heat duty, CO_2 production rate, CO_2 lean loading and capture level. In their statistical analysis, the selection of process variables in the model was affected by experts' opinion. However, the statistical model cannot describe the irregular non-linear relationships between process variables. To tackle these problems, they explored a neural network model later and then compared its performance with the previous statistical model [8]. From that research, they found using neural network model was able to predict CO_2 production rate with much higher accuracy than the statistical model. Hence, the artificial intelligence based models become increasingly significant and draws much more attention in modelling the post-combustion CO_2 capture plants. Meanwhile, Sipocz et al. [9] has also used a single-hidden layer feedforward artificial neural network (SLNNs) to model the chemical absorption process and explore the relationship between process variables and quality variables. Similarly, the results in the paper have shown that artificial neural network models are able to simulate the rigorous complex steady state process with minimum time demand and high accuracy. However, the SLNNs model would possibly encounter some problems such as over-fitting of the training data and poor generalization performance. To overcome these shortcomings, Li et al. [10] have presented a new learning algorithm, called bootstrap aggregated neural networks (BANNs), with a combination of several single-hidden layer neural networks together to model post-combustion CO_2 capture process. The modelling technique was verified to predict CO_2 capture level and CO_2 production rate with higher accuracy and reliability than the traditional neural network models, especially performs better in the aspect of long-term predictions. Besides, BANNs model is able to measure prediction reliability by prediction confidence bounds. Later, they have further explored a fast learning algorithm on the basis of BANNs, called bootstrap aggregated extreme learning machine (BAELM) [11]. The structure of ELM is the same as SLFNNs, but the network weights are calculated differently. In ELM, the weights between input and hidden layers are randomly assigned while the weights between hidden layer and output layer are obtained by one-step regression. Their research results show that the training time of BAELM is 5 times lower than that of BANNs. Furthermore, the generalization performance of BAELM is better than BANNs, as indicated by lower mean squared errors (MSE). Nevertheless, all of the above mentioned neural networks have only one hidden layer. They may possibly have difficulty in handling strongly correlated industrial process variables. In other words, the single hidden layer neural networks face difficulties when representing complex and highly-varying relationship and are easy to converge to local optima [12].

A solution to these problems has been proposed, which is a deep multi-layers neural networks model inspired by the structure of human brain [13]. The multiple hidden layers represent different levels of latent features of input data. As a result, it can deal with the data which has complicated irregular relationship. It is true that, before 2006, the attempts to apply deep multi-layers neural networks were failed because of unsuccessful training strategies. The reason is that the gradient-based method is starting from random initialization, thereby getting stuck near poor solution. Nonetheless, in 2006, Hinton et al. [13] put forward a greedy layer-wise unsupervised learning algorithm for deep belief networks (DBN), which pre-train one layer at a time in a greedy way. Simply, DBN is a generative model, in which the lower layers represent the low-level features from inputs and the upper layers extract the high-level features that explain the input samples. With the comparison of random initialization, the results show that the initial parameters of DBN are much closer to optimal solutions [14]. Since then, increasing attention has been paid to DBN model and it contributes a lot to image recognition [15] and time series forecasting [16]. However, it has never been applied to modelling of CO_2 capture processes.

The rest of paper is arranged as follows. Section 2 presents the theoretical knowledge of DBNs and their component layers, Restricted Boltzmann Machine (RBM). Section 3 introduces the details of a post-combustion CO_2 capture plant. Then, the comparative result analysis between SLNNs and DBNs is revealed in Section 4. Finally, Section 5 gives conclusions and future works.