Transfer Learning Can Be An Optimizer for Chemical Process: A Case Study of Ammonia Synthesis

Ming Zhu

Abstract—Artificial intelligence (Al) is of growing interest in chemical industry as numerous published articles show explicitly. In this work, we propose a transfer learning method for dynamic optimization on industrial size plant. Due to several included features, this A1 tool makes the process optimization in the industrial practice easy, comfortable, and robust applicable. To demonstrate the capability of transfer learning, the optimization of the industrial ammonia synthesis process is described in this work. The achieved optimization results show that 12% more ammonia can be produced with 22.9% less of hydrogen consumption, 12% less of nitrogen supply and 5bar lower of reaction pressure. It will significantly benefit engineers and project managers working in the field of clean ammonia synthesis, as it provides more liexibility for energy supplier such as wind or solar power. The utilization of transfer learning is not limited in ammonia synthesis, it has the potential to be applied in otherchemical processes.

***Index Terms*—Artificial intelligence, deep learning, transfer learning, ammonia synthesis, dynamic optimization.**

1. Introduction

Ammonia has been caught global attention recently because it is considered as hydrogen cartier — the potential energy resource. Ammonia produced based on wind power or solar energy may reshape the configuration of global energy, which is called "liquid sunshine" [1]. Traditional ammonia synthesis is though to be energy consuming and accompanied with large carbon emission. In recent years, clean ammonia synthesis with low carbon emission and low energy consumption is gradually arising, which is expected to bring a revolution in ammonia production. However, the uncertainty of wind power and solar energy limits the application of clean ammonia synthesis. Therefore, optimization tools are highly demanded for industrial application of clean ammonia synthesis.

Correspondence concerning this article should be addressed to Ming Zhu(ORCid 0000-0003-2812-5406), with the Department of Chemical Engineering, Nanjing Tech University, Nanjing, Jiangsu 211816, P. R. China (e-mail:zhux0447 @hotmaiL.com).

1.1. Ammonia synthesis and process optimization

Ammonia synthesis has been studied for over a century, and some progresses are just made recently. Most works focus on materials development, e.g. catalyst [2-4] and adsorbent [5,6]. Fewer works concern about process integration, i. e. combining adsorption with reaction [7], plasma enhancement [8] and even electro-catalysis [1]. Very few work concern about dynamic optimization in industrial scale. But it is crucial for realizing clean ammonia production, especially when the supply of resource gases are under variation. In the process of ammonia synthesis (as shown in Fig. 1), hydrogen can be supplied viaelectrolysis of water. Nitrogen is obtained from air separation unit (e. g. pressure swing adsorption). The mixture of hydrogen and nitrogen are compressed to reactive pressure, and preheated to initiate the synthesis reaction. Reactor effluent gases are chilled and condensed through a group of cooler. The anhydrous ammonia is separated from recycle gas to yield high purity product. The unreacted syngas is recycled through the reaction loop. This process is regulated by a number of chemical-, fluid mechanic- and thermal dynamic laws which brings high level of complexity for the process optimization of clean ammonia synthesis plant.

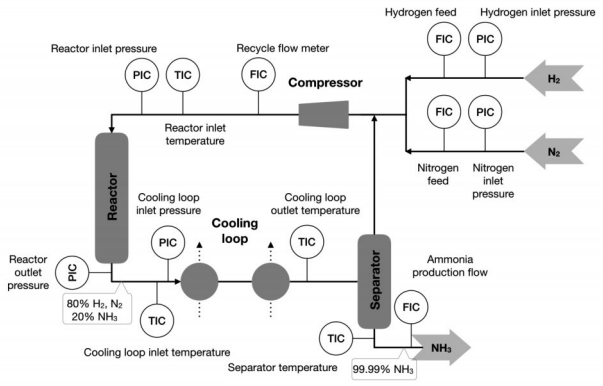

Fig. 1. Flow diagram of ammonia synthesis process.

The complexity of ammonia plant brings challenge towards process optimization. First, the actual industrial process is often characterized by multiple peaks and high dimensional parameters. Physical model based optimization technique (POT) works well on one dimensional search, but performs worse on high dimensional task [9]. For example, to find the optimal operation of distillation column, we tune

the feed tray of the column while keep other parameters fixed. But under actual industrial conditions, e.g. ammonia synthesis, it contains 11 operational parameters (as shown in Fig. 2), and all in instant fluctuation. Is that possible to improve ammonia production based on the dynamic operation? Second, the variation under practical conditions are almost discontinuous or non-derivable. Therefore, the gradient-based algorithm is infeasible due to thehighly dependence on derivability [10]. Lastly, the available time for obtaining solution is usually limited in the practice of chemical process optimization [10]. Therefore, computational efficiency should be ensured. For this reason, artificial intelligence (Al) has been considered as a valuable alternative modeling approach to replicate the rigorous model and at the same time obtain the same level of detail.

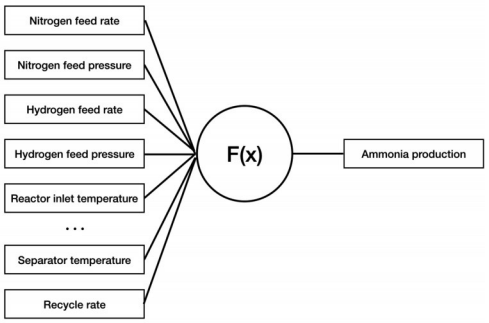

Fig. 2. Key operational parametexs for ammonia production.

1.2. Artificial intelligence and transfer learning

Artificial intelligence (Al), especially machine learning (ML) is thought to be revolutionary in chemical engineering. It is faster and more accurate to solve complex problems that expert system is slower to update new information and make corresponding changes [11]. There are several periods of AI's development. The first period is the application of neural networks (around 1990-2008) [11]. People used a feedforward neural network with input, hidden and output layers to solve nonlinear function approximation problems in an automated manner [12]. Then, Rumelhart, Hinton, and Williams developed a back-propagation algorithm to learn hidden patterns from input-output data [11]. Afterward, Bakshi and Stephanopoulos developed the WaveNet [13], which was an early work for nonlinear principal component analysis. It is worth to notice that the progress of computational capability in the recent decade initiates a new era of deep learning.

Deep learning is a broad concept that may include deep or convolutional neural nets (CNNs), reinforcement learning, transfer learning, and statistical machine learning. CNN is a filtering technique, well known in the domain of signals processing, for extracting features from a noisy signal [14-16]. After initial specification of the network architecture and the filter parameters such as the size and number of filters, a CNN learns during training, from a large dataset to conduct a successful performance based on appropriate filters. Another architectural innovation is the recurrent neural network (RNN). RNN is an extension of the multilayer perceptron with feed back connections [17, 18]. Long short-term memory (LSTM) network is suited for making prediction based on time series data, whereas overcomes the vanishing of gradient descent in RNN [17, 19]. Gated recurrent unit (GRU) network architec ture can enhance processing of sequential information with only two gates—reset gate and update gate [20, 21]. As for reinforcement learning, it is more like training a pet, which can get positive feedback from desired performance [22]. But it requires large number of resource data, which may not be achieved in chemical process. Statistical ML is the combination of mathematical methods from probability and statistics with ML techniques. It includes a number of useful techniques such as LASSO, Support Vector Machine, random forests, clustering, and Bayesian belief networks [11].

Transfer learning (TL) is the adaptation of one process to other similar process, so that the annotation reviews can be reduced [23]. TL is currently used in the material field to analyze the data and speed up material development. DeCost et al. [24] utilized a transferred deep convolutional network to learn representations of microstmctures and then used these representations to infer underlying annealing conditions. Wu et al. [25] used TL to refine the relationship between polymer chemistry and thermal conductivity. Wu et al. [26] then generalized the TL strategy to other properties of polymers and inorganic materials. Ma et al. [27] transferred the knowledge from 13 506 metal-organic framework (MOF) adsorption to small sample data and built surrogate models. Sun et al. [28] used TL to predict desorption of alkanediol/solvent/zeolite system, and reported an improved accuracy. The essence of TL in material research is laid on knowledge transfer. It can be a new insight into structure-property relationship and drastically accelerate material evolution.

So far, TL is used for model prediction. There are four types of transfer learning based on literatures: parameter-based transfer learning, which transfers parameters between models; instances-based transfer learning, which focuses on the instance weighting strategy; feature-based transfer learning, which focuses on transforming the original features to the new feature representation; relational-based transfer learning, which focuses on transferring the logical relationship or rules. The recent work of Ma et al. [27] reported using parameterbased TL to improve prediction of MOF adsorption data based on a small dataset training. The source model was a deep neural networks (DNNs) trained on H_2 adsorption data with 13 506 MOF structures at 100bar and 243K. The prediction target were adsorption capacity of CH_4 at 100bar and 298K, and Xe/Kr at 5bar, 298K. They reported that for Ha adsorption at 100bar and 130K, TL worked in 89.3% of the cases. For CH_4 adsorption, TL worked in 82.0% of the cases. For Xe/Kr adsorption, TL worked in nearly 50.0% of the cases. As well known, in a DNN-based model, the relationship between material structure and property is encoded in the weights and biases. Considering different adsorption mechanisms for MOF material, it may not be appropriate to use TL for model prediction

across different gas species. However, we will present in this work that TL can be used as an optimizer for chemical process.

1.3. The contribution of this work

This study made two novel contributions. First, we declare that TL may not only be used as prediction model, but also an optimization tool for chemical process to accomplish high dimensional tasks. The case study of clean ammonia synthesis with high dimensional variation is highlighted in this paper to demonstrate the capability of TL. Second, we propose an improved operation for ammonia synthesis on an industrial size plant. Such dynamic optimization in industrial scale are rare reported in the literature. These two contributions address many of the limitations mentioned in section 1.1 and 1.2.

2. MODEL DESCRIPTION

This section describes the model used for dynamic optimization of ammonia synthesis, which includes preparation of operational data, encoder-decoder architecture for mapping input-output relations, and transfer learning for optimization.

2.1. Data preparation

The original data is collected from on-site operation of Jiangsu Nuomeng Inc., China, which is denoted as series B. It contains dataset of 50 000 sampling points covering a period of 800hrs. Operational parameters include hydrogen feed rate, hydrogen feed pressure, nitrogen feed rate, nitrogen feed pressure, etc., as listed in Fig. 2. These parameters can be divided into two categories: Cumulative parameters, as it cumulates with time, such as hydrogen feed rate and nitroen feed rate; State parameters, e.g. nitrogen feed pressure, reactor inlet temperature, and reactor outlet pressure, etc.. The output parameter in this case is ammonia production because it is main concerning of this work.

The series A is the comparison group which can be generated from variational method. This includes two steps: First, a prior distribution $p\theta(z)$ is generated from a dataset of $z(i)$; Then, a dataset of $x(i)$ is generated from a conditional distribution $p\theta(x|z)$. A recognition distribution $q\phi(z|x)$ is used to mapping $x(i)$ from the intractable tree posterior $p\theta(z|x)$. According to literature [29], the log-likelihood of $x(i)$ can be written as the sum of KL divergence and the lower bound between the true posterior $p\theta(z|x)$ and the approximation $q\phi(z|x)$.

$$logp_\theta(x^{(i)}) = D_{KL}[q_\Phi(z|x^{(i)})||p_\theta(z|x^{(i)})] + L(\theta,\Phi;x^{(i)}) \quad (1)$$

$$L(\theta,\Phi;x^{(i)}) = E_{q_\Phi(z|x^{(i)})}[-log(q_\Phi(z|x^{(i)})) + logp_\theta(z|x^{(i)})] \quad (2)$$

After re-parameterization [29], the resulted distribution of $x(i)$ can be written as:

$$logp_\theta(x^{(i)}) = \frac{1}{2}\sum_{j=1}^{J}(1 + log((\sigma^{(i)})^2) - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2) + \frac{1}{D}\sum_{d=1}^{D} logp(x^{(i)}|z^{(i,d)}) \quad (3)$$

2.2. Encoder-decoder architecture for mapping inputoutput relations

The relations between multiple input and single output of ammonia production rate is replicated via encoder-decoder architecture containing LSTM and GRU layers (Fig. 3). A set of operational parameters are selected as decision variables such as hydrogen feed rate, hydrogen feed pressure, nitrogen feed rate, nitrogen feed pressure, etc.. The operational parameters are input to the first layer of LSTM to read the sequences and encode them in a fixed length vector. The encoded vector is then transferred to the second layer of GRU to decode the information and coordinate it with anunonia production rate. To output the terminal task, two layers of dense neural networks are employed at the end. The number of cells used in the LSTM layer, GRU layer and first layer of dense neural network are 100, 100 and 10 respectively. It is used to give a forward step-size for modeling prediction with LSTM and GRU cells. In this case, we just want to replicate the inputoutput relations, thus the forward step-size can be set zero, as a supervised learning process.

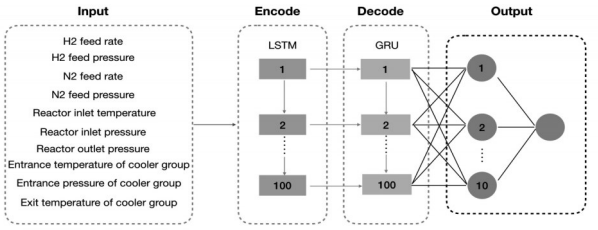

Fig. 3. The encoder-decoder architecture for mapping input-output relations.

2.3. Transfer learning for optimization

In this case, we use transfer learning (TL) to find optimal solution for ammonia synthesis. TL is used to play the role of modeling prediction based on transferring knowledge from small amount of data set. However, its intrinsic nature lying in the transferring knowledge from weights and bias makes it a strong tool for optimization. For optimization, TL transfers operational pattern to one of the other operations. If the inherited operational pattern is superior than the transferred pattern, the output, which is ammonia production rate in this case, will be higher than the output of transferred pattern. We notice that the significant differences exist in the modeling prediction of material adsorption across different gas species, as it works 89.3% for hydrogen prediction, 82.0% for methane prediction, and nearly 50% for Xe/Kr prediction [27]. It is not TL has low accuracy, but that TL indicates relative adsorption capacity on different materials.

TABLE I
ALGORITHM FOR TL OPTIMIZATION

For series A and B,
1. Supervised learning of A-A and B-B using encoder-decoder architecture.
2. Transfer learning to compare results:
  a. Transfer learning of A-B and B-A;
  b. arg max {A-B, A-A, B-A, B-B},
    If A-A,
    A is superior than B
    If B-B,
    B is superior than A
    Else, A equals to B
    Regeneration of A, and return to step 1.
End

As an optimization tool, cross TL is essential to identify an optimal solution from original operation and comparison group. The process contains two steps: First, the input-output relations is replicated via supervised learning, which obtains the results of A-A and B-B. Second, the cross TL is adopted to obtain the results of A-B and B-A. The TL calculation is also composed of two steps: The first step, the model is trained with input data and saves its weights as input-output relations. Secondly, the pretrained model is used to predict a similar process while keep the parameters of ultimate layers fixed. The corresponding algorithm for TL optimization is showed in Table I. The concept design for TL optimization is showed in Fig. 4.

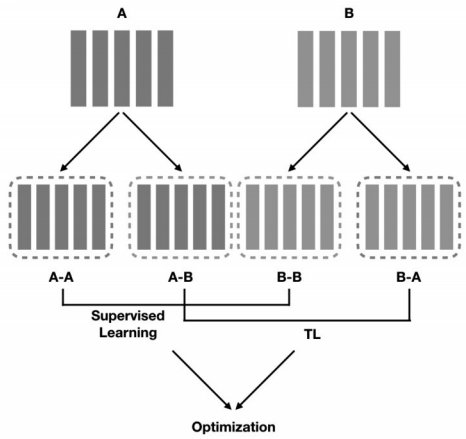

Fig. 4. Concept design for TL optimization.

3. RESULTS

In clean ammonia synthesis, hydrogen is supplied from electrolysis of water, and the energy is provided by solar or wind power. Due to the uncertainty of energy supplier, there are two types of hydrogen feed strategy such as series A and B (as shown in Fig. 5). Series A represents a low hydrogen feed strategy with hydrogen consumption in one and a half hour of 194.4moles. The hydrogen consumption is calculated by integrating hydrogen feed rate with time. In the same period, hydrogen consumption for series B is 239.1moles, which is 22.9% higher than series A. Series A varies in a range of 0269.5moles/hr, with average feed rate of 128.9moles/hr. Series B has a rate impulse up to 411.7 moles/hr. According to Haber reaction theory [7], This impulse of hydrogen feed should react with an increase of nitrogen feed to keep the stoichiometry as 3: 1 ($3H_2+N_2\rightarrow 2NH3$). Therefore, an increase of hydrogen usually indicates an increase supply of nitrogen.

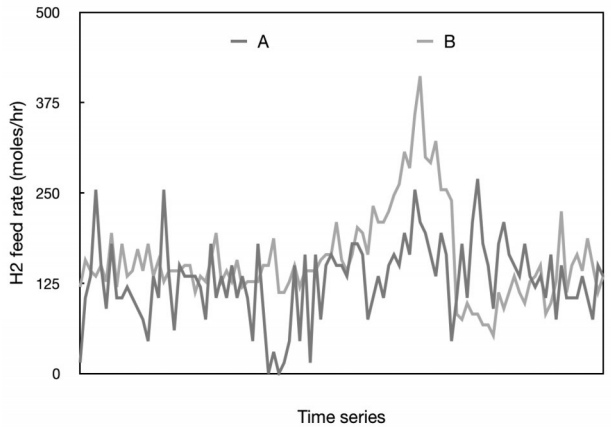

Fig. 5. Hydrogen feed rate of series A and B.

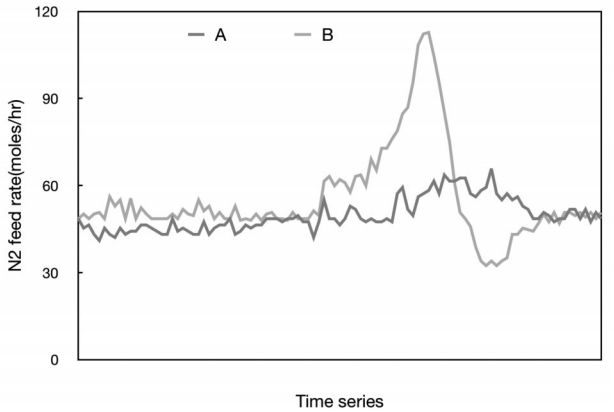

Fig. 6. Nitrogen feed rate of series A and B.

Nitrogen used for ammonia synthesis is supplied from air separation unit, e.g. pressure swing adsorption. To keep the stoichiometry as 3: 1, nitrogen supply should be increased accompany with the impulse of hydrogen feed (series B as shown in Fig. 6). Therefore, in Fig. 6, the nitrogen consumption of series B is 83.6moles in one and a half hour by time integration, while it is 73.8moles for series A. The nitrogen feed rate of series A varies in a range of 41.0-65.8moles/hr, while it is 32.4-112.8 moles/hr for series B. However, the increase of nitrogen and hydrogen feed does not ensure an increase of ammonia production. As shown in Fig. 7, a lower feed strategy of series A obtains higher ammonia production rate than series B. The total ammonia production in one and a half hour for series A is 38.3moles, 12% higher than that of series B. It is good news for clean ammonia synthesis, because one can use less hydrogen and nitrogen while produce more ammonia. To further understand the increase of ammo-

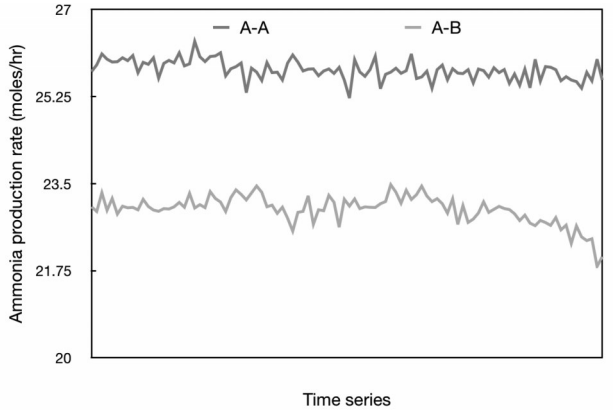

Fig. 7. Ammonia production rate of series A and B.

nia production, we need to trace back to the state operational parameters.

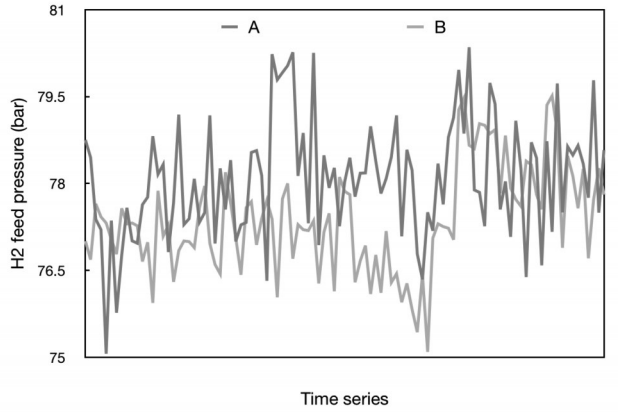

Fig. 8. Hydrogen feed pressure of series A and B.

We trace back to the hydrogen and nitrogen feed pressure of series A and B, as shown in Fig. 8 and Fig. 9 respectively. For hydrogen feed pressure, series A adopts a higher pressure strategy. The average pressure of series A in Fig. 8 is 78.1 bar, and 54% pressure points are over 78bar. Compared with series A, hydrogen feed pressure of series B is lower. The average pressure of series B in Fig. 8 is 77.4 bar, 0.6bar lower than series A. It indicates that higher pressure for hydrogen supply is beneficial for ammonia synthesis in this case. Higher feed pressure can reduce the hydrogen demand by improving the efficiency of hydrogen supply. It is especially useful for the renewable energy supplier such as wind and solar power, because it allows more flexibility for hydrogen generation.

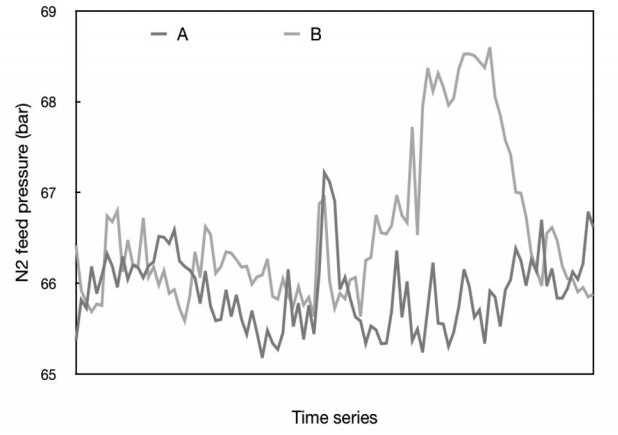

Fig. 9. Nitrogen feed pressure of series A and B.

As shown in Fig. 9, the nitrogen feed pressure of series A is lower than series B. The average pressure of series A in Fig. 9 is 65.9 bar, while it is 66.6bar for series B. The maximum feed pressure of series B is 68.6bar. Lower feed pressure of nitrogen can not only reduce the power consumption of gas compressor, but also increase the conversion rate for ammonia synthesis. In this case, iron-based catalyst is used for Haber reaction. The iron-based catalyst is rate limited of nitrogen dissociation [30]. Therefore, low feed pressure of nitrogen is superior to increasing ammonia conversion rate.

The state parameters-temperature and pressure for synthesis reaction are shown in Fig. 10 and Fig. 11, respectively. Fig. 10 shows the reactor inlet temperature of series A and B. The inlet temperature of series A varies in a range of 710.4-713.7K. The variation range of series B is 710.5-714.0K. The average temperature of series A in Fig. 10 is 711.8K, which is the same as series B. It indicates the reaction temperature can not be reduced based on the same iron catalyst. The reaction temperature is limited by chemical potential, therefore, to reduce reaction temperature requires the development of catalyst. However, based on the same catalyst used, the reaction pressure can be reduced (series A as shown in Fig. 11). The average reaction pressure is ll2.0bar for series A, 5bar lower than the one of series B.

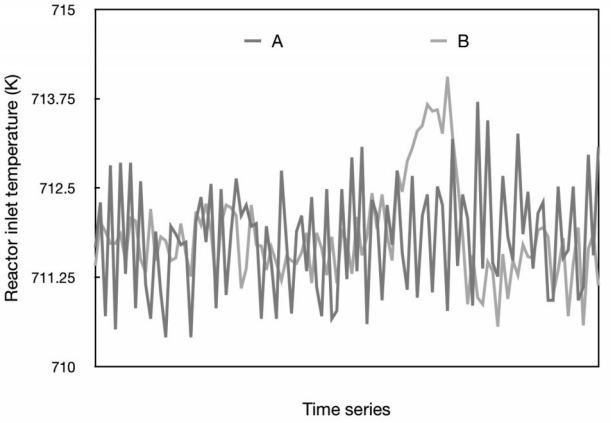

Fig. 10. Reactor inlet temperature of series A and B.

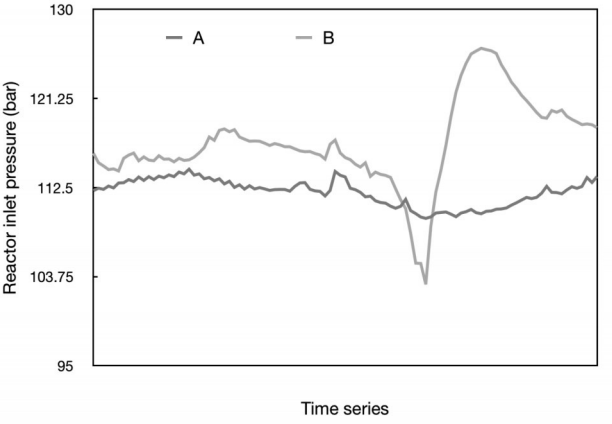

Fig. 11. Reactor inlet pressure of series A and B.

As we know that series A obtains higher ammonia production, it is worth to learn the operational pattern of series A in pre-cooling section, separation unit and recycle unit (Fig. 12-15). Generally, series A employs a distributed cooling strategy, which balance the cooling duty between cooler group and separator. However, series B put more cooling duty on the separation unit. As shown in Fig. 14, the inlet temperature of separator is increasing (series B), which cor-

responds to the sharp decrease of entrance temperature of cooler group in Fig. 12 and decrease of recycle rate in Fig. 15. As a result, the distributed cooling strategy makes series A a more stable operation than series B in the recycle unit.

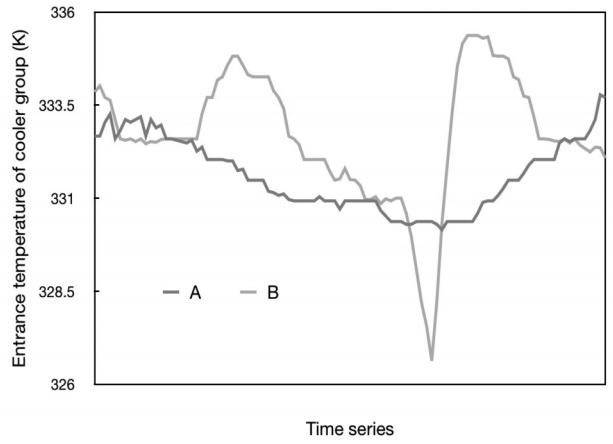

Fig. 12. Entrance temperature of cooler group.

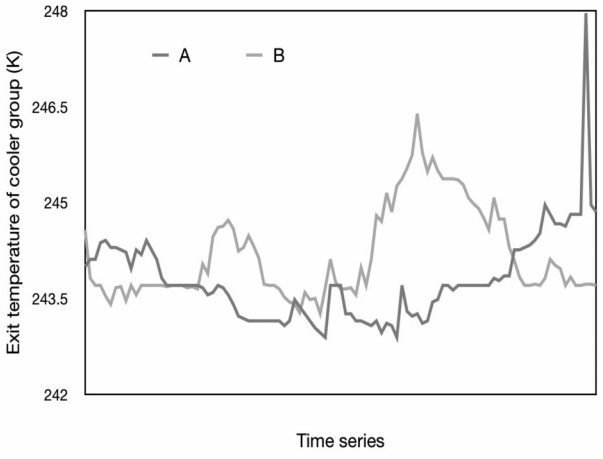

Fig. 13. Exit temperature of cooler group.

4. DISCUSSION

To evaluate our proposed method for process optimization, we use several conventional evaluation tools such as mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE) and coefficient of variation (CV). These tools are defined as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2 \tag{4}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i| \tag{5}$$

$$RMSE = \sqrt{MSE} \tag{6}$$

$$CV = \frac{RMSE}{\bar{x}} \tag{7}$$

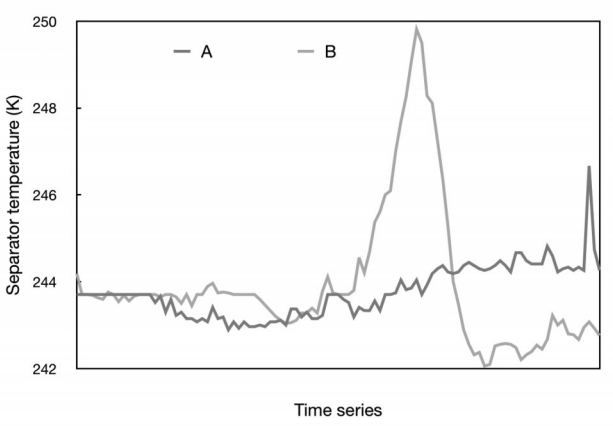

Fig. 14. Operational parameters for separator.

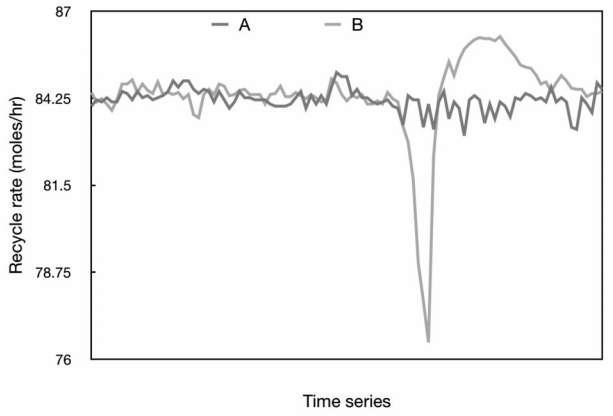

Fig. 15. Recycle rate for ammonia synthesis.

Where x_i is the original data of output, Yi is the mapping reslts from input parameters. $\bar{x}$ is the average of original data, and N is the total number of observation.

The aforementioned evaluation criteria are utilized to evaluate whether the input parameters are selected appropriately for replicating ammonia production. If an input parameter is essential for mapping the input-output relations, the error is increased when the certain parameter is permuted. In Table II, we list 12 types of operational parameters which are conventional measurement points in ammonia plant. In fact, not all these parameters are necessary for mapping inputoutput relations. As shown in Table II, reactor outlet temperature is redundant because no error increase is observed. As omitting other parameters, we observe error increase to different extents. The results show that it is proper to use the rest 11 parameters as the key parameters for mapping inputoutput relations.

In Fig. 16, MSE is utilized to decide if it is proper to use the encoder-decoder architecture for mapping the input-output relations. We use the encoder-decoder architecture to train input dataset under supervised learning. For comparison, we employ several other time series models such as RNN, LSTM and GRU. The encoder-decoder architecture uses LSTM in the first layer as the encoder, and GRU in the second layer as the decoder. As shown in Fig. 16, the encod-

TABLE II
SENSITIVITY ANALYSIS FOR OPERATIONAL PARAMETERS

Without parameter	MSE	MAE(moles/hr)	RMSE (moles/hr)	CV (%))
H_2 feed rate	0.00480	0.04428	0.06928	0.26854
H_2 feed pressure	0.00483	0.04405	0.06950	0.26937
N_2 feed rate	0.00493	0.04669	0.07021	0.27215
N_2 feed pressure	0.00486	0.04438	0.06971	0.27021
Reactor inlet temperature	0.00490	0.04361	0.07000	0.27132
Reactor outlet temperature	0.00477	0.04272	0.06907	0.26769
Reactor inlet pressure	0.00485	0.04613	0.06964	0.26993
Entrance temperature of cooler group	0.00495	0.04364	0.07036	0.27270
Entrance pressure of cooler group	0.00481	0.04348	0.06935	0.26881
Exit temperature of cooler group	0.00525	0.04311	0.07246	0.28084
Separator temperature	0.00509	0.04529	0.07134	0.27653
Recycle rate	0.00503	0.04508	0.07092	0.27489

er-decoder architecture has the best performance. It is more accurate and converges faster than single layer of LSTM and GRU. It is even more accurate than RNN to treat time series problem. In process optimization, we concern more about the ammonia production rate under different operational patterns. The model is acceptable as long as its accuracy does not exceed ammonia production rate. According to Table H and Fig. 17, the listed evaluation criteria are far below ammonia production rate and the comparison result of A-B. Other modeling tool will not change the optimization result. Therefore, the results demonstrate the effectiveness of the encoder-decoder architecture for mapping relation.

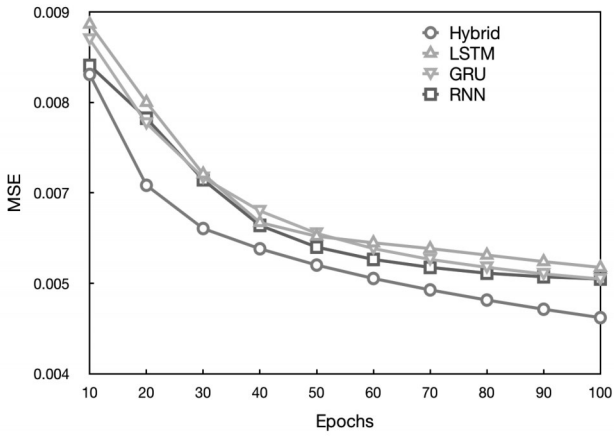

Fig. 16. Model evaluation for mapping the input-output relations.

To decide a better solution from series A and B, we use cross comparison and list the results of A-A, A-B, B-A and B-B in Fig. 17. As shown in Fig. 17, the result of A-A obtains the highest ammonia production rate on time average. B-A and B-B shows the same production rate at the intermediate level. The result of A-B has the lowest production rate of 22.981moles/hr. Therefore, we can conclude that pattern A is the optimal operation. It is discouraged to directly compare the result of A-A with B-B, because chemical processes are incomparable in some cases. For example, a 10-Liter reactor produces more than a 5-Liter reactor, but one cannot say the operation of 10-Liter reactor is better than the 5-Liter reactor. In fact, due to the worse mixing, the 10-liter reactor is usually less efficient. Therefore, the cross comparison is employed to ensure the comparability of two operations.

Compared with model optimization based on chemical laws, the advantage of TL is laid on its efficiency at high dimen-sional task, which is very time consuming for conventional physical or chemical models. Although the TL model is more like a "black box optimization algorithm", it is able to improve the convergence velocity and behavior for academia and industrial application. Additionally, due to the limited series for comparison, the results are tend to be local sub-optimal. There might be better operation than A or worse operation than B, which requires further study in the future work. However, it is impressive that less hydrogen supply but fed with higher pressure can enhance the ammonia production, and 12% promotion of ammonia production rate can be well acceptable for industrial practice.

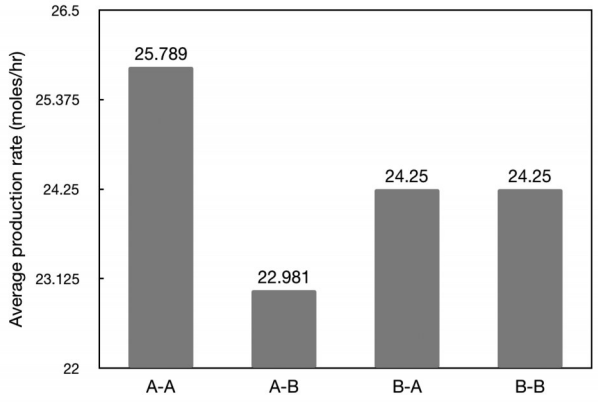

Fig. 17. Comparison of different operational patterns.

5. CONCLUSION

In this paper, we demonstrate that transfer learning can be an optimization tool for chemical process such as ammonia synthesis. The raw data are collected from on-site operation of industrial-scale chemical plant. The comparison group are generated from variational method. Encoder-decoder architecture containing LSTM and GRU layers is adopted to extract operational patterns from time series data based on supervised learning. An improved operation can be filtered through cross algorithm of transfer learning.

The industrial practice of ammonia synthesis brings a challenging problem which is computationally demanding and difficult to be solved using traditional methods. Transfer learning combined with encoder-decoder architecture can replicate such complex process without losing accuracy. The transfer learning models are much faster, the computational time for the TL model is less than 1% of that needed for the rigorous simulator, since the output from the TL is not an iterative solution, and easier to use, once trained. The use of TL model for optimization has the strength that is especially helpful for dynamic process applied on an industrial size plant. However, there is still a great need for further efforts on process models and global optimization algorithms using artificial intelligence architecture.

As a results, the ammonia production rate can be increased by 12% with 22.9% less of hydrogen supply, 12% less of nitrogen supply, and 5bar lower of reaction pressure. The outcome of this work will significantly benefit engineers and

project managers working in the field of clean ammonia synthesis, as it can be used as a reference leading to an improvement of operational strategy under high dimensional uncertainty. Not only for ammonia synthesis, the method proposed in this work has the potential to be applied in other chemical applications, e. g. styrene process, gasoline blending process, carbon dioxide capture process, etc..

ACKNOWLEDGMENT

The authors would like to thank the financial support by the National Science Foundation of the Jiangsu Higher Education Institutions of China (project No. 19KJB530006). The authors are willing to thank the academic support by Dr. Fei-Yue Wang, who is also the Professor and Director of the Key Laboratory of Complex Systems and Intelligence Science, CAS, and Fellow of IEEE, INCOSE, IFAC, ASME, and AAAS.

REFERENCES

[1] Service R. E,"Liquid sunshine—Ammonia made from sun, air, and water could turn Australia into a renewable energy superpower", Science,361(6398), pp. 120-123, 2018.

[2] Inoue Y., Kitano M., Kishida K., Abe H., Niwa Y., Sasase M., Fujita Y., Ishikawa H., Yokoyama T., Hara M., Hosuno H., "Efficient and stable ammonia synthesis by serf-urganized flat Ru nanoparticles on calcium amide", ACS Catalysis, 6, pp. 7577-7584, 2016.

[3] Lu Y., Li J., Tada T., Toda Y., Ueda S., Yokoyama T., Kitano M., Hosono H., "Water durable electride Y5Si3: Electronic structure and catalytic activity fur ammonia synthesis", J. Ara. Chem. Soc., 138, pp. 3970-3973, 2016.

[4] Takuya N., Tomofumi T., Hosono H., "Transition metal-doped Runanoparticles loaded on metal hydrides fur efficient ammonia synthesis from first principles", Journal of Physical Chemistry C, 124(2), pp.1529-1534, 2020.

[5] Wagner K., Zhu M., Malmali M., Smith C., McCormick A., Cnssler E. L., Seaton N. C. A., "Column absorption for reproducible cyclic separation in small scale ammonia synthesis", AIChE J., 63(7), pp. 3058-3068, 2017.

[6] Pan X., Zhu M., Mei H., Liu Z., Sben T., "Ammonia absorption enhancement by metal halide impregnated hollow mesoporons silica spheres",ChemistrySelect, 5, pp. 5720-5725, 2020.

[7] Malmali M., Wei Y., McCormick A., Cussler E. L., "Ammonia synthesis at reduced pressure via reactive separation", Ind. Eng. Chem. Res., 55, pp. 8922-8932, 2016.

[8] Iwamoto M., Aldyama M., Aihara K., Degnchi T.,"Ammonia synthesis on wool-like An, Pt, Pd, Ag, or Cu electrode catalysts in nonthermal atmosphuric-pressure plasma of N_2 and H_2", ACS Catalysis, 7, pp. 6924-6929, 2017.

[9] Sipocz N., Tobiesen E A., Assadi M., "The use of artificial neural network models for CO_2 capture plants", Applied Energy, 88, pp. 2368-2376, 2011.

[10] Wang X., Hun D., Lin Y., Du W., "Recent progress and challenges in process optimization: Review of recent work at ECUST", The Canadian Journal of Chemical Engineering, 96, pp. 2115-2123, 2018.

[11] Vankatasubramanian V.,"The promise of artificial intelligence in Chemical Engineering: Is it here, finally?" AIChE J., 65(2), pp. 466-478, 2019.

[12] Gupta A., "Introduction to deep learning: Pm 1", Chem. Eng. Prog., pp. 22-29, 2018.

[13] Bakshi B. R.,Stephanopoulos G.,"Wave-net:A muitiresolntion, hierarchical neural network with localized learning", AIChE J., 39, pp. 57-81, 1993.

[14] CS231n, "Convolutional neural networks for visual recognition", Available at: http://cs231n. github. io/convolntional-netwurks/, 2018.

[15] Gupta A., "Introduction to deep learning: Part 2", Chem. Eng. Prog., pp. 39-46, 2018.

[16] Lin Y., Yao L., Xin Z., "Geographical discrimination and adulteration analysis for edible oils using two-dimensional correlation spectroscopy and convolutional neural networks (CNNs)", Spectrechimica acta. Part A, Molecular and bimolecular spectroscopy, 246, Art. no. 118973, 2021.

[17] Schmidhnber J.,"Deep learning in neural netwurks: An overview", Neural Netw., 61, pp. 85-117, 2015.

[18] Shi H., Xn M., Li R., "Deep learning for household load fur forecastinga novel pooling deep RNN", IEEE Trans. Smart Grid, 9(5), pp. 5271-5280, 2018.

[19] Almalaq A., Zhang I. J., "Evolutionary deep learning-based energy consumption prediction for buildings", IEEE Access, 7, pp. 1520-1531, 2019.

[20] Almalaq A., Has J., Zhang J. J., Wang E Y.,"Parallel building: A complex system approach for smart building energy management", IEEE/CAA Journal of Antomatica Smica, 6(6), pp. 1452-1461, 2019.

[21] Kumar S., Hussain L., Bans0ce S., Reza M., "Energy load furecasting using deep learning approach-LSTM and GRU in spark cluster", In Proc. 5th Int. Conf. Emerging Applications of Information Technology (EAIT), pp. 1-4, 2018.

[22] Skymiad, "A beginner's gnide to deep reinfomement learning", Available at:http://skymind.ai/wiki/dcep-minforcement-leamiag.

[23] Blitzer J., Dredze M., Pereira F.,"Biographies, Bollywood, boomboxes and blenders: Domain adaptation fur sentiment classification", In Proccedings of the 45th Annual Meeting of the Association of Computational Linguistics, pp. 440-447, 2007.

[24] De. Cost B. L., Francis T., Holm E. A., "Exploring the mierostmctore manifold: Image texture representations applied to ultrahigh carbon steel microstmcteres", Acta Mater., 133, pp. 3040, 2017.

[25] Wu S., Kundo Y., Kakimoto M. A., Yang B., Yamade H., Kuwajima L., Lampard G., Hongo K., Xu Y., Shiomi J., Schick C., Morikawa J.,Yoshida R., "Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm", npj Comput. Mater., 5, Art. no. 66, 2019.

[26] Yamade H., Liu C., Wu S., Koyama Y., Ju S., Shiomi J., Morikawa J.,Yoshida R.,"Predicting materials properties with little data using shotgun transfer learning", ACS Central Science, 5, pp. 1717-1730, 2019.

[27] Ma R., Colon Y. J., Lno T., "Transfer learning study of gas adsorption in Metal-Organic-Framewurks", ACS Appl. Mater. Interfaces, 12, pp. 34041-34048, 2020.

[28] Sun Y., DeJaco R. F., Siepmann J. I.,"Deep neural netwurk learning of complex binary sorpfion equilibria from molecular simulation data",Chem. Sci., 10, pp. 4377-4388, 2019.

[29] Cheng E, He Q. P., Zhao J., "A novel process mouimring approach based on variational recurrent autoencoder", Computers and Chemical Engineering, 129, Art. no. 106515, 2019.

[30] Hara M., Kitano M., Hosono H., "Ru loaded alumina as a catalyst for ammonia synthesis", ACS Catalysis, 7, pp. 2313-2324, 2017.

Ming Zhu Doctor (Tianjing University), and Project manager (Nanjing Tech University). Visiting Scholar in University of Minnesota, twin cities, United States. Research focuses on process intensification, mass transfer, separation and purification teclmologies. Project manager of National Science Foundation of the Jiangsu Higher Education Institutions of China. Core member of projects of National Science Foundation of China, National Basic Research Program of China, and Minnesota Environment and National Resources Trust Fund. Research works were published on “Chemical Engineering Journal”, “AIChE Journal”, “Industrial Engineeaing and Chemistry Research”, “Journal of Chemical Engineering and Industry (China)” and “Chinese Journal of Environmental Engineering”, etc.. Reviewer of "Chemistry Select", a Wiley Europe publication.