Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Surrogate Modelling and Optimization for Complex Liquefied Natural Gas Refrigeration Cycles

Thomas Savage [*], **Fernando Almeida-Trasvina** [*], **Antonio Del-Rio Chanona** [**], **Robin Smith** [*], **Dongda Zhang** [*,**]

[*] *Centre for Process Integration, University of Manchester, Manchester, M1 3AL, UK.*
[**] *Centre for Process Systems Engineering, Imperial College London, London, SW7 2AZ, UK.*

Abstract: In this paper, surrogate modelling and optimization is investigated for use in large scale chemical processes. A novel CryoMan cascade liquefied natural gas (LNG) refrigeration cycle is selected as the case study which has been highlighted for potential use within industry. Given its high nonlinearity and dimensionality (31 input variables and 20 output variables with a number of physical constraints) and short time horizon for real-time decision-making, an time-efficient optimization scheme must be developed to maximize process performance. Therefore, various supervised and unsupervised learning techniques as well as surrogate model structures are explored in order to accurately capture the behaviour of this highly complex and interrelated process flowsheet. Optimal solutions identified by the surrogate models are validated against the rigorous process model. Following from the challenges encountered by artificial neural network based surrogate models, Gaussian processes were adopted and combined with partial least squares to simultaneously reduce dimensionality and capture the nonlinearity of the underlying chemical process. Through this innovative surrogate modelling strategy, overall time to optimize the LNG production process was reduced by orders of magnitude compared to the rigorous model based optimization methodology, hence significantly facilitating the industrial application of this new process.

Keywords: chemical engineering, large scale optimization problem, efficient strategies for large scale complex systems, machine learning, model reduction

1. INTRODUCTION

Developing disruptive digital technology to enable the design and operation of cost-effective and energy-efficient manufacturing systems is one of the grand research themes under the context of the 4$^{\text{th}}$ Industrial Revolution. Given the large amount of data accumulated from process industries, building data-driven models to enable rapid decision-making is of critical importance in order to guarantee the process performance and safety (Bhosekar and Ierapetritou, 2018).

Specific to the chemical industry, rigorous process models derived from mass and energy balances allow for accurate determination of states of a system, and can be directly constructed using multiple computer-aided software packages such as Aspen (Bhosekar and Ierapetritou, 2018). However, due to their complexity, the computational time to evaluate a rigorous model can be relatively large (e.g. potentially weeks), hence limiting their applications for process control and optimization (McBride and Sundmacher, 2019). As a result, surrogate models have been adopted to replace rigorous models and help with identifying optimal operating conditions.

At this moment, surrogate models are predominantly used to replace single unit operations (Henao and Maravelias, 2011; Quirante et al., 2015), whilst their applicability in terms of substituting more complex systems such as entire process flowsheets has not been well explored. Therefore, in this study, different cutting-edge surrogate models are tested to simulate and optimize a novel liquefied natural gas (LNG) refrigeration cycle, with their performance thoroughly compared against the optimal solution identified by the rigorous model. The structure of this paper is organized as follows. Section 2 introduces the investigated LNG production process. The approach to optimize the rigorous model and the associated primary challenges are also explained. Section 3 details the construction and optimization of different surrogate models. Section 4 summarizes the results of these surrogate models and compares them with the rigorous model's verification. A thorough discussion regarding the advantages of combining Gaussian processes and partial least squares for large scale complex chemical process simulation and optimization is presented in the Results and Discussion section.

2. INTRODUCTION TO THE CRYOMAN CYCLE

Commercial scale production of LNG involves the use of large, complex and energy-intensive refrigeration cycles. The costs associated with the energy for refrigerant compression (shaft work energy) dominate the overall operating costs of the LNG plant.

The CryoMan Cascade cycle recently developed by Almeida-Trasvina and Smith (2018) (shown in Figure 1) is a novel refrigeration configuration option, potentially able to bring significant savings in shaft work demand compared to current processes in the LNG industry. In this process, within the precooling cycle, a 'heavy' mixed refrigerant provides cooling in a series of two multi-stream heat exchangers (MSHEs); three stages are employed for refrigerant compression. In the liquefaction cycle, the 'light' mixed refrigerant is first fed into a liquid-vapour separator. The resulting outlet streams are partially mixed with each other to create the two refrigerant streams that provide cooling in a series of two MSHEs.

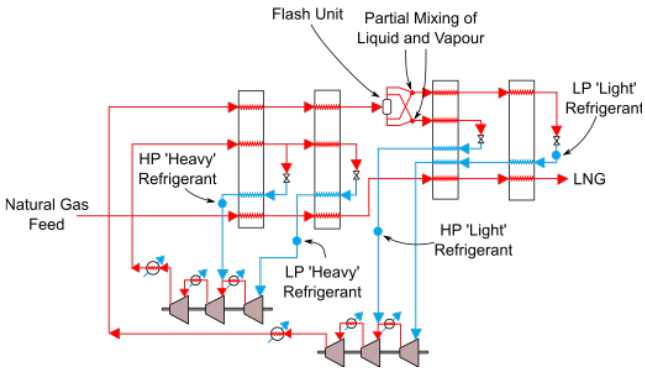

Figure 1. The CryoMan Cascade LNG refrigeration cycle.

2.1 Implementation of rigorous model

The models employed for the MSHE units are based on energy balances; feasible heat transfer is assessed given a practical value for the minimum approach temperature between hot and cold composite curves across the length of each MSHE. The compressors and pumps are modelled isentropically.

The rigorous model of the CryoMan Cascade cycle was implemented in Aspen HYSYS v8.2 and subsequently linked to MATLAB. The inputs include the refrigerant mass flowrates, refrigerant compositions, discharge pressure, refrigerant evaporating pressures, MSHE outlet temperatures, compression ratios and refrigerant split fractions (where applicable) for both precooling and liquefaction cycles. The outputs from the process are eight values of shaft work demand, four MSHE approach temperatures, vapour fractions of four streams (to assess wetness at inlet of compressors), and four compression ratios.

2.2 Optimization of the rigorous model

The optimization problem for the rigorous model is defined as follows:

$$min \left(\sum_{i=1}^{N} W_i \right) / m_{LNG} \tag{1}$$

$$s.t. \quad \Delta T_{min} \geq 2^o C \tag{2}$$

$$P_{rat} \leq 3.5 \tag{3}$$

$$\sum_{j=1}^{m} x_j = 1 \quad x_j \in X^{MR} \tag{4}$$

$$VF^{ref} = 1.00 \tag{5}$$

$$lb \leq \phi \leq ub \tag{6}$$

The specific shaft work, defined as the sum of the individual shaft works W_i divided by the mass flow rate of LNG m_{LNG}, is the objective function to be minimized. The 31 inputs to the rigorous model are represented by ϕ with the set of inputs having corresponding upper and lower bounds ub and lb, respectively. Constraints include minimum approach temperatures for MSHEs ΔT_{min}, maximum compression ratios P_{rat} to discourage mechanical damage to compressors, valid molar compositions represented by x_j, and no wetness within compressors by constraining vapour fractions at the inlet of compressors VF^{ref} to zero.

An evolutionary algorithm (EA) is first employed for the optimization of the rigorous model. The way in which constraints are dealt with in stochastic optimization is by applying penalties to invalid solutions. Successive Quadratic Programming (SQP) is next used to identify a local optimal solution around the best candidate resulting from the stochastic optimization.

3. DEVELOPMENT OF SURROGATE MODELS

Directly optimizing the rigorous model is time consuming (over 17 hours per run). This approach is not extremely computationally expensive, however, the real-time decision-making of a commercial LNG plant is often around once per 4 hours. Thus, using the Aspen based rigorous model cannot fulfill this practical constraint. As a result, surrogate models are adopted as a competent alternative to resolve this challenge. More importantly, if different configurations of LNG cycles are to be considered in parallel for operational optimization, being able to reduce the optimization time for each configuration will allow for more rapid evaluation of different case studies and structural modifications, and will lead to a reduced overall project execution time.

Different from using surrogate models to substitute a unit operation, building a surrogate model to replace an entire process flowsheet is more challenging due to the high nonlinearity of the underlying process and high dimensionality of the involved design variables. Therefore, in order to guarantee success, a number of surrogate model structures were proposed with different hypothesise and their performance was thoroughly compared in this work. Furthermore, taking advantage of both supervised (dealing with nonlinearity) and unsupervised (dealing with dimensionality reduction) machine learning techniques is another strategy proposed in this work, as this may greatly simplify the process complexity for surrogate model construction whilst obtaining high quality optimal solutions.

3.1 Construction of Artificial Neural Network based surrogate models

Artificial neural networks have been widely used as surrogate models to replace rigorous models for unit operation simulation (e.g. distillation columns) (McBride and Sundmacher, 2019; Abdul Jaleel and Aparna, 2019). ANNs provide a flexible structure to simulate nonlinear systems when given sufficient amounts of data. When an ANN is used for regression, its training procedure is to minimize the difference between model prediction and the given output, by optimizing the weights (w_{ij}) and biases (b_j) that connect the neurons between different layers.

An ANN consists of a number of hidden layers which can be connected through different manners. As this work does not involve image-reading or time-events, a feedforward neural network is selected. To construct an accurate feedforward neural network, choosing a suitable activation function is an important step. There exists many possible activation functions such as the *sigmoid*, *tanh* (Eq. 7), and *ReLu* functions, all of which have their situational benefits. As the current ANN falls into the category of regression, a *tanh* (hyperbolic tangent) activation function is preferred given its heuristic advantages. Training data was standardized through the standard procedure to take advantage of the nonlinear region within the activation function, and 20% of the data was reserved for model cross-validation to prevent overfitting (del Rio-Chanona et al., 2017).

$$y_j = tanh\left(\sum_i x_i \cdot w_{ij} + b_j\right) \tag{7}$$

To take into account the nonlinearity of this highly interconnected process flowsheet, different structures of ANN based surrogate models are constructed as shown in Fig. **??** and 3. These include: (1) a single ANN directly simulating the entire process (31 inputs and 20 outputs); (2) an ANN framework comprising 2 independent ANNs, one simulating the 8 shaft works, 4 compression ratios and 4 vapour fractions given the 31 inputs, and the other simulating the 4 temperature differences given the same inputs; (3) an ANN framework comprising 5 separate ANNs, one simulating the 8 shaft works, 4 compression ratios and 4 vapour fractions, and the other 4 each of which only simulating a specific temperature difference. The reasoning behind constructing different surrogate model structures is discussed in the Results and Discussion section. It is worth mentioning that the current ANNs were constructed through a thorough hyperparameter selection framework and a k-fold cross validation method. The optimal combination of essential key hyperparameters was identified to balance the model training and testing performance. A detailed explanation of the ANN model construction procedure is demonstrated in our recent publication (del Rio-Chanona et al. (2017)).

Figure 2. ANN based surrogate model structure 1.

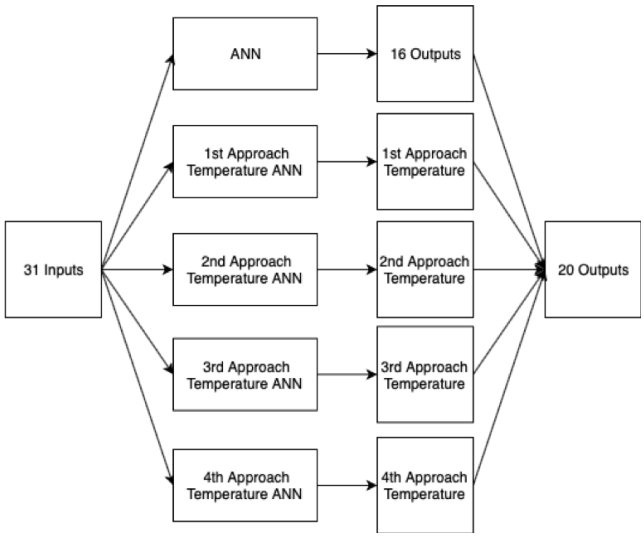

Figure 3. ANN based surrogate model structure 3.

3.2 Construction of Gaussian process based surrogate models

In contrast to an ANN which is a regression model, a Gaussian process (GP) is an interpolation model developed by Krige (1951) and popularized by the machine learning community since mid 2000s (Rasmussen and Williams, 2006). GPs aim to describe an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ using noisy observations, $y = f(\mathbf{u}) + \nu$, where $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is Gaussian distributed measurement noise with zero mean and a unknown variance σ_ν^2.

GPs consider a distribution over functions, and they can be seen as a generalization of multivariate Gaussian distributions,

$$f(\cdot) \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$$

where the mean function $m(\cdot)$ can be interpreted as the deterministic part of the function; and the covariance function $k(\cdot, \cdot)$ accounts for correlations between the function values at different points.

The focus therefore is on a constant mean function, $m(\mathbf{u}) := c$, along with the squared-exponential (SE) covariance function (Rasmussen and Williams, 2006) based on the assumption that the real system function f is smooth and stationary:

$$k(\mathbf{u}, \mathbf{u}') := \sigma_n^2 \exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{u}')^\mathsf{T} \mathbf{\Lambda} (\mathbf{u} - \mathbf{u}')\right)$$

where σ_n^2 is the covariance magnitude; and $\mathbf{\Lambda} := \text{diag}(\lambda_1 \cdots \lambda_{n_u})$ is a scaling matrix.

Maximum likelihood estimation is usually used to estimate the unknown hyperparameters $\mathbf{\Psi} := [c\ \sigma_n\ \sigma_\nu\ \lambda_1 \ldots \lambda_{n_u}]^\mathsf{T}$,

including σ_ν in case the measurement noise variance is also unknown. Other distinct characteristics of GPs compared to ANNs include: (1) GPs are Multiple Input Single Output (MISO) model whilst an ANN is a Multiple Input Multiple Output (MIMO) model; (2) GP is particularly efficient when learning from small sizes of data whilst ANN thrives from large sizes of data (Bradford et al., 2018). To investigate the efficiency of GPs for complex system simulation, in this work, a GP based surrogate model superstructure is constructed comprising of 7 independent GPs, each of which predicting a specific output including the 4 temperature differences, the total shaft work, 1 specific compression ratio and 1 specific vapour fraction. This superstructure also embeds an unsupervised learning algorithm - partial least squares (PLS) for dimensionality reduction. The detailed explanation of this superstructure is explained in the next section.

3.3 Introduction to Kriging Partial Least Squares

Partial Least Squares (PLS) is an unsupervised learning algorithm that seeks to find the multidimensional direction in input space $\mathbf{X}$ that minimizes the variance in the output dimension $\mathbf{Y}$ (Guido and Mueller, 2016). PLS allows for the reduction of input dimension whilst maintaining the majority of meaningful information within the data, by projecting both $\mathbf{X}$ and $\mathbf{Y}$ to new smaller dimensional spaces. In the method described by Bouhlel et al. (2016) to integrate PLS into GPs, PLS is performed over the data and the PLS weights (projected spaces of inputs and outputs) are used to construct a lower dimensional covariance matrix $\mathbf{K}(U)$ over which a Gaussian process model is constructed. By reducing the model dimensionality it allows for easier determination of the minimum log-likelihood $\mathcal{L}(\mathbf{\Psi})$ resulting in a lower computational time cost, and more accurate mapping of between the input space and the output space for large dimensional systems.

This integrated modelling strategy (embedding PLS into GP), namely KPLS, is adopted in this research to complete the construction of the GP based surrogate model superstructure mentioned in Section 3.2. The structure of this surrogate model is shown below in Figure 4.

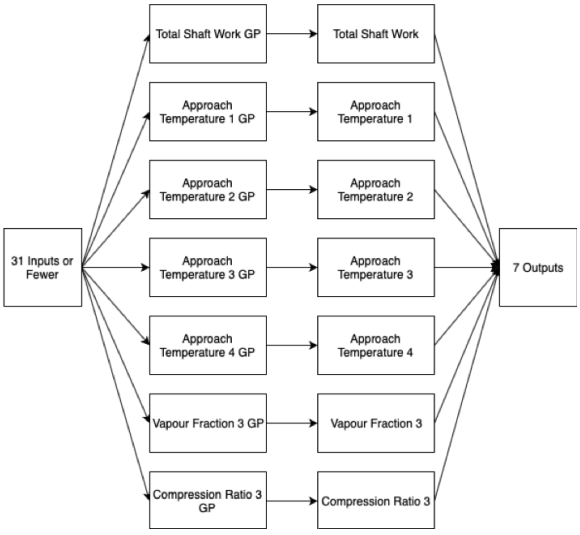

Figure 4. KPLS surrogate model superstructure with respect to the 7 independent GPs.

3.4 Data generation and selection

Three different datasets were generated with differing qualities and time costs in order to create the surrogate models:

- Dataset 1: 5000 random, bounded data points. Taking 10 minutes to generate, however are mostly infeasible.
- Dataset 2: 500 refined data points, generated within slightly restricted bounds.
- Dataset 3: 350 high-quality data points. Consisting of valid solutions after discarding invalid solutions, with an associated high time cost ($\tilde{3}$ hours).

As the accuracy of a surrogate model heavily relies on the data quality, Dataset 1 was found incapable of providing useful information for data-driven model construction, hence was discarded. However, having a sufficient amount of data is also a prerequisite to build an accurate surrogate model, thus purely using Dataset 3 is not able to meet this requirement. Therefore, Datasets 2 and 3 were combined in order to provide global representation of the rigorous model as well maintain a relatively large proportion of valid solutions, such that the surrogate model would be sufficiently accurate in promising valid areas.

For a highly interconnected system with 31 inputs and 20 outputs, 850 data points is not necessarily large. As a result, it is difficult to pre-determine which machine learning technique, ANN or GP, is more suited in this task. Although extra high quality data can be generated, the overall surrogate model construction time will be increased, diminishing its advantage over the rigorous model. The approach of limiting the data used to create a surrogate model may seem counter intuitive, however it places the emphasis on efficient and smart model creation with respect to the rigorous model, which can then be supplemented with additional data after construction.

3.5 Optimization of surrogate models

The optimization problem for surrogate models remains the same for the rigorous model. Specific to data-driven based surrogate models, stochastic optimization is the standard approach because of its ability to efficiently explore global search spaces, particularly when the function being optimized is highly nonlinear. Another advantage of stochastic optimization is that it does not rely on the derivative of the function being optimized, meaning that it is the only option able to optimize models where information about the gradient is unavailable or difficult to deduce (e.g. large artificial neural networks).

An evolutionary algorithm was employed to optimize the surrogate models, taking advantage of tournament selection and single point crossover. A mutation rate of 3% and selection percentage of 98% were used over a population size of 100. The optimization scheme was chosen to run for 500 generations. Whilst simply applying a constant penalty for any violation is feasible, the nature of the solution space specific to this case study results in mostly infeasible areas. Therefore, a linear penalty is implemented in order to guide the optimization initially towards a valid solution space. This linear penalty also contains a constant term in order to mathematically remain greater than the objective function for all invalid solutions.

Finally, the ANN based surrogate models were implemented using PyTorch 1.2.0 and the KPLS based surrogate models were implemented using the SMT toolbox by Bouhlel et al. (2019), both within Python 3.7.3 on a Windows 10 operating system. Total construction of the KPLS models takes on average 2 minutes using 850 data points, and that for the ANN based surrogate models takes approximately 5 minutes.

4. RESULTS AND DISCUSSION

4.1 Results of the single ANN surrogate model

Table 1. Results of the single ANN surrogate model. The optimal solution predicted by this surrogate model is validated against the rigorous model.

	SW_1	SW_2	SW_3	SW_4
	(MW)			
Prediction	12.26	21.43	9.93	0.38
Validation	10.56	21.14	11.94	0.28
	SW_5	SW_6	SW_7	SW_8
	(MW)			
Prediction	16.13	37.80	45.93	0.00
Validation	15.2	35.82	50.15	0.00
	VF_1	VF_2	VF_3	VF_4
	(-)			
Prediction	1.00	1.00	1.00	1.00
Validation	1.00	1.00	0.97	1.00
	P_1^{rat}	P_2^{rat}	P_3^{rat}	P_4^{rat}
	(-)			
Prediction	2.07	2.89	1.56	2.61
Validation	2.12	2.70	2.25	1.80
	ΔT_{min1}	ΔT_{min2}	ΔT_{min3}	ΔT_{min4}
	(K)			
Prediction	3.24	2.43	12.80	2.32
Validation	-0.04	-5.35	2.82	1.19

First of all, results of the single ANN (31 inputs, 20 outputs) are summarized in Table 1. The *prediction* results show the expected output as determined from the ANN model, and the *validation* results show the actual output when this 'optimal' solution is rigorously simulated. From the table, it is observed that this model has encouraging results with respect to predicting the individual shaft works. However, as shown in Figure 5, after validation with the rigorous model, it is found that MSHE approach temperatures (i.e. practical temperature constraints) are infeasible for the ANN predicted optimal solution. The single large ANN is here unable to capture the nonlinearity of the process, and its accuracy is verified to be low.

Fig. 5 shows the validated ΔT approach temperatures when results from the respective data-driven models are validated on the rigorous model. The two most simple ANN structures are shown to be unable to capture the nonlinear nature of the approach temperatures and both produce false optimal solutions that are infeasible ($< 2^{o}$C, below the red line).

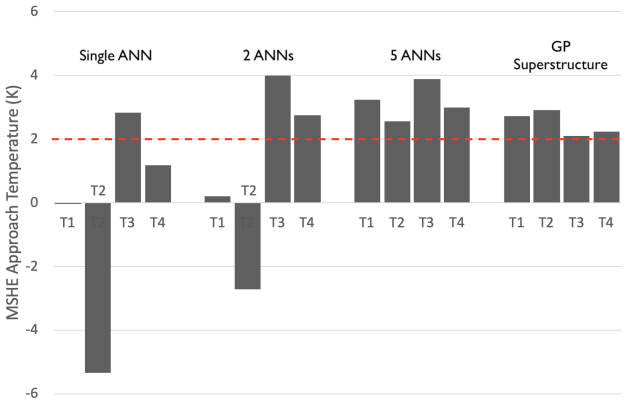

Figure 5. Plot of validated approach temperatures for 4 different surrogate model structures. Infeasible MSHE approach temperatures have values below the dotted red line.

4.2 Results of the ANN based surrogate model frameworks

A straightforward improvement to the single ANN surrogate model is to split this large ANN into a two separate ANNs to deal with the nonlinear temperature constraints. In other words, a surrogate model framework can be built to include two independent ANNs, one only predicting the four highly nonlinear MSHE approach temperatures and the other estimating the remaining sixteen outputs. It should be noted that whilst these approach temperatures do not directly contribute to the specific shaft work being optimized, they act as highly important practical constraints within the optimization problem and thus need to be accurately predicted.

Breaking a surrogate model up into a series of parallel sub-models may increase accuracy in capturing specific nonlinearities that are particularly prominent in a highly interrelated process flowsheet. Based on this hypothesis, a second surrogate model framework is constructed to further separate the MSHE approach temperature ANN into four separate ANNs, each trained to predict an individual approach temperature. This will allow the overall surrogate model to gain better accuracy still with regards to the individual approach temperatures, and ensure that a prediction made using the surrogate model remains feasible when validated with the rigorous model.

Indeed, through rigorous model verification, both of the ANN frameworks can well predict the shaft works and vapour fractions. From Figure 4, it is seen that although the first ANN framework (consisting of 2 ANNs) still violates the temperature constraints, the second framework (consisting of 5 ANNs) successfully meets all the constraints, hence resulting in feasible solutions. Nonetheless, the optimal total shaft work identified in the third surrogate model is 160 MW, much higher than the optimal solution (144 MW) identified using the rigorous model. Hence, the surrogate model is still not efficient in terms of process optimization. Overall, the results from the three ANN based surrogate models strongly suggest that due to the high nonlinearity and dimensionality of the underlying process and limited amount of data, it may not be an ideal approach using ANNs to directly construct surrogate models to simulate a whole process flowsheet.

Table 2. Results of the KPLS surrogate model superstructure. The optimal solution predicted by this surrogate model is validated against the rigorous model.

	SW_{total}	VF_3	P_3^{rat}	
	(MW)	(-)	(-)	
Prediction	146.3	0.994	3.01	
Validation	147.4	0.983	3.05	
	ΔT_{min1}	ΔT_{min2}	ΔT_{min3}	ΔT_{min4}
	(K)			
Prediction	2.60	3.99	2.52	2.82
Validation	2.73	2.92	2.10	2.23

4.3 Results of the KPLS superstructure model

Due to the system in question having a relatively large number of inputs and outputs, unsupervised learning techniques are adopted in order to reduce the dimension of the solution space (i.e. reducing the impact of the curse of dimensionality) for the construction and optimization of surrogate models. As there exists no constraint on each individual shaft work, these outputs are grouped into a single output representing the sum of the shaft works, simplifying the structure of the surrogate model. Likewise with reducing the shaft works into a total shaft work, it was observed from the ANN based surrogate models that the majority of data regarding 3 out of the 4 vapour fractions and 3 out of the 4 shaft works exist in feasible solution space all the time (i.e. redundant constraints). Due to this fact, it is decided to neglect to model these outputs in the first place to further reduce the dimension of the outputs.

The results of the KPLS superstructure model (consisting of 7 GPs) is presented in Table 2. From the table, it can be concluded that the total shaft work (147.4 MW) has been predicted well, and is comparable to the optimal result obtained using the rigorous optimization (144.4 MW). In addition, individual shaft works are found to be feasible after verification. All the constraints including vapour fractions and MSHE temperatures are satisfied. Most importantly, the full time to train and optimize this surrogate model only takes around 10 minutes (i.e. 2 minutes for model construction and 8 minutes for stochastic optimization) as opposed to 17 hours spent when optimizing the rigorous model. This directly demonstrates the superiority and practical advantages of using the integrated KPLS modelling strategy for the optimization and real-time decision-making of high dimensional complex chemical systems (e.g. entire process flowsheet).

5. CONCLUSIONS

To conclude, Gaussian processes and artificial neural networks can be used as building blocks in parallel to construct time-efficient surrogate models to represent highly nonlinear systems such as a highly interacted chemical process flowsheet. However, selection of surrogate model structure, fidelity of available data, and amount of data can greatly affect the accuracy and efficiency of surrogate models. To guarantee the success of surrogate model construction, knowledge of the real world system is required. Moreover, unsupervised learning techniques can also be taken advantage of in order to reduce the high dimensionality encountered in large scale processes simulation. Through the use of data-driven models and stochastic optimization algorithms, it is possible to reduce the computational time cost and meanwhile identify a high quality optimal solution for the operation of complex systems.

REFERENCES

Abdul Jaleel, E. and Aparna, K. (2019). Identification of realistic distillation column using hybrid particle swarm optimization and NARX based artificial neural network. *Evolving Systems*, 10(2), 149–166.

Almeida-Trasvina, F. and Smith, R. (2018). Design and optimisation of novel cascade refrigeration cycles for lng production. In A. Friedl, J.J. Klemeš, S. Radl, P.S. Varbanov, and T. Wallek (eds.), *28th European Symposium on Computer Aided Process Engineering*, volume 43 of *Computer Aided Chemical Engineering*, 621 – 626. Elsevier.

Bhosekar, A. and Ierapetritou, M. (2018). Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers & Chemical Engineering*, 108, 250–267.

Bouhlel, M., Bartoli, N., Morlier, J., and Otsmane, A. (2016). Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction. *Structural and Multidisciplinary Optimization*, 53, 935–952.

Bouhlel, M.A., Hwang, J.T., Bartoli, N., Lafage, R., Morlier, J., and Martins, J.R.R.A. (2019). A python surrogate modeling framework with derivatives. *Advances in Engineering Software*, 102662.

Bradford, E., Schweidtmann, A.M., Zhang, D., Jing, K., and del Rio-Chanona, E.A. (2018). Dynamic modeling and optimization of sustainable algal production with uncertainty using multivariate Gaussian processes. *Computers & Chemical Engineering*, 118, 143–158.

del Rio-Chanona, E.A., Fiorelli, F., Zhang, D., rashid Ahmed, N., Jing, K., and Shah, N. (2017). An efficient model construction strategy to simulate microalgal lutein photo-production dynamic process. *Biotechnology and Bioengineering*, 114(11), 2518–2527.

Guido, S. and Mueller, A.C. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media.

Henao, C.A. and Maravelias, C.T. (2011). Surrogate-based superstructure optimization framework. *AIChE Journal*, 57(5), 1216–1232.

Krige, D.G. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6), 119–139.

McBride, K. and Sundmacher, K. (2019). Overview of Surrogate Modeling in Chemical Process Engineering. *Chemie Ingenieur Technik*, 91(3), 228–239.

Quirante, N., Javaloyes, J., and Caballero, J.A. (2015). Rigorous design of distillation columns using surrogate models based on Kriging interpolation. *AIChE Journal*, 61(7), 2169–2187.

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.